

Jerzy Leszczynski

Editor

Anna Kaczmarek-Kedziera

Tomasz Puzyn

Manthos G. Papadopoulos

Heribert Reis

Manoj K. Shukla

Co-Editors

SPRINGER
REFERENCE

Handbook of Computational Chemistry

Second Edition

 Springer

Handbook of Computational Chemistry

Jerzy Leszczynski

Editor

Anna Kaczmarek-Kedziera • Tomasz Puzyn •

Manthos G. Papadopoulos • Heribert Reis •

Manoj K. Shukla

Co-Editors

Handbook of Computational Chemistry

Second Edition

With 588 Figures and 123 Tables



Springer

Editor

Jerzy Leszczynski
Department of Chemistry and Biochemistry
Interdisciplinary Center for Nanotoxicity
Jackson State University
Jackson, MS, USA

Co-Editors

Anna Kaczmarek-Kedziera
Faculty of Chemistry
Nicolaus Copernicus University
Toruń, Poland

Tomasz Puzyn
Laboratory of Environmental Chemometrics
Faculty of Chemistry
University of Gdańsk
Gdańsk, Poland

Manthos G. Papadopoulos
Institute of Biology
Medicinal Chemistry and Biotechnology
National Hellenic Research Foundation
Athens, Greece

Heribert Reis
Institute of Biology
Medicinal Chemistry and Biotechnology
National Hellenic Research Foundation
Athens, Greece

Manoj K. Shukla
US Army Engineer Research
and Development Center
Vicksburg, MS, USA

ISBN 978-3-319-27281-8 ISBN 978-3-319-27282-5 (eBook)
ISBN 978-3-319-28572-6 (print and electronic bundle)
DOI 10.1007/978-3-319-27282-5

Library of Congress Control Number: 2011941760

1st edition: © Springer Science+Business Media B.V. 2012

© Springer International Publishing Switzerland 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

We are delighted to deliver a second, expanded, and brought up to date edition of the *Handbook of Computational Chemistry*. It is apparent that this fast-growing research area requires constant updates. The second edition of our book provides the reader with current, key references to both methodologies and the results of the computational chemistry studies. We selected various focus areas to be covered by the handbook, and this selection is a subjective choice of the editors since due to a vast number of studies published every year, only the most representative examples of applications could be incorporated in the five volumes of this edition.

An amazing transformation of chemistry from an exclusively experimental science to one with theoretical foundations started in the early twentieth century, when principles of quantum mechanics had been established. The most notable progress has been accomplished during the last 50 years during which time it has advanced through noticeable developments and applications of computational methods. This transformation has its roots in both theoretical breakthroughs (Heisenberg's 1925 and Schrodinger's 1926 first papers on quantum mechanics) and the development of the first computer – Electronic Numerical Integrator and Computer (ENIAC) – built in 1943 for the US Department of the Army. However, it took about 20 years after the creation of the ENIAC computers to deliver computational chemistry techniques to the scientific community. Such early methods, being semiempirical in nature, do rely on the large number of experimental parameters. Disappointedly, though such an approach unites experimental chemistry origins with new theoretical approaches, it proves to produce in various cases artificial computational results, and it also lacks reliable parameters for some elements. The next successful chapter for computational chemistry started in the 1970s with applications of nonempirical ab initio methods. The first ab initio computer code popular among non theoreticians – GAUSSIAN 70 – was developed in the 1970s by John A. Pople's group. It initiated a computational chemistry revolution that fired up in the 1980s when supercomputers became accessible to the general scientific community. Also during this period, density functional theory approaches gained a prominent position among efficient computational methods. The vital role of computational chemistry in many research areas was convincingly acknowledged in 1998 when the chemical and physical community celebrated the Nobel Prize that was awarded to two leading

computational experts. It was not by chance that Walter Kohn and John A. Pople were recognized by the Nobel Committee for their contributions to the development of efficient computational methods for quantum chemistry.

Over the past years, different methods of theoretical chemistry were successfully transformed into useful tools that could be applied in diverse areas of research, science, and technology. Owing to the fundamental methodological developments and continuous impressive progress in computational technologies in the last quarter of the twentieth century, several fast and user-friendly programs were developed and made accessible to a wide community of scientists. These near-automatic computer codes are evidence of how methods of theoretical chemistry could be applied as both diagnostic and predictive research tools which support efforts or even guide the direction of traditional experimental approaches. The user-friendly features of commercial codes combined with efficient visualization methods make them accessible to researchers with almost no educational background. Generally, only a limited knowledge and a little formal experience in theoretical chemistry are required to use such programs and to obtain some numerical data. This might cause unexpected outcomes. Most of the methods of theoretical (quantum) chemistry have a limited range of applications, and their use by an inexperienced amateur who performs computational studies without understanding such constraint may lead to serious problems and mistakes. Moreover, a variety of methods of different qualities exist, and a user without a good background in theoretical chemistry may feel lost at the choices from usually rich menus of available programs.

The multivolume *Handbook of Computational Chemistry* is primarily intended as a guide to help navigate among different computational methods currently in use. In order to accomplish this goal, it provides a compact description of the basis of computational chemistry along with many examples of applications of these methods in various areas. The handbook is designed for researchers who are just being introduced into computational methods, as well as for those who are searching for the best choice for a solution to a specific problem involving theoretical approaches.

The first two volumes briefly describe different methods used in computational chemistry without going into exhaustive details of theory. Basic assumptions common to the majority of computational methods based on either quantum or statistical mechanics are outlined. Particular attention is paid to the limits of their applicability. Importantly, this volume also establishes definitions of various acronyms and terminologies used in the area of computational chemistry.

Since we assumed that the readers of this book are interested in applications of computational methods, a broad range of the most important applications of computational chemistry is provided. The applications include descriptions of standard chemical calculations for model molecules under various conditions. The handbook provides information on prediction of various molecular properties as well as investigations of chemical reactions.

There are two classes of species, biomolecules and nanomaterials, that are of a vital interest not only to chemists but also to physicists, biologists, and material scientists. Two volumes are devoted to description of specific computational

methods that are designed to investigate such species. Numerous examples of applications as well as descriptions of specific research problems and their solutions for various types of biomolecules and nanoparticles are given in volumes 3 and 4 of the handbook. The reviewed topics will attract the attention of all those who are already working or planning to start research involving computational approaches.

The fifth volume, an entirely new addition to the second edition, is devoted to the cheminformatics – both a description of techniques and discussion of applications. This is an important enhancement of the current *Handbook of Computational Chemistry* since the techniques of cheminformatics are vital tools of many computational chemists, those in research laboratories and, especially, those working in industry. Techniques developed allow storage, indexing, search, and comparing of information on chemical compounds. Such information could be used in the well-known quantitative structure-property (activity) relationship (QSPR/QSAR) studies for the prediction of specific property/activity of compounds of interest. QSAR techniques have been successfully applied for over 50 years in the pharmaceutical industry, and at the beginning of this century they were also adopted to investigate properties, biological activities, and toxicity of a new, unique class of compounds – nanomaterials.

We welcome all comments and suggestions related to the current version of the handbook as well as recommendations for possible revisions in the future. Many of the suggestions obtained from the readers of the first edition of the handbook have been applied to improve the second edition, and we are grateful for all inputs and comments. Such feedback and active participation help us to focus on areas that are important for the majority of our readers and improve the context of the book.

The Editors

Contents

Volume 1

Part I Theory and Methodology	1
<i>Anna Kaczmarek-Kedziera and Heribert Reis</i>	
1 Computational Chemistry: From the Hydrogen Molecule to Nanostructures	3
Lucjan Piela	
2 Molecular Mechanics: Principles, History, and Current Status	21
Valeri Poltev	
3 The Position of the Clamped Nuclei Electronic Hamiltonian in Quantum Mechanics	69
Brian Sutcliffe and R. Guy Woolley	
4 Remarks on Wave Function Theory and Methods	123
Dariusz Kędziera and Anna Kaczmarek-Kedziera	
5 Adiabatic, Born-Oppenheimer, and Non-adiabatic Approaches	173
Monika Stanke	
6 Directions for Use of Density Functional Theory: A Short Instruction Manual for Chemists	225
Heiko Jacobsen and Luigi Cavallo	
7 Introduction to Response Theory	269
Thomas Bondo Pedersen	
8 Intermolecular Interactions	295
Alston J. Misquitta	
9 Molecular Dynamics Simulation: From “Ab Initio” to “Coarse Grained”	337
Chris Lorenz and Nikos L. Doltsinis	

10	Statistical Mechanics of Force-Induced Transitions of Biopolymers	397
	Sanjay Kumar	
Part II Applications of Computational Methods to Model Systems		
	421	
<i>Anna Kaczmarek-Kedziera and Heribert Reis</i>		
11	Molecular Structure and Vibrational Spectra	423
	Jon Baker	
12	Molecular Electric, Magnetic, and Optical Properties	497
	Michał Jaszuński, Antonio Rizzo, and Kenneth Ruud	
13	Weak Intermolecular Interactions: A Supermolecular Approach	593
	Mark Waller and Stefan Grimme	
14	Chemical Reactions: Thermochemical Calculations	621
	John D. Watts	
15	Calculation of Excited States: Molecular Photophysics and Photochemistry on Display	639
	Luis Serrano-Andrés and Juan José Serrano-Pérez	
16	Solvent Effects in Quantum Chemistry	727
	Gerald Monard and Jean-Louis Rivail	
17	Solvent Effects on Molecular Electric Properties	741
	Miroslav Medved', Šimon Budzák, Wojciech Bartkowiak, and Heribert Reis	
18	Auxiliary Density Functional Theory: From Molecules to Nanostructures	795
	Patrizia Calaminici, Aurelio Alvarez-Ibarra, Domingo Cruz-Olvera, Victor-Daniel Domínguez-Soria, Roberto Flores-Moreno, Gabriel U. Gamboa, Gerald Geudtner, Annick Goursot, Daniel Mejía-Rodríguez, Dennis R. Salahub, Bernardo Zuniga-Gutierrez, and Andreas M. Köster	
19	Guide to Programs for Nonrelativistic Quantum Chemistry Calculations	861
	Tao Zeng and Mariusz Klobukowski	
20	Relativistic Methods in Computational Quantum Chemistry	885
	Paweł Tecmer, Katharina Boguslawski, and Dariusz Kędziera	
21	Time-Dependent Density Functional Theory: A Tool to Explore Excited States	927
	Daniel Escudero, Adèle D. Laurent, and Denis Jacquemin	

- 22 Molecular Aspects of Solvation Investigated Using Statistical Mechanics** 963
Norio Yoshida and Katsura Nishiyama

Volume 2

- Part III Solid States and Nanomaterials** **981**
Manthos Papadopoulos and Heribert Reis

- 23 Photoactive Semiconducting Oxides for Energy and Environment: Experimental and Theoretical Insights** 983
Malgorzata Makowska-Janusik and Abdel-Hadi Kassiba
- 24 Structures and Stability of Fullerenes, Metallofullerenes, and Their Derivatives** 1031
Alexey A. Popov
- 25 Structures and Electric Properties of Semiconductor clusters** 1097
Panaghiotis Karamanis
- 26 Structures, Energetics, and Spectroscopic Fingerprints of Water Clusters $n = 2 - 24$** 1139
Soohaeng Yoo and Sotiris S. Xantheas
- 27 Fundamental Structural, Electronic, and Chemical Properties of Carbon Nanostructures: Graphene, Fullerenes, Carbon Nanotubes, and Their Derivatives** 1175
Tandabany C. Dinadayalane and Jerzy Leszczynski
- 28 Optical Properties of Quantum Dot Nano-composite Materials Studied by Solid-State Theory Calculations** 1259
Ying Fu and Hans Ågren
- 29 Modeling of Quasi-One-Dimensional Carbon Nanostructures with Density Functional Theory** 1297
Veronica Barone, Oded Hod, and Juan E. Peralta
- 30 Variation of the Surface to Bulk Contribution to Cluster Properties** 1339
Antonis N. Andriotis, Zacharias G. Fthenakis, and Madhu Menon
- 31 Theoretical Studies of Structural and Electronic Properties of Clusters** 1357
Michael Springborg
- 32 The Response of Extended Systems to Electrostatic Fields** 1415
Michael Springborg, Bernard Kirtman, and Mohammad Molayem

33	Modeling of Nanostructures	1459
	Hande Toffoli, Sakir Erkoç, and Daniele Toffoli	
Part IV	Biomolecules	1515
	<i>Manoj K. Shukla</i>	
34	Quantum Cluster Theory for the Polarizable Continuum Model (PCM)	1517
	Roberto Cammi and Jacopo Tomasi	
35	Spin-Orbit Coupling in Enzymatic Reactions and the Role of Spin in Biochemistry	1557
	Boris F. Minaev, Hans Ågren, and V. O. Minaeva	
36	Protein Modeling	1589
	G. Naray-Szabo, A. Perczel, A. Lang, and D. K. Menyhard	
37	Applications of Computational Methods to Simulations of Proteins Dynamics	1627
	Wieslaw Nowak	
38	Molecular Dynamics and Advanced Sampling Simulations of Nucleic Acids	1671
	Jeremy Curuksu, Srinivasaraghavan Kannan, and Martin Zacharias	
39	Model Systems for Dynamics of π-Conjugated Biomolecules in Excited States	1697
	Mario Barbatti, Matthias Ruckebauer, Jaroslaw J. Szymczak, Bernhard Sellner, Mario Vazdar, Ivana Antol, Mirjana Eckert-Maksic, and Hans Lischka	
40	Low-Energy Electron (LEE)-Induced DNA Damage: Theoretical Approaches to Modeling Experiment	1741
	Anil Kumar and Michael D. Sevilla	
41	Computational Modeling of DNA and RNA Fragments	1803
	Jirı Šponer, Manoj K. Shukla, Jing Wang, and Jerzy Leszczynski	
42	Metal Interactions with Nucleobases, Base Pairs, and Oligomer Sequences; Computational Approach	1827
	Jaroslaw V. Burda, Jirı Šponer, and Filip Šebesta	
43	Two Photon Absorption in Biological Molecules	1875
	M. Alaraby Salem, Melis Gedik, and Alex Brown	

- 44 Consequences of Electron Attachment to Modified Nucleosides Incorporated into DNA** 1895
Lidia Chomicz-Mańska, Paweł Wityk, Łukasz Golon, Magdalena Zdrowowicz, Justyna Wiczak, Kinga Westphal, Michał Zyndul, Samanta Makurat, and Janusz Rak
- 45 Molecular Dynamics Simulations of Large Systems in Electronic Excited States** 1917
Jakub Rydzewski and Wiesław Nowak
- 46 Ab Initio Investigation of Photochemical Reaction Mechanisms: From Isolated Molecules to Complex Environments** 1943
Igor Schapiro, Patrick Zakhia El-Khoury, and Massimo Olivucci

Volume 3

- Part V Chemoinformatics** **1995**
Tomasz Puzyn
- 47 Computer Representation of Chemical Compounds** 1997
Jarosław Polanski and Johann Gasteiger
- 48 Application of Quantum Mechanics and Molecular Mechanics in Chemoinformatics** 2041
Natalia Sizochenko, Devashis Majumdar, Szczepan Roszak, and Jerzy Leszczynski
- 49 Molecular Descriptors** 2065
Andrea Mauri, Viviana Consonni, and Roberto Todeschini
- 50 Unsupervised Learning Methods and Similarity Analysis in Chemoinformatics** 2095
Katarzyna Odziomek, Anna Rybinska, and Tomasz Puzyn
- 51 Recent Developments in 3D QSAR and Molecular Docking Studies of Organic and Nanostructures** 2133
Bakhtiyor Rasulev
- 52 Ontologies in Chemoinformatics** 2163
Janna Hastings and Christoph Steinbeck
- 53 Chemoinformatics Methods for Studying Biomolecules** 2183
Adam Liwo, Cezary Czaplewski, Stanisław Ołdziej, Bartłomiej Zaborowski, Dawid Jagieła, and Jooyoung Lee

54	Open Source Chemoinformatics Software including KNIME Analytics	2201
	Georgios Leonis, Georgia Melagraki, and Antreas Afantitis	
55	Prioritization of Chemicals Based on Chemoinformatic Analysis	2231
	Paola Gramatica	
56	Predicting ADME Properties of Chemicals	2265
	Hyun Kil Shin, Young-Mook Kang, and Kyoung Tai No	
57	Predictive QSAR Modeling: Methods and Applications in Drug Discovery and Chemical Risk Assessment	2303
	Alexander Golbraikh, Xiang Simon Wang, Hao Zhu, and Alexander Tropsha	
58	Quantitative Structure–Activity Relationships of Antimicrobial Compounds	2341
	F. P. Maguna, N. B. Okulik, and Eduardo A. Castro	
	Index	2359

Editorial Board

Theory and Methodology

Anna Kaczmarek-Kedziera

Heribert Reis

Applications of Computational Methods to Model Systems

Anna Kaczmarek-Kedziera

Heribert Reis

Solid States and Nanomaterials

Manthos G. Papadopoulos

Heribert Reis

Biomolecules

Manoj K. Shukla

Chemoinformatics

Tomasz Puzyn

Contributors

Antreas Afantitis InSilicoLab LP, Athens, Greece

Novamechanics Ltd, Nicosia, Cyprus

Hans Ågren Division of Theoretical Chemistry and Biology, School of Biotechnology, Royal Institute of Technology, Stockholm, Sweden

Organic Chemistry, Bogdam Khmel'nitskii National University, Cherkassy, Ukraine

Aurelio Alvarez-Ibarra Departamento de Química, CINVESTAV, Av. Instituto Politécnico Nacional, México, D.F., México

Antonis N. Andriotis Department of Physics and Astronomy and Center for Computational Sciences, University of Kentucky, Lexington, KY, USA

Institute of Electronic Structure and Laser, FORTH, Heraklio, Crete, Greece

Ivana Antol Laboratory for Physical–Organic Chemistry – Division of Organic Chemistry and Biochemistry, Rudjer Bošković Institute, Zagreb, Croatia

Jon Baker Parallel Quantum Solutions, Fayetteville, AR, USA

Mario Barbatti Institute for Theoretical Chemistry, University of Vienna, Vienna, Austria

Veronica Barone Department of Physics, Central Michigan University, Mount Pleasant, MI, USA

Wojciech Bartkowiak Department of Physical and Quantum Chemistry, Wrocław University of Technology, Wrocław, Poland

Katharina Boguslawski Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Toruń, Poland

Alex Brown Department of Chemistry, University of Alberta, Edmonton, AB, Canada

Šimon Budzák Department of Chemistry, Faculty of Natural Sciences, Matej Bel University, Banská Bystrica, Slovak Republic

Jaroslav V. Burda Faculty of Mathematics and Physics, Charles University in Prague, Prague 2, Czech Republic

Patrizia Calaminici Departamento de Química, CINVESTAV, Av. Instituto Politécnico Nacional, México, D.F., México

Roberto Cammi Dipartimento di Chimica, Università di Parma, Parma, Italy

Eduardo A. Castro INIFTA, División Química Teórica, Departamento de Química, Facultad de Ciencias Exactas, La Plata, Buenos Aires, Argentina

Luigi Cavallo Physical Sciences and Engineering Division, Kaust Catalysis Center, King Abdullah University of Science and Technology (Kaust), Thuwal, Saudi Arabia

Lidia Chomicz-Mańka Biological Sensitizers Laboratory, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Viviana Consonni Department of Earth and Environmental Sciences, University of Milano-Bicocca, Milan, Italy

Domingo Cruz-Olvera Departamento de Química, CINVESTAV, Av. Instituto Politécnico Nacional, México, D.F., México

Jeremy Curuksu Mathematics Department, Swiss Federal Institute of Technology, Lausanne (EPFL), Switzerland

Cezary Czaplewski Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Tandabany C. Dinadayalane Department of Chemistry, Clark Atlanta University, Atlanta, GA, USA

Nikos L. Doltsinis Institute for Solid State Theory and Center for Multiscale Theory and Computation, University of Münster, Münster, Germany

Victor-Daniel Domínguez-Soria Departamento de Ciencias Básicas, UAM-A, Avenida San Pablo, México, D.F., México

Mirjana Eckert-Maksić Laboratory for Physical-Organic Chemistry – Division of Organic Chemistry and Biochemistry, Rudjer Bošković Institute, Zagreb, Croatia

Patrick Zakhia El-Khoury Chemistry Department, Bowling Green State University, Bowling Green, OH, USA

Sakir Erkoç Department of Physics, Middle East Technical University, Ankara, Turkey

Daniel Escudero CEISAM, UMR CNRS 6230, Université de Nantes, Nantes, France

Roberto Flores-Moreno Departamento de Química, CUCEI, Universidad de Guadalajara, Guadalajara Jalisco, México

Zacharias G. Fthenakis Department of Physics and Astronomy and Center for Computational Sciences, University of Kentucky, Lexington, KY, USA

Institute of Electronic Structure and Laser, FORTH, Heraklio, Crete, Greece

Ying Fu Division of Theoretical Chemistry and Biology, School of Biotechnology, Royal Institute of Technology, Stockholm, Sweden

Gabriel U. Gamboa Departamento de Química, CINVESTAV, Av. Instituto Politécnico Nacional, México, D.F., México

Johann Gasteiger Computer-Chemie-Centrum, Universität Erlangen-Nürnberg, Erlangen, Germany

Melis Gedik Department of Chemistry, University of Alberta, Edmonton, AB, Canada

Gerald Geudtner Departamento de Química, CINVESTAV, Av. Instituto Politécnico Nacional, México, D.F., México

Alexander Golbraikh Laboratory for Molecular Modeling and Carolina Center for Exploratory Cheminformatics Research, Division of Medicinal Chemistry and Natural Products, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA

Łukasz Golon Biological Sensitizers Laboratory, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Annick Goursot Institut Charles Gerhardt, UMR 5253 CNRS, Ecole de Chimie de Montpellier, Montpellier, Cédex 5, France

Paola Gramatica QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Theoretical and Applied Sciences, University of Insubria, Varese, Italy

Stefan Grimme Organisch-Chemisches Institut, Westfälische Wilhelms-Universität Münster, Münster, Germany

Janna Hastings European Bioinformatics Institute, Hinxton, UK

Oded Hod Department of Chemical Physics, School of Chemistry, The Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel

Heiko Jacobsen KemKom, New Orleans, LA, USA

Denis Jacquemin CEISAM, UMR CNRS 6230, Université de Nantes, Nantes, France

Institut universitaire de France, Paris, France

Dawid Jagieła Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Michał Jaszunski Institute of Organic Chemistry, Polish Academy of Sciences, Warszawa, Poland

Anna Kaczmarek-Kedziera Faculty of Chemistry, Nicolaus Copernicus University, Toruń, Poland

Young-Mook Kang Bioinformatics and Molecular Design Research Center, Seoul, Republic of Korea

Srinivasaraghavan Kannan Bioinformatics Institute (A*STAR), Singapore, Singapore

Panaghiotis Karamanis Groupe de Chimie Théorique et Réactivité, ECP, IPREM UMR 5254, Université de Pau et de Pays de l' Adour, PAU Cedex, France

Abdel-Hadi Kassiba Institute of Molecules and Materials of Le Mans – UMR-CNRS 6283, Université du Maine, Le Mans, France

Dariusz Kędziera Department of Chemistry and Photochemistry of Polymers, Faculty of Chemistry, Nicolaus Copernicus University, Toruń, Poland

Bernard Kirtman Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA, USA

Mariusz Klobukowski Department of Chemistry, University of Alberta, Edmonton, AB, Canada

Andreas M. Köster Departamento de Química, CINVESTAV, Av. Instituto Politécnico Nacional, México, D.F., México

Anil Kumar Department of Chemistry, Oakland University, Rochester, MI, USA

Sanjay Kumar Department of Physics, Banaras Hindu University, Varanasi, India

A. Láng Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös Loránd University and MTA-ELTE Protein Modelling Research Group, Budapest, Hungary

Adèle D. Laurent CEISAM, UMR CNRS 6230, Université de Nantes, Nantes, France

Jooyoung Lee Center for In Silico Protein Structure and School of Computational Sciences, Korea Institute for Advanced Study, Seoul, Republic of Korea

Georgios Leonis Novamechanics Ltd, Nicosia, Cyprus

Jerzy Leszczynski Department of Chemistry and Biochemistry, Interdisciplinary Center for Nanotoxicity, Jackson State University, Jackson, MS, USA

Hans Lischka Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX, USA

Adam Liwo Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Center for In Silico Protein Structure and School of Computational Sciences, Korea Institute for Advanced Study, Seoul, Republic of Korea

Chris Lorenz Department of Physics, King's College London, Strand, London, UK

F. P. Maguna Facultad de Agroindustrias, Universidad Nacional del Nordeste, Chaco, Argentina

Devashis Majumdar Department of Chemistry, Jackson State University, Jackson, MS, USA

Malgorzata Makowska-Janusik Institute of Physics, Faculty of Mathematics and Natural Science, Jan Dlugosz University in Czestochowa, Czestochowa, Poland

Samanta Makurat Biological Sensitizers Laboratory, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Andrea Mauri Department of Earth and Environmental Sciences, University of Milano-Bicocca, Milan, Italy

Miroslav Medved' Department of Chemistry, Faculty of Natural Sciences, Matej Bel University, Banská Bystrica, Slovak Republic

Daniel Mejía-Rodríguez Departamento de Química, CINVESTAV, Av. Instituto Politécnico Nacional, México, D.F., México

Georgia Melagraki InSilicoLab LP, Athens, Greece

Novamechanics Ltd, Nicosia, Cyprus

Madhu Menon Department of Physics and Astronomy and Center for Computational Sciences, University of Kentucky, Lexington, KY, USA

Institute of Electronic Structure and Laser, FORTH, Heraklio, Crete, Greece

D. K. Menyhárd Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös Loránd University and MTA-ELTE Protein Modelling Research Group, Budapest, Hungary

Boris F. Minaev Organic Chemistry, Bogdam Khmel'nitskii National University, Cherkassy, Ukraine

Theoretical Chemistry, School of Biotechnology, Royal Institute of Technology, Stockholm, Sweden

V. O. Minaeva Organic Chemistry, Bogdam Khmel'nitskii National University, Cherkassy, Ukraine

Alston J. Misquitta School of Physics, Queen Mary, University of London, London, UK

TCM Group, Cambridge, UK

Mohammad Molayem Physical and Theoretical Chemistry, University of Saarland, Saarbrücken, Germany

Gerald Monard Theoretical Chemistry and Biochemistry Group SRSMC, Nancy-University CNRS Boulevard des Aiguillettes, Vandoeuvre-les-Nancy, France

G. Náray-Szabó Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös Loránd University and MTA-ELTE Protein Modelling Research Group, Budapest, Hungary

Katsura Nishiyama Faculty of Education, Shimane University, Matsue, Japan

Kyoung Tai No Bioinformatics and Molecular Design Research Center, Seoul, Republic of Korea

Department of Biotechnology, Yonsei University, Seoul, Republic of Korea

Wiesław Nowak Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Toruń, Poland

Katarzyna Odziomek Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

N. B. Okulik Facultad de Agroindustrias, Universidad Nacional del Nordeste, Chaco, Argentina

Stanisław Oldziej Laboratory of Biopolymer Structure, Intercollegiate Faculty of Biotechnology, University of Gdańsk and Medical University of Gdańsk, Gdańsk, Poland

Massimo Olivucci Dipartimento di Chimica, Università di Siena, Siena, Italy

Thomas Bondo Pedersen Department of Chemistry, Centre for Theoretical and Computational Chemistry, University of Oslo, Oslo, Norway

Juan E. Peralta Department of Physics, Central Michigan University, Mount Pleasant, MI, USA

A. Perczel Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös Loránd University and MTA-ELTE Protein Modelling Research Group, Budapest, Hungary

Lucjan Piela Department of Chemistry, University of Warsaw, Warsaw, Poland

Jarosław Polanski Organic Chemistry and Chemoinformatics, University of Silesia Institute of Chemistry, Katowice, Poland

Valeri Poltev Autonomous University of Puebla, Puebla, Mexico

Alexey A. Popov Leibniz-Institute for Solid State and Materials Research (IFW Dresden), Dresden, Germany

Tomasz Puzyn Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Janusz Rak Biological Sensitizers Laboratory, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Bakhtiyor Rasulev Center for Computationally Assisted Science and Technology, North Dakota State University, Fargo, ND, USA

Heribert Reis Institute of Biology, Medicinal Chemistry and Biotechnology, National Hellenic Research Foundation, Athens, Greece

Jean-Louis Rivail Theoretical Chemistry and Biochemistry Group SRSMC, Nancy-University CNRS Boulevard des Aiguillettes, Vandoeuvre-les-Nancy, France

Antonio Rizzo Istituto per i Processi Chimico Fisici del Consiglio Nazionale delle Ricerche, Area della Ricerca di Pisa, Pisa, Italy

Szczepan Roszak Advanced Materials Engineering and Modelling Group, Faculty of Chemistry, Wrocław University of Technology, Wrocław, Poland

Matthias Ruckebauer Max-Planck-Institut für Kohlenforschung, Mülheim an der Ruhr, Germany

Kenneth Ruud Centre for Theoretical and Computational Chemistry, Department of Chemistry, University of Tromsø, Tromsø, Norway

Anna Rybinska Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Jakub Rydzewski Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Toruń, Poland

Dennis R. Salahub Department of Chemistry, CMS Centre for Molecular Simulation, IQST Institute for Quantum Science and Technology, University of Calgary, Calgary, AB, Canada

M. Alaraby Salem Department of Chemistry, University of Alberta, Edmonton, AB, Canada

Igor Schapiro Chemistry Department, Bowling Green State University, Bowling Green, OH, USA

Filip Šebesta Faculty of Mathematics and Physics, Charles University in Prague, Prague 2, Czech Republic

Bernhard Sellner Laboratory for Physical–Organic Chemistry – Division of Organic Chemistry and Biochemistry, Rudjer Bošković Institute, Zagreb, Croatia

Luis Serrano-Andrés Instituto de Ciencia Molecular, Universitat de València, Valencia, Spain

Juan José Serrano-Pérez Department of Chemistry, Imperial College London Computational and Structural Research Group, London, UK

Michael D. Sevilla Department of Chemistry, Oakland University, Rochester, MI, USA

Hyun Kil Shin Department of Biotechnology, Yonsei University, Seoul, Republic of Korea

Manoj K. Shukla US Army Engineer Research and Development Center, Vicksburg, MS, USA

Natalia Sizochenko Department of Chemistry, Jackson State University, Jackson, MS, USA

Jiří Šponer Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno, Czech Republic

Michael Springborg Physical and Theoretical Chemistry, University of Saarland, Saarbrücken, Germany

School of Materials Science and Engineering, Tianjin University, Tianjin, China

Monika Stanke Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicholas Copernicus University, Toruń, Poland

Christoph Steinbeck European Bioinformatics Institute, Hinxton, UK

Brian Sutcliffe School of Science and Technology, Nottingham Trent University, Nottingham, UK

Service de Chimie Quantique et Photophysique, Université Libre de Bruxelles, Bruxelles, Belgium

Jaroslav J. Szymczak Research Lab Computational Technologies and Applications, University of Vienna, Vienna, Austria

Paweł Tecmer Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Toruń, Poland

Roberto Todeschini Department of Earth and Environmental Sciences, University of Milano-Bicocca, Milan, Italy

Daniele Toffoli Department of Chemistry, Middle East Technical University, Ankara, Turkey

Hande Toffoli Department of Physics, Middle East Technical University, Ankara, Turkey

Jacopo Tomasi Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Pisa, Italy

Alexander Tropsha Division of Medicinal Chemistry and Natural Products, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA

Mario Vazdar Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX, USA

Mark Waller Organisch-Chemisches Institut, Westfälische Wilhelms-Universität Münster, Münster, Germany

Jing Wang Department of Chemistry and Biochemistry, Interdisciplinary Center for Nanotoxicity, Jackson State University, Jackson, MS, USA

Xiang Simon Wang Laboratory for Molecular Modeling and Carolina Center for Exploratory Cheminformatics Research, Division of Medicinal Chemistry and Natural Products, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA

John D. Watts Department of Chemistry and Biochemistry, Jackson State University, Jackson, MS, USA

Kinga Westphal Biological Sensitizers Laboratory, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Justyna Wiczak Biological Sensitizers Laboratory, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Paweł Wityk Biological Sensitizers Laboratory, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

R. Guy Woolley School of Science and Technology, Nottingham Trent University, Nottingham, UK

Service de Chimie Quantique et Photophysique, Université Libre de Bruxelles, Bruxelles, Belgium

Sotiris S. Xantheas Chemical and Material Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA

Soohaeng Yoo Chemical and Material Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA

Norio Yoshida Department of Chemistry, Graduate School of Science, Kyushu University, Fukuoka, Japan

Bartłomiej Zaborowski Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Martin Zacharias Physik-Department T38, Technische Universität München, Garching, Germany

Magdalena Zdrowowicz Biological Sensitizers Laboratory, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Tao Zeng Department of Chemistry, Carleton University, Ottawa, ON, Canada

Hao Zhu Laboratory for Molecular Modeling and Carolina Center for Exploratory Cheminformatics Research, Division of Medicinal Chemistry and Natural Products, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA

Bernardo Zuniga-Gutierrez Departamento de Ciencias Computacionales, Universidad de Guadalajara, Guadalajara Jalisco, México

Michał Żyndul Biological Sensitizers Laboratory, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Part I

Theory and Methodology

Anna Kaczmarek-Kedziera and Heribert Reis

Computational Chemistry: From the Hydrogen Molecule to Nanostructures

1

Lucjan Piela

Contents

Exceptional Status of Chemistry	4
A Hypothetical Perfect Computer	7
Does Predicting Mean Understanding	10
Orbital Model	12
Molecular Surface	13
Molecular Recognition	14
Rules of the Nanoscale	16
Power of Computer Experiments	16
Concluding Remarks	17
Bibliography	18

Abstract

Quantum chemical calculations rely on a few fortunate circumstances like usually small relativistic and negligible electrodynamic (QED) corrections and large nuclei-to-electron mass ratio. The fast progress in computer technology revolutionized theoretical chemistry and gave birth to computational chemistry. The computational quantum chemistry provides for experimentalists the ready-to-use tools of new kind offering powerful insight into molecular internal structure and dynamics. It is important for the computational chemistry to elaborate methods, which look at molecule in a multiscale way, which provide first of all its global and synthetic description such as shape and charge distribution, and compare this description with those for other molecules. Only such a picture can free researchers from seeing molecules as a series of case-by-case studies. Chemistry represents a science of analogies and similarities, and

L. Piela (✉)

Department of Chemistry, University of Warsaw, Warsaw, Poland

e-mail: lucjan.piela@tiger.chem.uw.edu.pl

© Springer International Publishing Switzerland 2017

J. Leszczynski et al. (eds.), *Handbook of Computational Chemistry*,

DOI 10.1007/978-3-319-27282-5_1

3

computational chemistry should provide tools for seeing this. This is especially useful for the supramolecular chemistry, which allow chemists to planify and study intermolecular interactions. Some of them, involving concave and convex moieties, represent molecular recognition, which assures a perfect fitting of the two molecular shapes. A sequence of molecular recognitions often leads to self-assembling and self-organization, typical to nanostructures.

Exceptional Status of Chemistry

Contemporary science fails to explain the largest-scale phenomena taking place in the Universe, such as speeding up the galaxies (supposedly due to the undefined “black energy”) and the nature of the lion share of Universe’s matter (also unknown “dark matter;” estimated roughly as about 90 % of the existing matter).

Quantum chemistry is in by far better position, which may be regarded even as exceptional in the whole domain of sciences. The chemical phenomena are explainable by contemporary theories down to individual molecules (which represent the subject of quantum chemistry). It turned out, by comparing theory and experiment, that the solution to the Schrödinger equation (Schrödinger 1926) offers in most cases a quantitatively correct picture. Only molecules with very heavy atoms, due to the relativistic effects becoming important, need to be treated in a special way based on the Dirac theory (Dirac 1928). This involves an approximate Hamiltonian in the form of sum of Dirac Hamiltonians for individual electrons and the electron-electron interactions in the form of the (nonrelativistic) Coulomb terms, a common and computational successful practice ignoring, however, the resulting resonance character of all the eigenvalues (Brown and Ravenhall 1951; Pestka et al. 2008). When, very rarely, a higher accuracy is needed, one may eventually include the quantum electrodynamics (QED) corrections, a procedure currently far for routine application, but still feasible for very small systems (Łach et al. 2004).

This success of computational quantum chemistry is based on a few quite fortunate circumstances (for references, see, e.g., Piela 2014):

- Atoms and molecules are built of only two kinds of particles: nuclei and electrons.
- The constituents of atoms and molecules are treated routinely as point charges. Although nuclei have nonzero size (electrons are regarded as point-like particles), the size is so small that its influence is much below chemical accuracy (Łach et al. 2004).
- The QED corrections are much smaller than energy changes in chemical phenomena (like 1:10,000,000) and may be safely neglected in most applications (Łach et al. 2004).
- The nuclei are thousand times heavier than electrons, and therefore, except some special situations, they move thousand times slower than electrons. This makes possible to solve the Schrödinger equation for electrons, assuming that the nuclei do not move, i.e., their positions are fixed in space (“clamped nuclei”).

This concept is usually presented within the so-called adiabatic approximation. In this approximation, the motion of the nuclei is considered in the next step, in which the electronic energy (precalculated for any position of the nuclei), together with a usually small diagonal correction for coupling the nuclei-electron motion, plays the role of the potential energy surface (PES). The total wave function is assumed to be a product of the electronic wave function and of the function describing the motion of the nuclei. The commonly used Born-Oppenheimer (BO) (1927) approximation is less accurate than the adiabatic one, because it neglects the abovementioned diagonal correction, making the PES independent of nuclear masses. Using the PES concept, one may introduce the crucial idea of the spatial structure of molecule, defined as those positions of the nuclei that assure a minimum of the PES. This concept may be traced back to Hund (1927). Moreover, this structure corresponds to a certain ground-state electron density distribution that exhibits atomic cores, atom-atom bonds, and atomic lone pairs.

It is generally believed that the exact analytical solution to the Schrödinger equation for any atom (except the hydrogen-like atom) or molecule is not possible. Instead, some reasonable approximate solutions can be obtained, practically always involving calculation of a large number of molecular integrals, and some algebraic manipulations on matrices built of these integrals. The reason for this is efficiency of what is known as the algebraic approximation (“algebraization”) of the Schrödinger equation. The algebraization is achieved by postulating a certain finite basis set of functions $\{\Phi_i\}_{i=1}^{i=M}$ and expanding the unknown wave function as a linear combination of the Φ_i 's with unknown expansion coefficients. Such an expansion can be encountered in the one-electron case (e.g., linear combination of atomic orbitals, LCAO, introduced by Bloch (1928)) and/or in the many-electron case, e.g., the total wave function expansion in Slater determinants, related to configurations (Slater 1930), or in the explicitly correlated many-electron functions (Hylleraas 1929). It is assumed for good-quality calculations (arguments are as a rule of numerical character only) that a finite M chosen is large enough to produce sufficient accuracy, with respect to what would be with $M = \infty$ (exact solution). The abovementioned integrals appear, because, after the expansion is inserted into the Schrödinger equation, one makes the scalar products (they represent the integrals, which should be easy to calculate) of the expansion with $\Phi_1, \Phi_2 \dots \Phi_M$, consecutively. In this way the task of finding the wave function by solving the Schrödinger equation is converted into an algebraic problem of finding the expansion coefficients, usually by solving some matrix equation. It remains to take care of the choice of the basis set $\{\Phi_i\}_{i=1}^{i=M}$. The choice represents a technical problem, but unfortunately it contains a lot of arbitrariness and, at the same time, is one of the most important factors influencing cost and quality of the computed approximate solution. Application of functions Φ_i based on the Gaussian-type one-electron orbitals (GTO) (Boys et al. 1956) provides a low cost/quality ratio, and this fact is considered as one of the most important factors, which made the contemporary computational chemistry efficient.

Algebraization involves as a rule a large M and therefore the whole procedure requires fast computing facilities.¹ These facilities changed over time, from very modest manual mechanical calculators at the beginning of the twentieth century to what we consider *now* as powerful supercomputers. Almost immediately after formulation of quantum mechanics in 1926, Douglas Hartree published several papers (Hartree 1927) presenting his manual calculator-based solutions for atoms of rubidium and chlorine. However amazing it looks now, these were self-consistent *ab initio*² computations.

In 1927 Walter Heitler and Fritz Wolfgang London clarified the origin of the covalent chemical bond (Heitler and London 1927), the concept crucial for chemistry. The authors demonstrated, in numerical calculations, that the nature of the covalent chemical bond in H_2 is of quantum character, because the (semiquantitatively) correct description of H_2 emerged only after inclusion the exchange of electrons 1 and 2 between the nuclei in the formula $a(1)b(2)$ (a, b are the 1s atomic orbitals centered on nucleus a and nucleus b , respectively) resulting in the wave function $a(1)b(2) + a(2)b(1)$. Thus, taking into account also the contribution of Hund, 1927 is therefore the year of birth of computational chemistry.

Perhaps the most outstanding manual calculator computations were performed in 1933 by Hubert James and Albert Coolidge for the hydrogen molecule (James and Coolidge 1933). This variational result has been the best one in the literature over the period of 27 years.

The 1950s mark the beginning of a new era – the time of *programmable* computers. Apparently, just another tool for number crunching became available. In fact, however, the idea of programming made a revolution because:

- It liberated humans from tedious manual calculations.
- It offered large speed of computation, incomparable to that of any manual calculator.
- The new data storage tools soon became of massive character.
- It resulted in more and more efficient programs, based on earlier versions (like staying “on the shoulders of the giants”), offering possibilities to calculate dozens of molecular properties.
- These programs and data soon became transportable, and also they could be copied and diffused worldwide.
- It allowed the dispersed, parallel, and remote calculations.
- It resulted in the new branch of chemistry: computational chemistry.³

¹Computational chemistry contributed significantly to applied mathematics, because new methods had to be invented in order to treat the algebraic problems of the previously unknown scale (like for M of the order of billions); see, e.g., reference Roos (1972).

²That is, derived from the first principles of the (nonrelativistic) quantum mechanics

³It is difficult to define what computational chemistry is. Obviously, whatever involves calculations in chemistry might be treated as part of it. This, however, sounds like a pure banality. The same is with the idea that computational chemistry means chemistry, which uses computers.

- It allowed performing calculations by anyone, even not trained in chemistry, quantum chemistry, mathematics, etc.

These achievements appeared gradually in the chemical practice. The first ab initio Hartree-Fock calculations (based on the ideas of Douglas Hartree (1927) and Vladimir Fock (1930)) on programmable computer for diatomic molecules have been performed at the Massachusetts Institute of Technology in 1956, using basis set of Slater-type orbitals. The first calculations with Gaussian-type orbitals have been carried out by Boys and coworkers also in 1956. An unprecedented spectroscopic accuracy has been obtained for the hydrogen molecule in 1960 by Kołos and Roothaan. In the early 1970s began the era of gigantic programs offering computation of many physical quantities at various levels of approximation.

We currently live in this era with computational possibilities growing exponentially (the notorious “Moore law” of doubling the computer power every 2 years).⁴ This enormous progress revolutionized our civilization in a very short time. The revolution in computational quantum chemistry changed chemistry in general, because computations became feasible for molecules of interest for experimental chemists. The progress has been accompanied by achievements in theory, however mainly of the character related to computational needs. Nowadays fair accuracy computations are possible for molecules composed of several hundreds of atoms, spectroscopic accuracy is achievable for molecules with a dozen of atoms, while the QED calculations can be performed for the smallest molecules only (few atoms). In many instances, we now reach the stage, when computing a molecular property is faster, less expensive, and more informative than to measure it.

A Hypothetical Perfect Computer

Suppose we have at our disposal a computer, which is able to solve the Schrödinger equation exactly for any system and in negligible time.⁵ Thus, we have free access to the absolute detailed picture of any molecule. This means we may predict with high accuracy and confidence the value of any property of any molecule. We might be tempted to say that being able to give such predictions is the ultimate goal of science: *We know everything about our system. If you want to know more about the world, take other molecules and just compute, you will know.*

It is questionable whether this problem needs any solution at all. If yes, the author sticks to the opinion that computational chemistry means quantitative description of chemical phenomena at the molecular/atomic level.

⁴The speed as well as the capacity of computer’s memory increased about 100 billion times over the period of 40 years. This means that what now takes an hour of computations would require in 1960 about 10,000 years of computing.

⁵In addition, we assume the computer is so clever that it automatically rejects those solutions, which are not square integrable or do not satisfy the requirements of symmetry for fermions and bosons. Thus, all nonphysical solutions are rejected.

Let us consider a system composed of 6 carbon nuclei, 6 protons, and 42 electrons. Suppose we want to know the geometry of the system for the ground state. Our computer answers that it does not know what we mean by the term “geometry.” We are forced to be more precise now and say that we are interested in the carbon-carbon (CC) and carbon-hydrogen (CH) distances. The computer answers that it is possible to compute only the *mean* distances and provides them together with the proton-proton, carbon-electron, proton-electron, and electron-electron distances, because it treats all the particles on equal footing basis. We look at the CC and CH distances and see that they are much larger than we expected for the CC and CH bonds in benzene. The reason is, that in our perfect wave function, the permutational symmetry is correctly included. This means that the average carbon-proton distance takes into account *all* carbons and *all* protons, the same with other distances. To deduce more, we may ask for computing other quantities like angles, involving three nuclei. Here too we will be confronted with numbers including averaging over identical particles. These difficulties do not necessarily mean that the molecule has no spatial structure at all, although this can also happen. The numbers produced would be extremely difficult to translate into a useful 3D picture even for quite small molecules, not speaking about such a floppy molecule as a protein.

In many cases we would obtain a 3D picture we did not expect. This is because many molecular structures, we are familiar with, represent higher-energy metastable electronic states (isomers). This is the case in our example. When solving the time-dependent Schrödinger equation, we are confronted with this problem. Let us use as a starting wave function the one corresponding to the benzene molecule. In time evolution, we will stay probably with a similar geometry for a long time. However, there is a chance that after a long period, the wave function changes to that corresponding to three interacting acetylene molecules (3 times HCCH). The Born-Oppenheimer optimized ground electronic state corresponds to the benzene [-230.703 a.u. in the Hartree-Fock approximation for the 6-311-G(d) basis set]. The three isolated acetylene molecules (in the same approximation) have the energy -230.453 a.u. and the molecule (also with the same formula C_6H_6) $H_3C-C \equiv C-C \equiv C-CH_3$ -230.592 a.u. Thus, the benzene molecule seems to be a stable ground state, while the three acetylenes and the diacetylene are metastable states within the same ground electronic state of the system. All the three physically observed realizations of the system $6C + 6H$ are separated by barriers, and this is the reason why they are observable.

What is, therefore, the most stable electronic ground state corresponding to the flask of benzene? This is a quite different question, which pertains to systems larger than a single molecule. If we multiply the number of atoms in a single molecule of benzene by a natural number N , we are confronted with new possibilities of combining atoms into molecules, not necessary of the same kind and possibly larger than C_6H_6 . For a large N , we are practically unable to find all the possibilities. In some cases, when based on chemical intuition and limiting to simple molecules, we may guess particular solutions. For example, to lower the energy for the flask of benzene, we may allow formation of the methane molecules and the graphite

(the most stable form of carbon). Therefore, the flask of benzene represents a metastable state.

Suppose we wish to know the dipole moment of, say, the HCl molecule, the quantity that tells us important information about the charge distribution. We look up the computer output and we do not find anything about dipole moment. The reason is that all molecules have the same dipole moment in any of their stationary state Ψ , and this dipole moment equals to zero; see, e.g., reference Piela (2014), p. 735. Indeed, the dipole moment is calculated as the mean value of the dipole moment operator, i.e., $\boldsymbol{\mu} = \langle \Psi | \hat{\boldsymbol{\mu}} | \Psi \rangle = \langle \Psi | (\sum_i q_i \mathbf{r}_i) | \Psi \rangle$; index i runs over all electrons and nuclei. This integral can be calculated very easily: the integrand is antisymmetric with respect to inversion and therefore $\boldsymbol{\mu} = \mathbf{0}$. Let us stress that our conclusion pertains to the *total* wave function, which has to reflect the space invariance with respect to inversion, leading to the zero dipole moment. If one applied the transformation $\mathbf{r} \rightarrow -\mathbf{r}$ only to some particles in the molecule (e.g., electrons), and not to the other ones (e.g., the nuclei), then the wave function will show no parity (it would be neither symmetric nor antisymmetric). We do this in the adiabatic or Born-Oppenheimer approximation, where the electronic wave function depends on the electronic coordinates only. This explains why the integral $\boldsymbol{\mu} = \langle \Psi | \hat{\boldsymbol{\mu}} | \Psi \rangle_e$ (the integration is over electronic coordinates only) does not equal to zero for some molecules (which we call polar ones). Thus, to calculate the dipole moment, we have to use the adiabatic or the Born-Oppenheimer approximation.

Now we decide to quit the perfect computer and the perfect picture of the molecule it produces, because the computer output looks like almost useless, paradoxically because it is “too exact,” similarly as a map of a terrain with the scale 1:1. We desperately need an approximation (a model), which will tell us the most important features of the molecule under consideration. This is why we introduce the Born-Oppenheimer approximation. Its first consequence is the molecular geometry, this one which leads to a minimum of the electronic energy. There is a problem though that usually we have many such minima of different energy, each minimum corresponding to its own electronic density distribution. Any such distribution corresponds to some particular chemical bond pattern.⁶ In most cases the user of computers does not even think of these minima, because he performs the calculations for a predefined configuration of the nuclei and forces the system (usually not being aware of it) to stay in its vicinity. This is especially severe for large molecules, such as proteins. They have an astronomic number of stable conformations,⁷ but often we take one of them and perform the calculations for this one. It is even difficult to say that we select this one, because we rarely even consider the other conformations. In this situation, we usually take as the starting point a crystal structure conformation (we *believe* in its relevance).

⁶Bond patterns are the same for different conformers.

⁷For a dipeptide, one has something like ten energy minima, counting only the backbone conformations (and not counting the side-chain conformations for simplicity). For very small protein of, say, a hundred amino acids, the number of conformations is therefore of the order of 10^{100} , a very large number exceeding the estimated number of atoms in the Universe.

Moreover, usually one starts calculations by setting a starting electronic density distribution. The choice of this density distribution may influence the final electronic density and the final geometry of the molecule. In routine computations, one guesses the starting density according to the starting nuclear configuration chosen. This may look as a reasonable choice, except when small deformation of the nuclear framework leads to large changes in the electronic density.

In conclusion, in practice computer gives the solution, which is close to what the computing person considers as “reasonable” and sets as the starting point.

Does Predicting Mean Understanding

The existing commercial programs allow us to make calculations for molecules, treating each molecule as a new task, as if every molecule represented a new world, which has nothing to do with other molecules. We might not be satisfied with such a picture. We might be curious:

- Whether, living in the 3D space, the system has or has not a certain shape (“geometry”)
- If yes, why the shape is of this particular kind
- If the shape is fairly rigid or rather flexible
- If it is reasonable to assign the system some external surface
- If there are some characteristic substructures in the system
- How these substructures interact one another interact
- How they influence the calculated global properties, etc.
- If the same substructures are present in other molecules
- If the presence of the same substructures does determine similar properties

It is of fundamental importance for chemistry that we do not study particular cases, case by case, but derive some general rules. Strictly speaking these rules are false for, because of approximations made, they are valid to some extent only. However, despite this, they enable chemists to operate, to understand, and to be efficient. If we relied uniquely on exact solutions of the Schrödinger equation, there would be no chemistry at all, and people would lose the power of rationalizing this branch of science, in particular to design syntheses of new molecules. Chemists rely on molecular spatial structure (nuclear framework), on the concepts of valence electrons, chemical bonds, electronic lone pairs, etc. All these notions have no rigorous definition, but they still are of great importance describing a *model* of molecule. A chemist predicts that two OH bonds have similar properties, wherever they are. Moreover, chemists are able to predict differences in the OH bonds by considering what the neighboring atoms are in each case. It is of fundamental importance in chemistry that a group of atoms with a certain bond pattern (functional group) represents an entity that behaves similarly, when present in different molecules.

We have at our disposal various scales at which we can look at details of the molecule under study. In the crudest approach, we may treat the molecule as a point mass, which contributes to the gas pressure. Next we might become interested in

the shape of the molecule, and we may approximate it first as a rigid rotator and get an estimation of rotational levels we can expect. Then we may leave the rigid body model and allow the atoms of the molecule to vibrate about their equilibrium positions. In such a case, we need to know the corresponding force constants. This requires either to choose the structural formula (chemical bond pattern) of the molecule together with taking the corresponding empirical force constants or to apply the normal mode analysis, first solving the Schrödinger equation in the Born-Oppenheimer approximation (we have a wide choice of the methods of solution). In the first case, we obtain an estimation of the vibrational levels; in the second, we get more reliable vibrational analysis, especially for larger atomic orbital expansions. If we wish we may consider anharmonicity of vibrations.⁸

At the same time, we obtain the electronic density distribution from the wave function Ψ for N electrons⁹

$$\rho(\mathbf{r}) = N \sum_{\sigma_1=\frac{1}{2}}^{\frac{1}{2}} \int d\tau_2 d\tau_3 \dots d\tau_N |\Psi(\mathbf{r}, \sigma_1, \mathbf{r}_2, \sigma_2, \dots, \mathbf{r}_N, \sigma_N)|^2.$$

According to the Hellmann-Feynman theorem (Hellmann 1937; Feynman 1939) ρ is sufficient to compute the forces acting on the nuclei. We may compare the resulting ρ calculated at different levels of approximation. The density distribution ρ can be analyzed in the way known as Bader analysis (Bader 1994). First, we find all the critical points, in which $\nabla\rho = 0$. Then, one analyzes the nature of each critical point by diagonalizing the Hessian matrix calculated at the point¹⁰:

- If the three eigenvalues are negative, the critical point corresponds to a maximum of ρ .
- If two are negative and one positive, the critical point corresponds to a covalent bond.

⁸The low-frequency vibrations may be used as indicators to look at possible instabilities of the molecule, such as dissociation channels, formation of new bonds, etc. Moving all atoms, first according to a low-frequency normal mode vibration and continuing the atomic displacements according to the maximum gradient decrease, we may find the saddle point and then, sliding down, detect the products of a reaction channel.

⁹The integration of $|\Psi|^2$ is over the coordinates (*space and spin ones*) of all the electrons except one (in our case the electron 1 with the coordinates \mathbf{r}, σ_1) and in addition the summation over its spin coordinate (σ_1). As a result one obtains a function of the position of the electron 1 in space: $\rho(\mathbf{r})$. The wave function Ψ is antisymmetric with respect to exchange of the coordinates of any two electrons, and, therefore, $|\Psi|^2$ is symmetric with respect to such an exchange. Hence, the definition of ρ is independent of the label of the electron we do not integrate over. According to this definition, ρ represents nothing else but the density of the electron cloud carrying N electrons and is proportional to the probability density of finding an electron at position \mathbf{r} .

¹⁰Strictly speaking the nuclear attractors do not represent critical points, because of the cusp condition (Kato 1957). This worry represents rather a pure theoretical problem, when, what is a common practice, one uses Gaussian atomic orbitals as the basis set.

- If one is negative and two positive, the critical point corresponds to a center of an atomic ring.
- If all three are positive, the critical point corresponds to an atomic cavity.

The chemical bond critical points correspond to *some* pairs of atoms, and there are other pairs of atoms, which do not form bonds.¹¹ The Bader analysis enables chemists to see molecules in a synthetic way, nearly independent of the level of theory that has been used to describe it, focusing on the ensemble of critical points. We may compare this density with the density of other molecules, similar to our one, to look whether one can see some local similarities. We may continue this, getting a more and more detailed picture down to the almost exact solution of the Schrödinger equation.

It is important in chemistry to follow such way, because at its beginning as well as at its end, we know very little about chemistry. We learn chemistry traveling along the way.

Orbital Model

The wave function for identical fermions has to be antisymmetric with respect to exchange of coordinates (space and spin ones) of any two of them. This means that two electrons having the same spin coordinate cannot occupy the same position in space. Since wave functions are continuous, this means that electrons of the same spin coordinate avoid each other (“Fermi hole” or “exchange hole”). This Pauli exclusion principle does not pertain to two electrons of opposite spin. However, electrons repel one another (Coulombic force) at any finite distance, i.e., they have to avoid one another because of their charge (“Coulomb hole” or “correlation hole”). It turned out (references in Piela (2014), p. 710) that the Fermi hole is by far more important than the Coulomb hole. A high-quality wave function has to reflect the Fermi and the Coulomb holes. The corresponding mathematical expression should have the antisymmetrization operator in front, and this will take care of the Pauli principle (and introduce a Fermi hole). Besides this, it should have some parameters or mathematical structure controlling somehow the distance between any pair of electrons (this will introduce the Coulomb repulsion). Since the Fermi hole is much more important, it is reasonable to consider first a wave function that takes care of the Fermi hole only. The simplest way to take the Fermi hole into account is the orbital model (approximation). Within the orbital model, the most advanced is the Hartree-Fock method. In this method, the Fermi hole is taken into account by construction (antisymmetrizer). The Coulomb hole is not present, because the Coulomb interaction is calculated through averaging the positions of the electrons.

¹¹Although this is only after assuming a reasonable threshold as a criterion. This is a common situation in chemistry – strictly speaking there is no such thing as chemical bond in a polyatomic molecule.

The orbital model is incorrect (i.e., represents an approximation), because it neglects the Coulomb hole. Due to its simplicity, it gains, however, enormous scientific power, because:

- It allows to see the electronic structure as contributions of individual electrons, with their own “wave functions,” i.e., orbitals, mutually orthogonal, with a definite mathematical form, symmetry, energy (“orbital energy levels”), etc.
- We take the Pauli exclusion principle into account by not allowing occupations of an orbital by more than two electrons (if two, then of the opposite spin coordinates). The occupation of all orbital levels is known as the all-important orbital diagram.
- The orbital energy may be interpreted as the energy needed to remove an electron from the orbital (assuming that all the orbitals do not change during the removing, Koopmans theorem (Koopmans 1933/1934)).
- Molecular electron excitations may be often identified with changing the electron occupancy in the orbital diagram.
- We may even consider electronic correlation (still incorporating a concept of the Coulomb hole) either by allowing different orbitals for electrons of different spin or considering a wave function expansion composed of electron diagrams with various occupations.
- One may trace the molecular perturbations to changes in the orbital diagram.
- One may describe chemical reactions as a continuous change from a starting to a final molecular diagram. Theory and computational experience bring some rules, like that only those orbitals of the molecular constituents mix, which have similar orbital energies and have the same symmetry. This leads to important symmetry selection rules for chemical reactions (Woodward and Hoffmann 1965; Fukui and Fujimoto 1968) and for optical excitations (Cotton 1990).
- The orbital model provides a set of fundamental keywords, which create a chemical vocabulary – a tool to communicate among chemists (including quantum chemists). This language and the numerical experience supported by theory create a kind of quantum mechanical intuition coupled to experiment, which allows to compare molecules, to classify them, and to predict their properties in a descriptive way. A majority of theoretical terms in general chemistry stem from the orbital theory.

Molecular Surface

Many ideas of chemistry, although extremely helpful and sometimes indispensable, are more or less arbitrary. The all-important concept of molecular surface belongs to this class. The molecular surface, also representing a kind of synthetic view of molecule, can be defined in several distinct ways, like:

- The isosurface corresponding to a given electron density contour.
- The isosurface corresponding to a given interaction energy with an atom like helium atom.

- The surface (van der Waals surface) of the sum of the van der Waals spheres of all the atoms of the molecule, with the van der Waals radii assigned (by an empirical consensus).
- The accessible surface area (Lee and Richards 1971) is the area of the solvent accessible surface (SAS) that corresponds to rolling a ball representing a water molecule (its radius usually taken as 1.4 Å) over the van der Waals surface, i.e., extending the latter by the radius of the water molecule.

At any of these definitions, some molecules may possess multiple surfaces (external and internal), due to possible cavities. It is true that each of the definitions leads to a different surface. However, the role of such a molecular surface is in fact to define the molecular shape, and the resulting shapes do not differ much. In the all-important intermolecular interactions, the *overall* molecular shape belongs to primary and synthetic characteristics of the molecules involved and often decides about the strength of the interaction.

The molecular shape is a function of the quantum mechanical state of the molecule. The most important in chemistry and physics are the low-energy lying states, in most cases the conformational states separated by some rather low energy barriers and all corresponding to the ground electronic state. How to find these lowest-energy states among myriads of others (their number sometimes of the order of 10^{100}) represents a very serious problem of chemistry (Piela 2002). The most promising way is to use the concept of pseudoatoms (like united atoms), the appropriate force field, and the Monte Carlo approach to overcome the energy barriers (Koliński and Skolnick 2004).

Molecular Recognition

The interaction energy of two molecules in the vacuum is dominated by several physically distinct contributions (Jeziorski and Kołos 1977). The first two result from the first-order perturbation theory: the electrostatic interaction energy and the valence repulsion energy. The electrostatic interaction represents a Coulomb interaction of the two molecular charge distributions, which are identical (“frozen”) to those corresponding to the noninteracting molecules. The valence repulsion, a consequence of the Pauli exclusion principle, increases very strongly (exponentially) with shortening of the intermolecular distance, when the two molecular charge distributions begin to overlap. In the second order, one has the induction interaction (an electrostatic effect of relaxing the charge distribution) and the dispersion interaction (the intermolecular electron-electron Coulomb correlation effect) with some small electron-exchange corrections of these quantities.

Among these effects the most important are the electrostatic interaction and the valence repulsion. The first one is of the most long-range character and represents the only interaction that may change the sign (attraction or repulsion) upon rotation of one of the molecules by 180° (large anisotropy). This means the electrostatic interaction usually predefines to a large extent the mutual orientation

of the approaching molecules. The two molecules when approaching just try to fit electrostatically one another. This electrostatic fitting may be lost by an ill fitting of the inappropriate shapes of the molecules (which results in an excessive valence repulsion energy; this is why this effect may be critically important) or may be enhanced by a perfect fitting of the shapes over a large contact area of the interacting molecules (molecular recognition). The induction and dispersion interactions always represent attraction and have much smaller anisotropy. This is why in most cases they do not decide about the resulting structure, while always enhancing its stability.

The molecular surface plays a role of determining a no-access zone for other molecules or atoms. A violation of the zone's border is possible, but at a large energy price (valence repulsion). In addition to this repulsive effect, there is an energy gain or expense for approaching the molecular surface at a given surface point. This is related to the (long-range) electrostatic potential created by the charge distribution of the molecule. The electrostatic potential represents the interaction energy (attraction or repulsion) of the unit charge $+1$ a.u. with the electrostatic field created by the molecule. The sign and the absolute value of the electrostatic potential calculated at each point of the molecular surface represent an important information about accessibility of this part of the molecular surface by other molecules (with their own surface and their own electrostatic potential calculated on it). A large area of the molecule-molecule contact (fitting) together with the opposite signs of the corresponding electrostatic potentials (over the contact area) means a perfect fitting: of the shapes and of the electrostatic characteristics. Thus, the molecular surface "colored" by the electrostatic potential provides an essential (and synthetic) information about the ability of the molecule to interact with the other ones.

The above described contributions are responsible for some specific interactions like: the ubiquitous hydrogen bonds, the coordination interaction (interaction of the lone electron pair of the electron donor with an empty orbital of the electron acceptor), and, in aqueous solutions, the hydrophobic effect (hydrocarbon, or "fatty," residues having tendency to keep together in water). The hydrogen bond is dominated by the electrostatic energy, the coordination interaction (e.g., the coordination of a positive ion by a crown ether of suitable size) by the electrostatic attraction and the valence repulsion, and the hydrophobic effect by the hydrogen bond network of the liquid water. These three types of intermolecular interaction determine the majority of the appearing supramolecular structures, typically of the size of nanometers.

In the case of two interacting molecules, the resulting molecular recognition is characterized by one of the following scenarios differing by the role played by molecular flexibility of the interacting moieties:

- Keylock (both molecules are rigid and fit together just because the complementary molecular shapes in their ground conformational states)
- Template (a higher-energy conformation of a flexible molecule makes a perfect fit with the ground conformational state of a rigid molecule, the latter serving as a template)

- Hand glove (two flexible molecules do not fit one another unless they acquire some particular higher-energy conformational states)

In the two last cases, the expense of the necessary conformational excitation is more than compensated by the resulting interaction energy of the two moieties.

Rules of the Nanoscale

Note that all three types of the molecular recognition stem from the same general principle that a molecule with a convex surface (guest) may find a perfect fitting with a certain molecule with a concave surface (host). This phenomenon, which may now look even trivial, has been discovered as late as in the 1960s of the twentieth century. The very reason of why it was so late is that chemists had to learn how to synthesize large molecules and only such molecules can have a concave surface. Another reason is that the progress in organic chemistry has been limited in the past by the possibilities of computational chemistry, which finally enabled chemists to plan the complex molecular geometries and check whether they fit as expected.

This convex-concave paradigm represents a basis of the molecular self-assembling and self-organization. This type of interaction has some unique features, which make possible precise human operation in the nanoscale by planning the molecules to be synthesized using the tools of computational chemistry:

- A convex molecule of a particular shape may fit a *specific* concave one, much more precisely than two convex molecules are able to do (*molecular recognition*). This may be used to get a sure and direct contact of some particular functional groups to make a particular chemical reaction to proceed with a high reaction yield.
- Many interatomic contacts distributed over a *large convex-concave contact area* may assure a sufficiently strong intermolecular binding.
- A concave reactant (“host”) means also preventing some potential and unwanted reactants (“guest molecules”) to enter the cavity, just because these molecules do not recognize each other (a large energy penalty by the valence repulsion, a result of ill fitting).
- A sufficiently weak intermolecular binding in order to allow for an easy reorganization of the complex, if needed.
- Leaving the reaction cavity by the products, because they may have not enough space there.

Power of Computer Experiments

In experiments we always see quantities averaged over all molecules in the specimen, all orientations allowed, all states available for a given temperature, etc. In some experiments, we are able to specify the external conditions in such a way

as to receive the signal from molecules in a given state. Even in such a case, the results are averaged over molecular vibrations, which introduce (usually quite small) uncertainty for the positions of the nuclei, close to the minimum of the electronic energy (in the adiabatic or Born-Oppenheimer approximations).

This means that in almost all cases, the experimenters investigate molecules close to the minimum of the electronic energy (minimum of PES). What happens to the electronic structure for other configurations of the nuclei is a natural question, sometimes of great importance (e.g., for chemical reactions). Only computational chemistry opens the way to see what would happen to the energy and to the electronic density distribution if:

- Some internuclear distances increased, even to infinity (dissociation).
- Some internuclear distances shortened, and the shortening may correspond even to collapsing the nuclei into a united nucleus, or approaching two atoms which in the minimum of PES form or do not form a chemical bond. This allows us to investigate what happens to the molecule under a gigantic pressure.
- Some nuclei changed their mass or charge (beyond what one knows from experiment).
- We apply to the system an electric field, whose character is whatever we imagine as appropriate. Sometimes such a field may mimic the influence of charge distributions in neighboring molecules.

This makes from computational chemistry a quite unique tool allowing to give the answer about the energy and electronic density distribution (bond pattern) for any system and for any deformation of the system we imagine. This powerful feature can be used not only to see, what happens for a particular experimental situation, but also what would happen if we were able to set the conditions much beyond any imaginable experiment. This uniqueness is probably the most exciting feature of computational chemistry for it carries us to discoveries everyday.

Concluding Remarks

What counts in computational chemistry is looking at molecules at various scales (using various models) and comparing the results for different molecules. If one could only obtain an exact picture of the molecule without comparing the results for other molecules, we would be left with no chemistry. The power of chemistry comes from analogies and similarities, as well as from trends rather than from the ability of predicting properties. Such ability is certainly important for being efficient in any particular case, but predicting by computation does not mean understanding. We need computers with their impressive speed, capacity, and possibility to give us precise predictions, but also we need a language to speak about the computations, a model which simplifies the reality, but allows us to understand what we are playing with in chemistry.

Acknowledgements The author is very indebted to Professor Leszek Z. Stolarczyk, for joy to be with him, discussing all exciting aspects of chemistry, science, and beyond, a part of them included in the present paper.

Bibliography

- Bader, R. F. W. (1994). *Atoms in molecules. A quantum theory*. Oxford: Clarendon Press.
- Bloch, F. (1928). Über die Quantenmechanik der Electonen in Kristallgittern. PhD Thesis. University of Leipzig.
- Born, M., & Oppenheimer, J. R. (1927). Zur Quantentheorie der Molekeln, *Annalen der Physik*, 84, 457.
- Boys, S. F., Cook, G. B., Reeves, C. M., & Shavitt, I. (1956). Automatic Fundamental Calculations of Molecular Structure. *Nature*, 178, 1956.
- Brown, G. E., & Ravenhall, D. G. (1951). On the interaction of two electrons. *Proceedings of the Royal Society A*, 208, 552.
- Cotton, F. A. (1990). *Chemical applications of group theory* (3rd ed.). New York: Wiley.
- Dirac, P. A. M. (1928). The quantum theory of the electron. *Proceedings of the Royal Society of London*, A117, 610; (1928) *ibid.*, A118, 351.
- Feynman, R. P. (1939). Forces in Molecules. *Physical Review*, 56, 340.
- Fock, V. (1930). Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Zeitschrift für Physik*, 61, 126; (1930) „Selfconsistent field“ mit Austausch für Natrium. *ibid.*, 62, 795.
- Fukui, K., & Fujimoto, H. (1968). An MO-theoretical interpretation of nature of chemical reactions. I. Partitioning Analysis of Interaction Energy. *Bulletin of the Chemical Society of Japan*, 41, 1989.
- Hartree, D. R. (1927). The wave mechanics of an atom with a non-coulomb central field. Part I. Theory and methods. *Proceedings of the Cambridge Philosophical Society*, 24, 89.
- Heitler, W., & London, F. W. (1927). Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik. *Zeitschrift für Physik*, 44, 455.
- Hellmann, H. (1937). *Einführung in die Quantenchemie*. Leipzig: Deuticke.
- Hund, F. (1927). Zur Deutung der Molekelspektren. I- III. *Zeitschrift für Physik*, 40, 742; (1927) *ibid.*, 42, 93; (1927) *ibid.*, 43, 805.
- Hylleraas, E. A. (1929). Neue Berechnung der Energie des Heliums im Grundzustande, sowie des tiefsten Terms von Ortho-Helium. *Zeitschrift für Physik*, 54, 347.
- James, H. M., & Coolidge, A. S. (1933). The ground state of the hydrogen molecule. *Journal of Chemical Physics*, 1 825.
- Jeziorski, B., & Kołos, W. (1977). On Symmetry Forcing in the Perturbation Theory of Weak Intermolecular Forces. *International Journal of Quantum Chemistry*, 12, 91.
- Kato, T. (1957). On the eigenfunctions of many-particle systems in quantum mechanics. *Communications on Pure and Applied Mathematics*, 10, 151.
- Kołos, W., & Roothaan, C. C. J. (1960). Accurate electronic wave functions for the H₂ molecule. *Reviews of Modern Physics*, 32, 219.
- Koliński, A., & Skolnick, J. (2004). Reduced models of proteins and their applications. *Polymer*, 45, 511.
- Koopmans, T. C. (1933/1934). Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. *Physica*, 1, 104.
- Łach, G., Jeziorski, B., & Szalewicz, K. (2004). Radiative corrections to the polarizability of helium. *Physical Review Letters*, 92, 233001.
- Lee, B., & Richards, F. M. (1971). The interpretation of protein structures. Estimation of static accessibility. *Journal of Molecular Biology*, 55, 379.

- Pestka, G., Bylicki, M., & Karwowski, J. (2008). *Dirac- Coulomb equation: Playing with artifacts*. In S. Wilson, P. J. Grout, J. Maruani, G. Delgado-Berrio, & P. Piecuch (Eds.), *Frontiers in quantum systems in chemistry and physics* (pp. 215–238). Heidelberg: Springer.
- Piela, L. (2002). Deformation methods of global optimization in chemistry and physics. In P. M. Pardalos & H. E. Romeijn (Eds.), *Handbook of global optimization* (Vol. 2, pp. 461–488). Dordrecht/Boston: Kluwer Academic Publishers.
- Piela, L. (2014). *Ideas of quantum chemistry* (2nd ed.). Amsterdam: Elsevier.
- Roos, B. O. (1972). A new method for large-scale CI calculations. *Chemical Physics Letters*, *15*, 153.
- Schrödinger, E. (1926). Quantisierung als Eigenwertproblem. I-IV. *Annalen der Physik*, *79*, 361; (1926) *ibid.*, *79*, 489; (1926) *ibid.*, *80*, 437; (1926) *ibid.*, *81*, 109.
- Slater, J. (1930). Cohesion in monovalent metals. *Physical Review*, *35*, 509.
- Woodward, R. B., & Hoffmann, R. (1965). Selection rules for sigmatropic reactions. *Journal of the American Chemical Society*, *87*, 395, 2511.

Valeri Poltev

The ultimate justification for the many severe approximations and assumptions made in the present work comes from the fact that the agreement between the simple calculations and the available experimental data is as good as it is. N. L. Allinger, J. Amer. Chem. Soc., **81**, 5727, 1959

Contents

Introduction	22
Foundations and General Scheme of Molecular Mechanics. Atoms as Elementary Units of the Matter	23
General Expression and Terms of Molecular System Energy	24
Intramolecular Contributions to Molecular System Energy	25
Intermolecular and Nonbonded Intramolecular Interactions	26
General Remarks on Molecular Mechanics, Its Accuracy, and Applicability	27
A Bit of History: The "Precomputer" and Early Computer-Aided MM Calculations	29
First MM Applications to Three-Dimensional Structure and Thermodynamics of Organic Molecules	30
The Role of Molecular Crystal Study on the First Steps of Molecular Mechanics	31
Molecular Mechanics on the First Steps of Molecular Biology: Molecular Mechanics and Protein Physics	33
Molecular Mechanics on the First Steps of the Biophysics of Nucleic Acids	37
The Problems and Doubts of Further Development of the MM Approach	41
Approaches to Improvement of MM Formulae and Parameters	41
Various Schemes of Water Molecule in MM Calculations	44
Modern Molecular Mechanics Force Fields: Description, Recent Development, and Applications	47
Allinger's Force Fields and Programs	48
Merck Molecular Force Field (MMFF94)	49
The Force Fields and Programs Designed by Scheraga and Coauthors	50
Force Fields and Programs Developed by Kollman and Coauthors. AMBER	52
Other Popular Force Fields and MM Software. CHARMM, OPLS, GROMOS	55
Conclusions	58
Bibliography	59

V. Poltev (✉)

Autonomous University of Puebla, Puebla, Mexico
e-mail: poltev@fcfm.buap.mx; vpoltev@yahoo.com

Abstract

A short survey of the general principles and selected applications of molecular mechanics (MM) is presented. The origin of molecular mechanics and its evolution is followed starting from “pre-computer” and the first computer-aided estimations of the structure and potential energy of simple molecular systems to the modern force fields and software for the computations of large biomolecules and their complexes. Analysis of the current state of physicochemical study of biological processes suggests that MM simulations based on empirical force fields have an ever-increasing impact on understanding the structure and functions of biomolecules. The problem of “classic mechanics” description of essentially quantum properties and processes is considered. Various approaches to a selection of force field mathematical expressions and parameters are reviewed. The relation between MM simplicity and “physical nature” of the properties and events is examined. Quantum chemistry contributions to MM description of complex molecular systems and MM contribution to quantum mechanics computations of such systems are considered.

Introduction

Molecular mechanics (MM) deals with the classical (“mechanistic,” i.e., via Newtonian mechanics) description of molecular and supramolecular systems. The simplified assumptions and approximations enable one to use MM for wide applications to various systems, starting from simple low-molecular-weight molecules (such as numerous hydrocarbons) to large biomolecular complexes (such as those of proteins, nucleic acids, and membrane fragments) or material assemblies of many thousands of atoms. During a period a bit longer than a half of a century, the number of publications using this approach roused by many orders of magnitude. It was rather easy to follow nearly all the works related to MM calculations in the 1960s and in the early 1970s of the last century (and the author of this chapter tried to do so); today it is very difficult to follow all the papers related to the MM approach even in the rather restricted area of application (e.g., to biopolymer structure and function). Thus it is impossible to consider all the applications of MM in a short survey, and the choice of the material for this chapter is “a matter of taste” of the author who refers readers to other monographs and chapters for a more general description of the modern state and/or selected aspects of extended applications of MM; for general references see, e.g., monographs (Lewars 2011; Allinger 2010; Ramachandran et al. 2008; Cramer 2004; Leach 2001). We will follow the evolution of MM from the first “precomputer” and early computer-aided (i.e., before the era of personal computers) works to modern complex simulations impossible without supercomputers.

The MM method is used now not only in theoretical and computational works but also as a part of experimental studies (e.g., many X-ray and NMR-derived structures of proteins, nucleic acids, and their complexes with other molecules deposited in

Protein Data Bank (Berman et al. 2003) and Nucleic Acid Data Bank (Berman et al. 1992) are the results of MM refinements). The MM-like semiempirical terms are widely used now in some quantum mechanics-based studies. The systematic inclusion of such terms started from the addition of the dispersion term to the standard density functional theory computations, for creating the so-called DFT-D method (see, e.g., Grimme 2004; Antony and Grimme 2006; Jurecka et al. 2007, and references herein for the first DFT-D results). Recent developments and applications of different DFT-D methods (DFT-D2, DFT-D3) demonstrated their predictive power for biomolecular systems (see, e.g., Grimme 2011; Brandenburg and Grimme 2014). Since the beginning of this millennium, MM-like terms are frequently included into so-called enhanced semiempirical quantum mechanics methods (see, e.g., Yilmazer and Korth (2015) for a recent review). Both dispersion (D) and hydrogen-bond (H) correction terms are used in such methods usually referred to as SQM-DH methods.

The rapid extension of MM simulations is a result of development of new branches of natural science related to molecular biology with its numerous applications, to directed synthesis of new substances with desirable properties (including new drugs and new materials for the industry), and many other areas of so-called nanoscience. The MM simulations enable one to perform a preliminary selection of the compounds with desirable properties before experimental testing, and even before chemical synthesis itself. Sometimes it is impossible to rationalize completely experimental results without the construction of atom-level mechanistic model. In many cases, the model cannot be derived from experimental data only, e.g., the accuracy of X-ray diffraction patterns for biopolymer fragments is not sufficient for the extraction of precise atom coordinates due to a complexity of the system and inherent irregularities of atom positions (e.g., delocalization of water molecules). In other cases, the complex molecular system cannot be characterized by a single three-dimensional structure but by a set of conformations. MM simulations enable one to construct habitual atom-level molecular models of sufficiently complex systems favorable from energy viewpoint. These models can be visualized by computer graphics programs and are ready to use for further investigations and refinements (e.g., by quantum mechanics methods or by new experimental approaches).

Foundations and General Scheme of Molecular Mechanics. Atoms as Elementary Units of the Matter

The modern MM can be considered as an extension of simple atom and bond representation of molecules and their complexes in classical chemistry. In the beginning of the 1950s, scaled paper images and hand-made primitive models of atoms and bonds from wood and wire enabled subtle scientists to construct the first successful models of proteins and DNA. The pioneering papers of Nobel Prize authors can be mentioned in this relation (Pauling et al. 1951; Watson and

Crick 1953). Stick, ball-and-stick, and space-filling plastic models as well as their computer images are widely used until now for both teaching and investigation.

General Expression and Terms of Molecular System Energy

The MM approach is a generalization and quantitative representation of these models via mathematical expressions. A set of such expressions together with numerical coefficients is usually referred to as the force field; the MM method itself is frequently referred to as the force field method. Rigid balls of computer graphics or plastic (metal, wood, etc.) models are replaced by “soft” ones, and hard sticks are replaced by springs. “Hard Sphere” approximation can be considered as simplified (in many cases oversimplified) MM approach. Nevertheless, very important results were obtained via this approximation in the middle of the last century; we will mention some of them below.

The modern MM method and most of force fields mathematical expressions are the results of the efforts of many researchers for nearly half a century, trials and errors of both well-known and widely cited authors, as well as of those who contributed to specific parts of nonbonded energy computations for a selected class of molecules. In this section, we will describe general scheme and the main common assumptions of the MM method leaving any references to specific papers and contributions of the researchers for next sections. The main subject of this survey is rather simple MM scheme suggesting that the systems (molecules or their complexes) consist of atoms, each atom being simulated as a single point particle. Most computations for biologically important molecules use the so-called Class 1 model. More complicated schemes considering additional centers of interactions, additional terms, and more exact expressions (so-called Class 2 and Class 3 MM models) will be mentioned briefly. Molecular mechanics describes molecules in terms of “bonded atoms” (atoms in molecule), their positions are distorted from some idealized (equilibrium) geometry due to nonbonded interactions with the atoms of the same molecule and of other molecules. The MM approach implicitly uses the starting approximation for both classical and quantum computational chemistry schemes, namely, Born–Oppenheimer approximation. MM does not include electrons explicitly, but different configurations of electron shells of the same chemical element’s atoms are implicitly considered in force field parameters. Some authors consider not only atoms but lone pairs of electrons, chemical bonds, and other points in a molecule as centers of interactions. We will mention such “extension” of Class 1 models below.

The first principal approximation of MM method is the additivity for several levels of calculations, namely, the additivity of energy terms responsible for the contributions of different physical nature, and the additivity of contributions of the atoms. The simple (“minimalist”) representation of the potential energy of a molecule or molecular complex as a function of atom coordinates, R , can be represented by a sum of four main terms each term being a sum of many contributions (Eq. 1).

$$\Delta E(\mathbf{R}) = \sum E_b + \sum E_a + \sum E_t + \sum E_{nb} \quad (1)$$

The summing up in the terms is over all the chemical bonds (the first term), valence angles (the second term), torsion angles (the third term), and all pairs of atoms nonbonded to each other or to common third atom (the last term). The energy terms depend on mutual positions of atoms and on the adjustable constants (parameters); the parameters are suggested to be transferable between atoms and molecules of the same type. The transferability of the force field parameters is the second important assumption of MM, and the number of atom types of the same chemical element depends on particular force field.

Intramolecular Contributions to Molecular System Energy

The first and the second terms (sums of Eq. 1) refer to the energy changes due to variations of bond lengths, stretching (Eq. 2), and bond angles, bending (Eq. 3). These terms are usually modeled as harmonic potentials centered on equilibrium values of bond lengths and bond angles, respectively, i.e., “simple Hooke’s law” dependences are used.

$$E_{b_i} = k b_i (l_i - l_i^0)^2 \quad (2)$$

$$E_{a_i} = k a_i (\alpha_i - \alpha_i^0)^2 \quad (3)$$

In these equations, l_i and α_i , the current values of bond lengths and angles; parameters l_i^0 and α_i^0 , equilibrium values for bond lengths and angles; $k b_i$ and $k a_i$, stretch and bend force constants, respectively. These adjustable parameters depend on types of atoms forming the bond or the valence angle. Some force fields may also contain cubic and higher-order contributions to these terms, or sometimes more flexible Morse potential can be used instead, as well as “cross terms” can be included to account for correlations between stretch and bend components. In the latter case, the terms depending on both bond length and angle variations are added. The third sum of the expression (Eq. 1) refers to the changes of torsion energy; it is responsible for interactions of electron shells of two atoms A and D (and of two bonds A–B and C–D) which are connected through an intermediate chemical bond B–C.

$$E_{t_i} = k t_i (1 + \cos (n_i \phi_i - \delta_i)) \quad (4)$$

This periodic function (integer n_i being the periodicity) contains the current value ϕ_i of the angle i of the rotation around B–C chemical bond (the angle between the two planes defined by atoms A, B, and C and by B, C, and D) and three parameters ($k t_i$, n_i , and δ_i ; n_i is multiplicity and δ_i is the phase angle) for each type of torsion (for each combination of four neighbor atoms of the molecule). These parameters can be estimated from experimental data on the structure and properties of the

molecule considered and of related molecules, then they (as well as I_i^0 , α_i^0 , kb_i , and ka_i of Eqs. 2 and 3) should be adjusted by trial computations. Many force fields include terms responsible for “improper” torsions or out-of-plane bending, i.e., terms related to four atoms not forming consecutive chemical bonds, which function as correction factors for out-of-plane deviations (for example, they can be used to keep aromatic rings planar). These terms can be expressed via harmonic potentials like those for bond stretching and valence angle bending. Cross terms depending on both torsion angle and bond length or valence angle are added in some force fields. Three terms mentioned above are positive for practically all states of all the molecular systems (or equal to zero when the term corresponds to equilibrium state for contributed bond, angle or torsion).

Intermolecular and Nonbonded Intramolecular Interactions

The last sum of Eq. 1 refers to so-called nonbonded interactions, E_{nb} , of all the atom pairs not bonding to each other or to the same third atom, Eq. 5. Each atom-atom term is usually represented by a sum of electrostatic, Coulomb (the first term of Eq. 5), and van der Waals (the second and the third terms of Eq. 5) interactions.

$$E_{ij}(r_{ij}) = Kq_iq_j/r_{ij} - A_{ij}/r_{ij}^6 + B_{ij}/r_{ij}^{12} \quad (5)$$

This equation contains r_{ij} , the current distance between i and j atoms; q_i and q_j , effective atom charges; A_{ij} and B_{ij} , adjustable parameters responsible for dispersion (London) attraction and short-range repulsion interactions respectively. The atomic charges are usually derived using calculations via various quantum chemistry methods, effective dielectric constant implicitly accounting for surrounding can be used (this value may be distant dependent). The A_{ij} and B_{ij} coefficient can be preliminarily estimated via equilibrium inter-atomic distance and energy values at equilibrium for neutral pairs of atoms (ρ_{ij} and ϵ_{ij} respectively) and followed by the adjustment to reference experimental data. Most of the early force fields used for description of van der Waals interactions are Buckingham (6-exp) potential instead of Lennard-Jones (6-12) potential in Eq. 5. The total expression for nonbonded interaction term are usually referred to as (1-6-12) or (1-6-exp) potential relating to the dependency of the terms on the inter-atom distance. The Buckingham potential is more flexible (it has three adjustable parameters instead of two for 6-12 potentials for each atom pair type) and has more physics justifications for the range of experimentally observed distances (due to exponential dependence of electron wave functions on the distance from nuclei), but it is less convenient for computations. It has a maximum at short distance and then trend to negative infinite value. Majority of the modern force fields utilize 6-12 expressions for description of van der Waals interactions, the total atom-atom potential being referred to as 1-6-12 one. Some force fields substitute 6-12 potential by 10-12 one for the interactions of hydrogen atoms of hydrogen bonds in order to describe more sharp distance dependence in the most important area of energy minimum corresponding to H-bond formation (referred

to as 1-10-12 potential). More complex expressions (including those dependent on the angles between two straight lines connecting three atoms of H-bond) were used for H-bond description in some early potential sets. The nonbonded terms of the intramolecular energy related to 1-4 interactions (i.e., the interactions between atoms in a molecule separated by three chemical bonds) are frequently accounted for with a coefficient less than 1 (1-4 scaling) as these interactions are already included into torsion term (Eq. 4). To reduce the number of adjustable A_{ij} and B_{ij} parameters of Lennard-Jones potential (and corresponding parameters of other potentials), the combination rules for equilibrium inter-atom distance, ρ , and minimal energy values, ϵ , for pairs of different atoms are usually applied.

$$\rho_{ij} = \rho_i + \rho_j; \epsilon_{ij} = (\epsilon_i \epsilon_j)^{1/2} \quad (6)$$

Some force fields apply the combination rules directly to the coefficients of van der Waals terms.

The calculations of potential energy via Eq. 1 are used to search for local energy minima (mutual atom positions corresponding to possible stable configurations), to construct and analyze multidimensional potential energy surfaces (PES), to follow trajectory of movement (in MD, molecular dynamics simulations), or to study averaged thermodynamic and geometry characteristics (via MC, Monte Carlo sampling) of the systems. Description of MD and MC methods is beyond the survey, while some results of these applications of MM method will be mentioned briefly.

General Remarks on Molecular Mechanics, Its Accuracy, and Applicability

The first computer (and all “precomputer”) applications of mechanistic approach to molecule conformations and interactions ignored certain energy terms (e.g., stretching, bending, torsion, or electrostatic ones). Some modern works can ignore certain terms as well in order to reduce the number of variables of energy function, e.g., considering bond lengths as constants (their changes in many cases are very small and have no influence on energy and geometry of minimal energy structures). The simplest one of such approaches considers bond lengths and valence angles as constants, ignores torsion energy (the contributions of the first three terms of Eq. 1 being zeros), and utilizes “hard sphere approach” to nonbonded interactions. This approach is a mathematical representation of plastic (or wood, iron, etc.) space filling mechanic models or their computer images. The configurations are forbidden when any two nonbonded atoms are closer to each other than a sum of van der Waals radii (these configurations have infinite positive energy), all other being allowed (with zero energy). Already this oversimplified approach enables one to obtain some important results, e.g., to reject certain configurations and even a possibility to synthesize the molecule with inevitable too close positions of nonbonded atoms. The first “Ramachandran maps” for proteins (which will be discussed in the next section) have demonstrated allowed and forbidden regions on two-dimensional plots of the

fragment of polypeptide chain. These maps were subsequently improved using more realistic MM functions or quantum mechanics calculations.

Most modern MM computations include additional terms besides those already mentioned. These terms refer to direct imposition of experimental data (e.g., NMR-derived restrains on inter-atom distances or global characteristics of the macromolecule) and to description of specific quantum effects not accounting for by standard MM force field formulae.

The complexity of mathematical expressions and the number of parameters depend on the systems considered. The problem of “which atoms pertain to the same type and which ones are of different types” is considered by the authors of specific force field and software depending on the tasks and computer resources. The atom type may depend not only on the chemical element and electron shell configuration but on neighbor atoms and on the structure of the whole molecular fragment (e.g., the carbons of six-member and five-member aromatic rings having the same three bonded atoms may be considered as pertaining to different atom types). The more broad the applications that are planned for the force field, the greater the number of atom types that should be involved and the more complex force field formulae that should be constructed. The first works which dealt with the tasks related to specific systems (e.g., the conformations of saturated hydrocarbons or peptide fragments) usually contained a few parameters; the modern force fields may contain thousands of parameters (in spite of use of combination rules mentioned above). The mathematical forms of the equations and numbers of atom types and adjustable parameters represent a compromise between the simplicity (for elaboration and for use) and the accuracy of the results obtained.

Various physical considerations can be used for preliminary estimation of mathematical expressions and parameter values (rather simplified considerations were used in Eqs. 2, 3, 4, and 5). It is important to emphasize that neither dependences nor values of parameters can be “derived” (directly calculated) from universal principles or measured by any experimental method. The stretch and bend constants (of Eqs. 2 and 3) can be evaluated using infrared spectra; equilibrium bond lengths and valence angles can be estimated from X-ray data for crystals of simple molecules. The A_{ij} coefficients of the attraction part of van der Waals interactions can be evaluated (and really were calculated and used without refinement in the first MM works) via approximate formulae for dispersion interactions; however, their exact values for a certain class of systems should be adjusted by comparison with experimental data or with the most exact quantum chemistry results after trial computations for reference set of related systems. The same is valid for other terms and their parameters. Some parameters have rather simple physical meaning and restricted areas of possible values (e.g., equilibrium distances between bonded and nonbonded atoms or barriers to rotation about the bonds); other parameters have only approximate relation to physical values (atom charges, B_{ij} coefficients of Lennard-Jones potential). As all the parameters are adjustable ones, only the values of total energy and the equilibrium geometry of the molecular system can be compared with experimental data, and consequently have the strict physical meaning, not the individual contributions or the values of the individual parameters.

As various force fields utilize different reference sets of data, the individual parameters are not transferable between different force fields even in cases where they use the same mathematical expressions. Different force fields may result in the nearly equal energy and geometry of local minima configurations but rather different values of the individual term contributions. Thus individual terms of the energy may have very approximate physical interpretation, although in some cases it is interesting to evaluate the certain energy contributions and to follow their changes for different molecular complexes and different configurations (and many researchers include these evaluations to their publications).

It is worth mentioning that preliminary consideration of MM scheme has resulted already in some doubts and objections. Generally speaking, the classical description of the essentially quantum molecular systems cannot be exact and full. Most of the terms in Eqs. 2, 3, 4, and 5 refer to the first approximation or to the first term of expansion of the corresponding interaction energy. The atoms are not points, they have dipole and quadrupole moments (not only charges), and charge distribution in a molecule is continuous. The polarization or electron delocalization interactions are not considered in the classical “minimalist” MM approach, and the contributions of three-body and four-body interactions can be essential ones. Many attempts have been undertaken to overcome these inherent difficulties of the MM method as well as to justify the assumptions and simplifications; we will consider some of these attempts below. Few remarks for justification of the main principles of MM method are described here.

The possibility of consideration of atoms as elementary subunits of the molecular systems is a consequence of Born-Oppenheimer or adiabatic approximation (“separation” of electron and nuclear movements); all the quantum chemistry approaches start from this assumption. The additivity (or linear combination) is a common approach to a construction of complex functions for a physics-based description of the systems of various levels of complexity (cf. orbital approximation, MO LCAO approximation, basis sets of wave functions, and some other approximations in quantum mechanics). The final justification of the method is good correlation of the results of its applications with the available experimental data and a possibility to predict the characteristics of molecular systems before experimental data become available. It can be achieved after careful parameter adjustment and proper use of the force field in the area of its validity. The contributions not considered explicitly in the force field formulae are included implicitly into parameter values of the energy terms considered.

A Bit of History: The “Precomputer” and Early Computer-Aided MM Calculations

The quantitative estimations of molecular properties via simple atom level mechanics representations originate from the communications of Hill (1946, 1948), Westheimer and Mayer (1946), and Barton (1948, 1950). All these papers refer to conformations of organic molecules. It is interesting to mention that mathematical

expression of potential energy suggested in the pioneering work of Hill (1946) contains common for all the modern force fields stretch and bend components (Eqs. 2 and 3 of previous section) as well as the Lennard-Jones terms of nonbonded interaction energy. Westheimer and Mayer (1946) suggested the use of exponential term for description of steric repulsion. The first calculations for selected conformations of rather simple (“medium size”) molecules, such as diphenil derivatives in paper (Westheimer and Mayer 1946), cis-decalin and steroids in the papers of Barton (1948, 1950) were performed manually or using desk calculators. Some researchers constructed “hand-made” models of steel or wood (e.g., of cyclic saturated hydrocarbons in the papers of Allinger (1959)) for careful measurements of geometry parameters. The importance of quantitative estimations of nonbonded interactions for considerations of three-dimensional structure of organic molecules was emphasized starting from the first mechanical considerations, which was clearly shown by Bartell in 1960. He illustrated the preference of “soft sphere” over “hard sphere” approach to the analysis of hydrocarbon structure and suggested one of the first (6-12) parameters for hydrocarbons (Bartell 1960). As mentioned above, rather approximate calculations clearly demonstrated the utility of MM approach to the problems of organic chemistry as well as the need for further extensive computations and searching for more reliable parameters. We will refer to all these quantitative theoretical considerations of molecular properties as MM, not depending on the use of this term by the authors or on the method of estimation of different types of interactions.

Rapid expansion of MM method starting from the 1960s was provoked by an introduction of computers into all the branches of natural science. In this section, we will briefly consider some examples of the first computer-aided applications of MM method to three research areas, namely, the physical organic chemistry (these works can be considered as a continuation of the “precomputer” papers mentioned above), the structure and properties of molecular crystals, and the interactions and conformations of biopolymers.

First MM Applications to Three-Dimensional Structure and Thermodynamics of Organic Molecules

The first paper on computer study of organic molecule conformations was related to saturated hydrocarbons by Hendrickson (Hendrickson 1961). The angle bending, torsion, and (6-exp) van der Waals contributions to the conformation energy were taken into account, while constant values were assigned to bond lengths. The computer calculations of cyclo-alkanes containing 5, 6, and 7 carbon atoms enabled him to consider various conformers and to reproduce and rationalize the experimental data. During the 1960s and the beginning of the 1970s such computations were performed by several groups of investigators. Hendrickson (1962, 1973, and references herein), Allinger et al. (1971), and Allinger and Sprague (1973) extended MM approach to more complex hydrocarbons including those with delocalized electronic systems. Allinger’s computations took into account bond stretching in

addition to terms used by Hendrickson. Wiberg (1965) was the first researcher who suggested use of the Cartesian coordinates of atoms (instead of “internal coordinates,” i.e., bond lengths, bond angles, torsions) for nonbonded (“strain”) energy minimization. Such approach enables researchers to use the same scheme and the same software to study completely different systems. Most modern MM programs use such an approach. The electrostatic term has not been included in the papers mentioned in this paragraph as hydrocarbons are nonpolar molecules; it was introduced later when more broad sets of molecules became to be considered.

The most important results of early MM computations of organic compounds can be illustrated by Engler et al.’s paper (1973) titled “Critical Evaluation of Molecular Mechanics.” The calculations for various hydrocarbons have been performed using two rather different force fields, their own and that of Allinger et al. (1971). The two force fields have substantially different parameters as different sets of experimental characteristics were used for parameter adjustment. It results in significant difference of separate terms of the energy, which may vary by several times or even be positive for one force field and negative for another one. Nevertheless, total energies and relative values for various conformations calculated by two force fields are rather close to each other for most of the compounds considered. Two phrases of the paper’s abstract can be considered as characteristics of the whole situation with the MM approach for that time: “Most of the available data are reproduced with an accuracy rivaling that achieved by the experimental methods” and “The molecular mechanics method, in principle, must be considered to be competitive with experimental determination of the structures and enthalpies of molecules.”

Since the abovementioned publications, the theoretical conformational analysis using MM force fields has become an inherent part of physical studies of organic molecules. It became clear that in spite of difficulties with the parameter choice, the MM calculations are not only a useful tool to rationalize experimental observations, to reproduce and to predict the structure and energy characteristics of “medium size” organic molecules with experimental accuracy, but they can also help in elaboration of pathways of chemical synthesis. Different mathematical expressions and parameters transform into *force fields*, i.e., to complete and verified sets of the formulae and constants and ready-to-use computer programs. One of the first such force fields and programs was MM1 program based on the paper by Allinger and Sprague (1973), which then was followed by MM2, MM3, and MM4 force fields and software for broad class of organic molecules. The inclusion of new atom types and new terms of energy as well as parameter refinement continued till last years, and we will discuss these program sets in the next sections together with other modern force fields and software.

The Role of Molecular Crystal Study on the First Steps of Molecular Mechanics

The first publications on MM approach to molecule conformations mentioned above deal with *intramolecular* interactions. The first works on such approach to

intermolecular interactions belong to one of the pioneers and founders of MM, A. I. Kitaigorodsky (other spellings, Kitaigorodskii, Kitajgorodskij, Kitaygorodsky), though he (and all the researchers at that time) did not use the term “Molecular Mechanics.” Nearly at the same period (the 1950s) as Hill and Westheimer, usually mentioned as initiators of MM approach, he suggested to consider mechanical models of molecular systems quantitatively via mathematical expressions. In the 1950s, he foresaw the applications of the method to various problems of physics and physical chemistry of organic and biological molecules, including the problems of crystal structure, adsorption, and conformational transitions. We refer to the most frequently cited works of Kitaigorodsky (1960, 1961, 1973). Kitaigorodsky used structure and thermodynamic data on the crystals of organic molecules for adjustment of parameters of atom-atom potential functions for calculations of non-bonded interaction energy. This term of the total energy is dominant for molecular crystals. The method of atom-atom potentials was suggested as a generalization of the “principle of close packing” of molecules in molecular crystals he discovered earlier; his book (Kitaigorodski 1959) describing this principle is still widely cited. Later he demonstrated that this principle is a consequence of more general atom-atom approach to potential energy of molecular crystals. Kitaigorodsky was the first researcher who suggested to consider the interactions of nonbonded atoms in a molecule and of such atoms of different molecules via the same mathematical expressions and parameters (applications of MM approach to biopolymers would be impossible without this suggestion). The studies of molecular crystals via MM approach enabled one to derive potential functions for nonbonded interactions and to test the fidelity and accuracy of the approach itself using extensive sets of available quantitative data on structure and energy of molecular crystals.

Some energy terms were ignored in the first MM works on conformations of organic molecules, but it was impossible to predict a priori what types of contributions to intramolecular interactions energy (sums in Eq. 1) can be reasonably neglected for specific types of molecules. Considering crystal structure of nearly rigid molecules (e.g., aromatic rings) it is possible to ignore (at least in the first approximation) all the terms but the last one in Eq. 1. This assumption is just a consequence of examination of the geometry of series of related molecules in crystals (e.g., aromatic hydrocarbons consisting six-atom rings); all of them have nearly equal values of bond lengths and angles not depending on molecular complexity and the type of crystal packing. For hydrocarbons, it is possible (again at least for the initial studies) to ignore electrostatic contributions as well (the molecules have no dipole moment, and all the estimations of atom charges result in small values, less than 0.1 of electron charge). Kitaigorodsky suggested step-by-step selection of potential functions starting from molecules of nearly neutral atoms of two types (hydrocarbons) and introducing next atom types one-by-one for selecting next parameters. His pioneering works on molecular crystals were followed by more extensive computations of other authors. Some of these works were inspired by his ideas and followed his methodology (e.g., first papers of Williams (1966, 1967)), other researchers performed computations using different expressions for nonbonded potentials and additional terms of the energy (e.g., Lifson and Warshel

1968; Warshel and Lifson 1970). The common set of (6-exp) parameters for C...C, C...H, and H...H intermolecular interactions was obtained by Williams (1967) from energy and structure calculations for crystals of both aromatic and aliphatic hydrocarbons. Warshel and Lifson (1970) derived “consistent force field” parameters for description of both intermolecular and intramolecular interactions in crystals. This set contains both parameters for van der Waals interactions and other terms of energy mentioned in the previous section. Mason and coauthors (e.g., Craig et al. 1965; Rae and Mason 1968; Mason 1969) used combined approach to calculations of intermolecular interactions in crystals, namely, repulsive terms were calculated at atom-atom level, while attractive ones were considered on bond-bond level. Such approach and other changes of simple atom-atom scheme are not convenient for computations, and we would like to cite a phrase from Mason’s paper comparing his and Kitaigorodsky’s methods: “Although the representation of the attractive potential by spherically symmetric atom-atom interactions cannot be justified theoretically, it has been outstandingly successful in predicting the properties of crystals” (Mason 1969). The calculations for molecular crystals became a part of the selection and test of the parameters during elaboration of nearly all modern force fields. Later MM calculations for molecular crystals enable researchers to examine the approximations accepted in force field elaboration and to derive force field versions with additional terms of energy (including polarization and atom centers, e.g., Williams and Weller (1983)).

Molecular Mechanics on the First Steps of Molecular Biology: Molecular Mechanics and Protein Physics

Development of molecular biology and biophysics in the 1960s required a quantitative consideration of the conformations and interactions of proteins, nucleic acids, and biomolecules in general. The problems of biological importance were to rationalize the structure and conformational properties of the proteins and nucleic acids, namely, to understand the contributions of the subunits and to construct the models of the most favorable and biologically important conformations of the fragments. We already mentioned in previous section that the first successful models of regular structures of proteins and DNA duplex have been constructed using “hand-made” fragments of paper, wood, wire in the beginning of the 1950s. For understanding, explanation, and prediction of molecular level mechanisms of biopolymer functions, it was necessary to work out the method of quantitative simulation of the three-dimensional structure and properties of biomolecules. The first molecular mechanics considerations of biopolymer fragments were performed in 1960s. We will follow briefly such studies for proteins in a few paragraphs below.

The general problems on proteins which can be in principle solved via MM simulations are (1) the construction of three-dimensional structure of the macromolecule and prediction of the pathways of “protein folding” using restricted experimental data (ideally, the primary structure only); (2) the refinement of experimental structure (X-ray diffraction patterns usually do not supply us with information sufficient

for precise atom coordinate assignment). The whole problem of the proteins functioning cannot be solved via MM only as chemical reactions are beyond the MM approach. Changes in electronic structure in processes that involve bond-breaking and bond-forming, charge transfer, and/or electronic excitation require quantum mechanics considerations. Nevertheless, MM approach is useful for the problems of enzyme-substrate complex formation and of molecular recognition which are crucial for the protein functions.

First of all, two important results on protein structure obtained before classical MM computations should be mentioned. The atom-level structures of α -helical and β -sheet fragments of polypeptide chains have been designed by Linus Pauling and Robert Corey in 1951 using “hand-made” models of wood (Pauling and Corey 1951; Pauling et al. 1951). Such regular structures are the intrinsic parts of the majority of the proteins. The success in the construction of these first models of the regular protein fragments (as well as of the DNA duplex) primarily depended on correct subunits geometry and a potential to predict the correct scheme of hydrogen-bond formation, all other contributions to intramolecular interactions being of secondary importance. The regular structures imply H-bond formation between N-H donors and C=O acceptors of the peptide groups of the same chain for α -helix, and of other chain or of distant parts of the same chain for β -sheet. As it was pointed in more than half of century (Eisenberg 2003), “In major respects, the Pauling-Corey-Branson models were astoundingly correct, including bond lengths that were not surpassed in accuracy for >40 years.”

The second important “precomputer” result refers to construction via hard-sphere models of two-dimensional map for possible (and impossible) conformations of the dipeptide unit, the so-called Ramachandran plot (Ramachandran et al. 1963). The polypeptide backbone consists of repeating peptide group connected via C_α atoms, the latter being connected to hydrogen atom and amino-acid residue (the first atom of the residue is designated as C_β atom). The peptide group is practically planar, and the only single bonds in the peptide chain are those between C_α atom and two neighbor peptide groups (C_α -N-H of one group, the torsion angle for rotation about this bond is designated as Φ , and C_α -C=O of the next group, the torsion angle for rotation about this bond is designated as Ψ). Using mathematical expressions for atom coordinates suggesting fixed bond lengths and bond angles, Ramachandran et al. (1963) have constructed the two-dimensional Φ - Ψ maps with two sets of van der Waals atom radii, “normally allowed” and “outer limit,” i.e., normal and shortened radii. The main part of the maps for all the amino-acid residues except glycyl corresponds to the conformations forbidden due to shorten atom-atom contacts, while α -helix and β -sheet fall into the allowed regions. For glycyl residue, the allowed regions were considerably more extended. It took about half a year to construct these maps using a desk calculator (Ramachandran 1990); later such maps were computationally constructed using various force fields and quantum mechanics methods; superposition of Φ - Ψ combinations corresponding to experimentally or computationally constructed three-dimensional structure of proteins and peptides on the Ramachandran plot helped to check and rationalize the protein models nearly for half a century. The consideration of dependence

of Ramachandran plot on amino acid residue via hard sphere approach enables Leach et al. (1966a) to evaluate semi-quantitatively the parts of two-dimensional map corresponding to allowed conformations for different amino-acid residues (from about 50 % of all conceivable conformations for glycyl to only 5 % for valyl). The evaluation of steric restrictions emphasizes their important role as a determinant in protein conformation; the consideration of α -helices demonstrated that the preference of the right-handed ones in comparison to left-handed helices is due essentially to interactions of the C_{β} atom of the side chains with atoms in adjacent peptide units of the backbone (Leach et al. 1966b). The “hard sphere” approach was applied to search for allowed conformations of cyclic oligopeptides, for example (Némethy and Scheraga 1965).

The main conclusion of the “hard sphere” works was that steric effects are one of the most important factors in determining polypeptide conformations. “Many conformations of a polypeptide can be classed as energetically unfavorable without consideration of other kinds of interactions; however, the method breaks down to the extent that it cannot discriminate between those conformations which are sterically allowed . . . The contribution that this method has made is that more than half of all the conceivable polypeptide conformations are now known to be ruled out by steric criteria alone” (Scott and Scheraga 1966a). Nearly at the same time (the middle of 1960s) as abovementioned hard sphere considerations of peptides, the first applications of MM formulae to protein structure were started using “soft atoms,” and the first parameters of potential functions suitable for peptide subunits calculations were suggested. The first such works were published by three groups of researchers, namely, those of De Santis and Liquori (De Santis et al. 1965), Flory (Brant and Flory 1965), and Scheraga (Scott and Scheraga 1966a). All these works were preceded by the publications of the authors related to synthetic polymers. The bond lengths, valence angles, and planar peptide group in all the early studies of polypeptides were fixed.

De Santis et al. (1963) have carried out the calculations of van der Waals term of the energy using different potentials for regular conformations of linear polymers as functions of torsion angles of monomer units. The deepest minima of the conformational diagrams were found very near to the experimental structures, as obtained by X-ray fiber diffraction methods, for a series of polymers investigated including polyethylene, poly(tetrafluoroethylene), poly(oxy-methylene), and polyisobutylene. When the calculations were extended to the polypeptides (polyglycine, poly-L-alanine, and poly-L-proline), good agreement with the experiments was obtained as well (De Santis et al. 1965). After nearly 30 years De Santis wrote “While other contributions to the conformational energy are included in the calculations, the dominant role of the van der Waals interactions remains well established as the main determinant of the conformational stability of macromolecules” (De Santis 1992).

Brant and Flory (1965) have carried out the calculations on peptide unit energy in which they included torsion potentials, van der Waals (6-exp) interactions, and dipole-dipole interactions between the permanent dipole moments of near-neighbor peptide groups. The conformational energies were used by Brant and Flory

in statistical mechanical calculations of the mean-square unperturbed end-to-end distance of various polypeptide chains. They concluded that it was necessary to include the electrostatic interactions in the energy calculations to obtain agreement between the calculated and experimentally determined chain dimensions and thus that these interactions are important in determining polypeptide conformations.

The first publication of Scheraga's group related to "soft atoms" MM simulations related to proteins (on helical structures of polyglycine and poly-L-alanine (Scott and Scheraga 1966a)) included besides contributions considered in paper of Brant and Flory (1965) explicit accounting of hydrogen bonds and interactions of atom charges of peptide group. Rather complex expression for dipole-dipole energy of interactions between peptide groups used by Brant and Flory (1965) was substituted by Coulomb atom-atom terms, and (6-exp) expressions for van der Waals interactions were substituted in this paper and subsequent publications by (6-12) potentials, but additional complicated expression for description of hydrogen bond energy was added. The method and parameters for calculation of torsion and van der Waals energies were earlier developed by Scott and Scheraga (1966b) for normal hydrocarbons and polyethylene. It is interesting to note that both groups of authors (those of Flory and of Scheraga) applied without further refinement approximate Slater-Kirkwood equation for calculation of coefficient for attractive term of van der Waals energy (the coefficients were expressed via polarizabilities and effective numbers of electrons of the interacting atoms) and van der Waals radii for calculation of repulsion term parameters. Approximate character of such parameters is evident, but rather good agreement with experimental results was obtained in the both publications (Brant and Flory 1965; Scott and Scheraga 1966a). The existence of polyglycine and poly-L-alanine regular α -helices was rationalized as a consequence of favorable both H-bond and van der Waals interactions, while helices of other possible types (ω and 3_{10}) lie in relatively high-energy regions, and for isolated helices these structures have been excluded (Scott and Scheraga 1966a).

The abovementioned calculations of Scheraga et al. on protein simulations were followed by a series of publications related to both improvements of the method (both parameter adjustments and new term additions) and of simplifications of computations. Such changes in the methods enable them to obtain preliminary results on peptides of up to 20 amino acid residues already in 1967 (Gibson and Scheraga 1967a, b). The improvements of the method refer to enlarging the equilibrium radii of atoms, assigning partial charges to each atom (but neglecting the electrostatic contribution when the atoms are not close together), and introduction of orientation-dependent term for H-bond description instead of complicated formulae of the previous paper (Scott and Scheraga 1966a). The simplified account for solvent contribution to the energy has been included considering the effect of removing the nearest neighbor to peptide water molecules (Gibson and Scheraga 1967a). The simplification of the model refers to introducing "extended atom," i.e., CH, CH₂, and CH₃ groups were considered as the single atom to reduce the number of atom-atom interactions that had to be computed (this approach was used later in many force fields and such potential functions are commonly referred to as "united atom" potentials, we will discuss this simplification below). Scheraga's group prolongs

the improvement of the parameters and the method continuously till recent years. We will discuss these studies below, now we will mention some improvements performed in the first half of the 1970s. The most important improvements refer to description of H-bond potential by (10-12) expression (McGuire et al. 1972) (some other force fields adopted such dependence for description of interaction between hydrogen atom capable to participate in H-bond and atom-acceptor of H-bond), to calculation of partial charges via semiempirical CNDO/2 method (Yan et al. 1970), and to adjustment of repulsion part of van der Waals interactions to crystal data for extensive set of molecules including amino acid crystals (Momany et al. 1974, 1975). The results of the crystal calculations led to an internally self-consistent set of interatomic potential energies for interactions between all types of atoms found in polypeptides (ECEPP force field (Momany et al. 1975)). The bond lengths and bond angles were obtained from a survey of the crystal structures and were considered as fixed for each of amino acid residue (Scheraga's force fields differ in this feature from most other widely used force fields). The preliminary set of potential functions they used to refine X-ray structures of some proteins, e.g., lysozyme (Warne and Scheraga 1974), α -lactalbumin (Rasse et al. 1974), and rubredoxin (Warne et al. 1974).

We will not consider details of these papers; instead we will mention the first publication on MM refinement of X-ray protein structure of Levitt and Lifson (1969). The refinements of protein structures (myoglobin and lysozyme) have been performed via two steps by minimization of the energy function containing all the MM terms and an additional term describing deviation of the atom coordinates from their values obtained from X-ray diffraction studies (penalty function). Three energy terms (corresponding bond stretching, bond angle bending, and torsion potentials) supplemented by penalty functions were used at the first step. At the second step nonbonded and hydrogen-bonded interactions were included, and penalty functions were omitted. The authors mention that energy function parameters are preliminary ones, intended for protein simulation in aqueous solution, differed significantly from those of Scheraga and subject to further improvement. The refined structures were close to X-ray diffraction ones and do not contain short atom-atom contacts or unusual bond lengths and valence angles. Starting from rough model coordinates, the method minimizes the assumed total potential energy of the protein molecule to give a refined conformation (Levitt and Lifson 1969).

Molecular Mechanics on the First Steps of the Biophysics of Nucleic Acids

The problems related to MM approach to nucleic acid structure and functions differ in some aspects from those for proteins. The main computational tasks for the 1960s in the area of nucleic acids were to rationalize the structure of native and modified DNA duplexes, t-RNAs, synthetic polynucleotides, i.e., to understand the contribution of the subunits (the bases, sugars, ionized phosphate groups, counter-

ions, and surrounding water) to three-dimensional structure, and to evaluate the contributions of interactions of different physical nature to structure and functions of nucleic acids. The extensive applications of MM approach to nucleic acids were started a few years after those for proteins. The same is true for the adjustment of MM parameters for calculations of nucleic acids. Such a situation can be explained by “special role” of DNA in cells as a heredity material as well as by expectations of “specific” forces between the bases as conjugated molecules with delocalized electron system. Many researchers considered the electron exchange via H-bonds in base pairs and exchange interactions between stacked bases as the crucial contributions to nucleic acid functioning. It took some years before pioneers of quantum mechanics approach to biochemical problems would write: “The possible contributions of resonance energy stabilization through electronic delocalization across the hydrogen bond for horizontal interactions and of overlapping of their π -electronic cloud or of charge-transfer complexation for the vertical interactions appear to be of much smaller order the magnitude than the stabilization increments due to van der Waals-London forces” (Pullman and Pullman 1968).

The author of this survey was the first researcher since the application of MM formulae in 1966 to consider the interactions of nucleic acid subunits. We will now briefly follow the route of investigations of nucleic acid interactions to “classic” MM approach. The first quantitative estimation of interactions of bases in DNA duplex via formulae of intermolecular interactions was performed by De Voë and Tinoco (1962). They used so-called molecule dipole approximation, i.e., permanent and induced electric point dipoles were placed into the center of each base. The energy of interaction between two paired or stacked bases was suggested to be a sum of the interactions of permanent dipoles (Coulomb electrostatic interactions), permanent dipole-induced dipole (polarization interactions), and of induced dipoles (London dispersion interaction). Such an approximation does not enable them to obtain even qualitatively correct results (e.g., the energy of interaction between the bases in Watson-Crick A:T pair was positive), but the computational problem of evaluation of interaction energy changes on the formation of the specific nucleic acid conformation has been defined in this chapter. That is why it is cited in hundreds of publications during nearly a half of century, many of them (including those of the author of this survey) were inspired by the paper of De Voë and Tinoco (1962). Shortly it became clear that such an approximation is invalid for this system not depending on dipole location as well as after replacement of point dipoles by “real” ones, i.e., by dipoles of definite size. Bradley et al. (1964) suggested that the main contribution into base-base interaction energy is of Coulomb electrostatic nature, which can be calculated in “monopole approximation.” The point charges were calculated via semiempirical methods of quantum chemistry and placed on each atom of the bases. This paper was followed by the papers of Nash and Bradley related to calculations of base-base interactions. One of their papers should be mentioned in relation to the general progress and development of MM approach. Starting from various mutual in-plane positions for all the combinations of RNA bases, the calculations and search for minima of electrostatic energy was performed (Nash and Bradley 1966). Van der Waals energy was taken into account

in “hard sphere” approximation, i.e., any non-hydrogen atom pair should not be closer than sum of van der Waals radii. As a result of the calculations, 27 energy minima were obtained which have two or more short N-H...O or N-H...N contacts (with N...N or N...O distances about 3 Å) corresponding to nearly linear or bifurcated hydrogen bonds. The results enable Nash and Bradley to rationalize the available up-to-date experimental data on base crystals data as well as base-base complex formation in solutions and polynucleotides. It appears that the observed base pairings with two H-bonds in crystals always correspond to one of the computed geometries of lowest energy (Nash and Bradley 1966). It is important to emphasize that Bradley et al. (1964) “monopole” approach to calculations of the base interaction was the first attempt to consider the electrostatic energy in the framework of the modern MM scheme. The abovementioned calculations of Scheraga’s group with inclusion of point charges of peptide group were started 2 years after the Bradley’s publication.

The subsequent studies on the quantitative evaluation of the interactions and three-dimensional structure of nucleic acid fragments were continued via two routes. One of the approaches suggested construction of MM force field parametrized to nucleic acids interactions; another approach implied calculations using more rigorous physics concepts and molecular characteristics in a hope to understand physical nature of interactions. The first works on this way were performed by Pullman’s group and referred to interactions of the bases in fixed positions (e.g., Pullman et al. 1966). Considering the energy of interactions as a sum of electrostatic, polarization, and dispersion terms, two approximations were applied, “dipole” and “monopole”. The monopole approximation was applied to electrostatic energy only (following Bradley et al. 1964), while two other terms were really calculated in molecular dipole approximation. This approach was later augmented by inclusion of the short range repulsion term and representation of the molecule as a set of “many-centered multipole distributions obtained from ab initio SCF calculations (charges, dipoles, and quadrupoles located on the atoms and the middles of segments joining pairs of atoms)” (Langlet et al. 1981). We will not consider subsequent progress on this way of nucleic acid computations as new schemes (e.g., Gresh et al. 1986 and references herein) do not correspond to the MM approach. Nevertheless, it is worth to mention that the sophisticated schemes of base-base computations had no advantage for rationalization and prediction of experimental results in comparison to “standard” MM scheme reviewed in the next paragraph.

In the middle of the 1960s, the author of this survey proposed the first atom-atom scheme and numerical parameters for the calculation of interactions of nucleic acid bases (the terms “force field” and “molecular mechanics” were not widely used in those years). This scheme had three source points, namely, idea of De Voe and Tinoco (1962) on quantitative estimations of interactions of nucleic acid subunits, “monopole” approximation of Bradley et al. (1964) for electrostatic interactions of the bases, and Kitaygorodsky (1961) atom-atom approach to calculations of van der Waals interactions in molecular crystals. At that time, there was no “MM-like” calculation scheme suitable for quantitative considerations of nucleic acid

interactions and structure. We refer here only to later papers summarizing some preliminary adjustments and applications of the approach (Poltev and Sukhorukov 1970; Poltev and Shulyupina 1986). When calculating the interactions of nucleic acid bases or other conjugated heterocyclic molecules the only term $\sum E_{nb}$ of Eq. 1 was considered, corresponding to all the pair-wise interactions of atoms not pertaining to the same molecule. The bond lengths and valence angles were fixed and corresponded to averaged experimental geometry for each of the bases (like those for amino-acid residues and peptide backbone in Scheraga's force field (Momany et al. 1975)). The atom-atom terms contained electrostatic, London attraction, and short-range repulsion contributions. This scheme was the first one assigning different parameters of van der Waals terms to the atoms of the same chemical element but different electron shell configuration (e.g., aromatic and aliphatic carbons, pyridine and pyrrol nitrogens, keto and enol oxygens). These parameters were preliminarily evaluated from approximate London formulae (for A_{ij} coefficient of Eq. 5) and enlarged van der Waals radii of atoms (for B_{ij} coefficient of Eq. 5 or corresponded values for the firstly used 6-exp potentials). The final values of the parameters were selected after step-by-step calculations and comparison with experimental data for crystals containing the atoms of some types only (starting from graphite, then aromatic hydrocarbons, pyrazine, benzoquinone, purine, and finally the bases). The effective charges of atoms were calculated via simple semiempirical quantum chemistry method (of Huckel and Del Re for π -electrons and σ -electrons, respectively) with the parameters adjusted to reproduce experimental values of dipole moments for related molecules. The primary scheme (Poltev and Sukhorukov 1970) contained polarization term as well, but it was shortly eliminated due to its small value for many cases and to avoid difficulties related to nonadditivity of atom-atom and molecule-molecule interactions. The parameter set was than amplified by consistent parameters for sugar-phosphate backbone (Zhurkin et al. 1980) and for water-DNA interactions (Poltev et al. 1984). This simple MM scheme specially adjusted to nucleic acid interactions enables us to rationalize extended set of experimental data and to forecast some nucleic acid properties before experimental evidence. Variability of DNA helix parameters (Khutorsky and Poltev 1976), dependence of mutual base-pair positions in energy minima on nucleotide sequence (Polozov et al. 1973), pathways for all the base-substitution errors (Poltev and Bruskov 1978), and DNA duplexes with mispairs (Chuprina and Poltev 1983) were predicted by such calculations before experimental data became available. The parameters of this simple scheme were refined later several times (like any other force field scheme discussed above and below) via adjustment to new experimental data and quantum mechanics considerations. We will mention here only one more series of early MM calculations related to nucleic acids performed by Indian scientists Renugopalakrishnan, Lakshminarayanan, and Sasisekharan (coauthor of Ramachandran et al. (1963) paper on Φ - Ψ plot for proteins). They suggested the first complete set of parameters for calculations of the conformations of nucleic acid fragments (Renugopalakrishnan et al. 1971, and references herein). The atom charges crucial for the base interactions have been calculated via the same procedure as in our earlier papers (e.g., Poltev and Sukhorukov 1970).

The Problems and Doubts of Further Development of the MM Approach

As follows from the previous section, the usefulness of MM approach was already demonstrated in the first decade of its extensive applications and modifications. At the same time, the problems of justification and of pathways of future development of the approach had arisen. From the very beginning, it was clear that the method cannot provide us with exact description of the structures and processes due to its semiempirical nature. The interesting problem of the method is the relative extent of empiricism and physical meaning of the scheme and its parameters.

Approaches to Improvement of MM Formulae and Parameters

Several ways of improvement of MM scheme were proposed during the 1960s and the 1970s. Two hypothetical approaches to refinement of MM representation and force fields can be considered. In the framework of the first approach, it is suggested to describe the physical nature of the system in the most exact way, e.g., to employ more exact equations for description of the interaction energy (e.g., not restricting expressions to first term in the expansion for bond stretching and angle bending, use of better grounded expressions for van der Waals interactions, etc.) and/or more detailed representation of the system considering not only point atoms but other centers of interactions, not restricting the scheme with scalar values but introducing vectors (such as dipole moments) or tensors (such as anisotropic polarizability). In the framework of the second approach, it would be possible to consider energy expressions as entirely empirical ones and to adjust the parameters to experimental data (or, later, to reliable quantum mechanics results) until the closest coincidence (e.g., via the least square method). Unsuccessful attempts of consistent and strict use of only one of the approaches had been undertaken during the first two decades of the MM computer simulations.

The second approach mentioned above is practically impossible to realize; usually, there is no sufficient quantitative experimental data not only to derive the form of dependences of the energy on inter-atomic distances but to obtain the coefficients of already accepted potential form (e.g., of Lennard-Jones). Even in a case we have sufficient number of experimental values, the equations for parameter calculations have no single definite solution. It was demonstrated in the paper of Momany et al. (1968). The authors of the abovementioned paper tried to adjust the coefficient of (6-12) potentials for C...C, C...H, and H...H interactions for benzene molecule to the energy and structure data of benzene crystal. Using three sets of experimental data for different temperatures, they obtained entirely repulsive contribution from C-C interactions. It means that qualitative estimations (relative values of attraction terms for various atoms, of equilibrium radii, of atom charges in the same or similar molecules, etc.) should be taken into account in adjusting the parameters of the force field. Most of the modern force fields originate

from previous sets of expressions and parameters derived from simple qualitative considerations (e.g., simplified formulae of dispersion term, van der Waals radii of atoms) and take into account these considerations automatically (e.g., using previous approximate parameters' values as starting points).

As for the first approach, the more complex expressions of the energy were suggested since the end of 1960s, and such expressions are used in some modern force fields, i.e., next terms in the energy expansions for stretching, bending, and van der Waals contributions. The inclusion of additional terms requires new parameters. Usually the number of reliable quantitative values is not sufficient, and such an approach becomes useful in case the area of applications of the force field is restricted to certain types of similar molecules (e.g., hydrocarbons). In any case, the expressions for each term remain the approximate ones, and more coefficients should be adjusted to target data. The more complex representation of the molecular system as compared to a set of point atoms was proposed by several researchers during the 1960s and the 1970s (including the author of this survey). Such a representation looks the most natural for atoms with lone pairs of electrons (inclusion of additional points for electrostatic interactions of pyridine nitrogen or keto oxygen located at the centers of lone-pair orbitals) or for atoms with π -orbitals (additional points for aromatic heterocycles located above and below the ring planes to account for quadrupole moments of the molecules). Additional lone pair centers for electrostatic interactions are included in some modern force fields for better reproduction of electric field around the molecules and the directionality of hydrogen bonds.

Although the introduction of such centers looks physically based, an inconsistency of such an approach can be detected rather easily. When assigning the negative charged interaction centers to some electron orbitals (and hence positively charged centers to nuclei locations), it looks natural to assign additional centers to chemical bonds (sites of the highest electron density between the nuclei). As a result, we have several times more centers (more computer resources are needed) and several times more adjustable parameters. Such an approach was consistently implemented into force field suggested by Scheraga and coauthors. The first version was titled "empirical potential using electrons and nuclei" (EPEN ; Shipman et al. 1975) and an improved version EPEN/2 was released in 3 years (Snir et al. 1978). In this model, the molecular interactions are modeled through distributed interaction sites, which account for the nuclei and electronic clouds. The positive charges are located at the atomic nuclei, and the negative charge centers are located off the nuclei. The bonding electrons are located along the bonds, while the lone-pair and π - electrons are located off the bonds. The energy of interaction is approximated by the sum of the coulomb interactions between all point charge centers, an exponential repulsion to represent electron–electron overlap repulsion, and an R^{-6} (R = distance) attraction to simulate attractive energies between the nuclei. The distances between the charge locations are the method parameters; they are fixed in the molecule fragments. The parametrization was optimized by least square fits to spectroscopic, crystallographic, and thermodynamic data. This approach (as well as other approaches considering many in addition to atom centers of interactions)

was not widely used in MM calculations. The situation can be accounted for by restrictions imposed by fixed geometry of the molecule fragments as well as by need for the parameter adjustment for molecules other than considered by the approach authors.

A separate problem is the choice of the force field parameters for most exact reproduction of experimental and/or quantum mechanics data. All the experimental values are determined with a definite accuracy; all the quantum mechanics data refer to gas phase and do not include implicitly third body interactions. The use of quantum mechanics data only (not depending on level of theory) results in errors in reproduction of condensed phase properties. The adjustment of parameters to characteristics of condensed phase (most experimental data refer to crystals or solutions) leads to wrong values of characteristics for separate molecules or their simple complexes in gas phase. Hence, various experimental and quantum mechanics data should be considered with different weights in the process of the adjustment of force fields, the weights being dependent on areas of intended applications of the force field.

Recently, a method and software package named ForceBalance for systematic force field optimization has been published (Wang et al. 2013a). The method enables researchers to parametrize a wide variety of functional forms using flexible combinations of target data. These data may come from experimental measurements or from quantum mechanics calculations. Various properties are included in a single objective function, which is then integrated over the entire configuration space. The method was used by the authors (Wang et al. 2013a) for developing a new polarizable water model with 27 parameters adjustable to quantum mechanics data for single water molecule and H-bonded water clusters. The water model contains five fluctuating charge sites including three virtual sites. The positions of virtual sites have been fully optimized; they do not coincide with lone pair centers. An interesting suggestion was made for description of van der Waals interactions. The authors propose specially modified (6-exp) potential. As it was mentioned in the first section, “classical” (6-exp) Buckingham function widely used on the first steps of MM-like calculations has unphysical dependence on interatomic distance at small distances. The authors of the model (Wang et al. 2013a) have chosen empirical van der Waals potential $E_{vdw}(r)$ with three parameters.

$$E_{vdw}(r) = 2\varepsilon \left(\rho^6 / (\rho^6 + r^6) \right) / (1 - 3 / (\gamma + 3)) [3 \exp(\gamma(1 - r/\rho)) / (\gamma + 3) - 1] \quad (7)$$

This function depends on equilibrium interatomic distance, ρ ; minimal energy values, ε ; and a dimensionless constant, γ , describing the steepness of the repulsion. Like Buckingham function, it tends to an attractive r^{-6} potential at large interatomic distances, tends to a repulsive exponential term at small distances, and has well-defined unique minimum. The authors note that their water model is quite accurate for some, but not all, properties of liquid water. We should note that this disadvantage is common for all the force fields derived from quantum mechanics results

only, without adjustment to experimental data. Improvements of water model using ForceBalance software and both quantum mechanics and experimental data will be considered below.

Various Schemes of Water Molecule in MM Calculations

In view of importance of water for life and for description of the biomolecular interactions as well as of “anomalous” physical properties of water in liquid and solid states, hundreds of models and their modifications for water molecule have been proposed. We will mention here only few of most frequently used and cited rigid models and some recent polarizable models. Simple atomic structure of the water molecule enables researchers to perform quantum mechanics calculations of rather high level unreachable for most organic molecules. The results of these computations for a single molecule and for H-bonded water clusters together with availability of quantitative experimental data for bulk water at various temperatures and pressures stimulate elaboration of new models and permanent improvement of existent ones. These models can be used for demonstration of possibilities and restrictions of the MM method in general, for comparison of various approaches to MM model construction.

Most water models describe the molecule by the number of interaction centers not coinciding exactly with atom centers. This diversion from simple MM scheme can be understood as a consequence of small size of the molecule and the inability of any simple charge distribution on atom centers to adequately describe the molecular dipole and quadrupole of the monomer at the same time. Inclusion of additional centers or/and displacement of interaction centers from atoms is necessary to reproduce charge distribution of this simple molecule. (The author of this survey noticed this fact earlier for heterocyclic compounds, but fortunately in this case the contributions of permanent quadrupole interactions can be implicitly accounted for by other terms of interaction energy.)

The first water molecule model (frequently referred to as BF model) has been proposed by Bernal and Fowler (1933). This model may be considered as the earliest molecular model suitable for MM calculations, and it was used for Monte Carlo calculations of liquid water performed in nearly a half of century for comparison with some later models (Jorgensen et al. 1983). The model contains four interaction sites: three charged sites, namely, two positive centers located at the hydrogen atoms and one negative center (so-called M-site) displaced by 0.15 Å from oxygen atom in the direction of H-O-H bisector, and one center for van der Waals interactions located on oxygen atom. It is interesting to note that Bernal and Fowler (1933) performed approximate evaluations of intermolecular interaction energy by the (1-6-12) potential functions. Jorgensen et al. (1983) used nearly the same position of centers and slightly modified charges in TIP4P four-site water model (one of their popular Transferable Intermolecular Potentials). Matsuoka et al. (1976) derived more complex four-site model from potential surface of water dimer constructed via ab initio quantum mechanics calculations with configuration interactions.

The analytical expression for interaction potential function between two water molecules consists of pair-wise Coulomb interactions of three charges located similar to BF or TIP4P models and exponential interactions of all the three atoms of one molecule with the atoms of the other one. This model (the so-called MCY model) was widely used during the first decade of liquid water simulations.

We will not consider many models including earliest ones which consist of three, five, or more centers but only briefly mention some of the rigid models cited till the present time and used in modern force fields. Stillinger and Rahman (1974) suggested the five-site model with two positive centers (hydrogen atoms), two negative centers (located at oxygen lone pairs), and one center (oxygen atom) for Lennard-Jones interactions. The model has been used successfully to study a wide variety of properties of liquid water. Mahoney and Jorgensen (2001), after extensive Monte Carlo simulations and parameters adjustment, suggested five-site model TIP5P which enables them to describe the density anomaly of liquid water better than previously existing models. Nada and Van der Eerden (2003) have designed six-site model of water molecule. Three of these sites are the three atoms of the molecule (they interact through a Lennard-Jones potential), and three other sites are point charges (O atom is electrically neutral; the two H atoms carry positive charges). The model can be considered as a combination of TIP4P and TIP5P models of Jorgensen. The Monte Carlo simulations of ice and water show that the six-site model reproduces well the real structural and thermodynamic properties of ice and water near the melting point.

Two three-site rigid water models, namely, TIP3P of Jorgensen et al. (1983) and SPC of Berendsen et al. (1981) should be specially mentioned, as these models (or their modifications) are used in modern force fields for biomolecular simulations. These models have (in accord with general MM principles for class I models) three interaction sites centered on the nuclei. Each site has a partial charge for computing the Coulomb energy and only one site (oxygen atom) for Lennard-Jones interactions. The TIP3P model demonstrates better reproduction of experimental thermodynamic and structural data for bulk water (Jorgensen et al. 1983).

Further progress in the study of water properties and its interactions with other molecules is related to new models introducing flexibility to the molecule (incorporating bond stretching and angle bending terms into interaction energy) and with the inclusion of electronic polarization. These effects are important for the quantitative description of water properties in liquid and solid states as well as of phase transitions. We mention here a few of recent water models beyond class I MM methods.

The model elaborated by Ren and Ponder (2003, 2004), AMOEBA (Atomic Multipole Optimized Energetics for Biomolecular Applications), includes full intramolecular flexibility, permanent atomic monopole, dipole, and quadrupole moments placed on each atomic center and buffered 7-14 potential for pairwise atom-atom van der Waals interactions. Polarization effects are explicitly treated in the AMOEBA force field via mutual induction of dipoles at atomic centers. The functional forms for bond stretching and angle bending were taken from the Allinger's MM3 force field and include anharmonicity through the use of higher-order deviations from ideal bond lengths and angles. An additional intramolecular

term, the Urey-Bradley term, consists of a simple harmonic function. The force field does not contain off-atom sites. It was parametrized by hand to fit results from *ab initio* calculations on gas phase clusters up to hexamers (Ren and Ponder 2003) and temperature and pressure dependent bulk phase properties (Ren and Ponder 2004). The model enables authors to reproduce well both the high-level *ab initio* results for various gas phase clusters and bulk water thermodynamics data, including the second virial coefficients over the whole temperature range. The authors mention that AMOEBA water model is approximately eight times slower than the TIP3P one (for MD simulations of a periodic system of 216 water molecules); nevertheless, they consider it as “the first step in the derivation of a next-generation polarizable force field for biomolecular simulation.” Later, AMOEBA force field was extended to organic molecules (Ren et al. 2011) and to proteins (Shi et al. 2013).

The iAMOEBA, i.e., “inexpensive AMOEBA” (Wang et al. 2013b) was developed using the ForceBalance automated procedure (Wang et al. 2013a). A direct approximation to describe electronic polarizability reduces the computational cost relative to a fully polarizable AMOEBA model, and the application of systematic optimization ForceBalance method results in raising the accuracy of iAMOEBA.

Recently (Laury et al. 2015), a set of improved parameters for the AMOEBA water model was developed using the automated procedure, ForceBalance (Wang et al. 2013a). An adjustment of 21 force field parameters has been performed to both high-level *ab initio* results for gas phase clusters, containing 2–20 water molecules, and six condensed phase properties at a range of temperatures (249–373 K) and pressures (1 atm to 8 kilobar). The enthalpy of vaporization and density dependence on temperature were given a greater weight (as compared to the other properties) in the optimization. The model represents a significant refinement of the original AMOEBA parametrization, namely, it provides a substantial improvement in the ability to predict experimental data, particularly for a number of liquid properties across a range of temperatures. The authors (creators of original AMOEBA models are among them) refer to new model as AMOEBA14 and to the original model (Ren and Ponder 2003, 2004) as AMOEBA03. The AMOEBA14 model has a similar accuracy level as the AMOEBA03 model for the interaction energy of the gas phase cluster ranging from dimers to 20-mers. They promise compatibility of AMOEBA14 water model with the previously determined AMOEBA force fields for organic molecules (Ren et al. 2011) and for proteins (Shi et al. 2013).

The authors of CHARMM force field and software package, widely used for biomolecular simulation (see description below), elaborate their own polarizable models of water based on classical Drude oscillator. Last model of their SWM (simple water model) series, six-site SWM6 one (Yu et al. 2013) was developed and optimized for liquid phase simulations under ambient conditions using CHARMM family of force fields. The atoms of the molecule kept its gas phase experimental geometry, and the positions of three other massless interaction sites correspond to the so-called M point located at a fixed distance from the oxygen atom along the bisector of the H–O–H angle, and two oxygen’s lone pairs (LP sites). The model is less computer time consuming as compared to more sophisticated AMOEBA; it has an improved balance between descriptions of microscopic and bulk phase

properties compared to most additive and polarizable water models developed previously. Besides, the local hydrogen bonding structures in the liquid phase are better reproduced in the model as indicated by the O-O distance distribution function. However, the model does not reproduce correctly the experimental data on liquid water density maximum. Authors note that the model is not developed for routine use but for studies when the local hydrogen bonding is important and high accuracy is needed.

We will not discuss many other recent sophisticated water models, including those developed using both experimental data and quantum mechanics computations (via both *ab initio* and DFT methods), such as Tröster's polarizable molecular mechanics (PMM) model (Tröster et al. 2014) and Pinilla's polarizable and dissociable force field (Pinilla et al. 2012). Selected old and recent MM water models, both simple and rather sophisticated, enable us to arrive at some conclusions on possibilities of molecular mechanics description of simple and complex molecular systems. Some of such conclusions are rather trivial, but additional evidence from the experience of hundreds of researchers on water molecule modeling can help to illustrate and reinforce them.

Till recent time, complete and reliable MM description of even this very simple biologically important molecule is a challenge for computationally oriented researchers. More complex expressions for separate energy terms grounded on physical theories, increase of the number of adjustable parameters, improved algorithms for parameter adjustment to target data, and the use of more reliable quantum mechanics results, all together help to better reproduce and rationalize experimental data, but some results are not accurately reproduced. Each model of the molecule, each force field, reproduces (and predicts) some characteristics better as compared to other force fields. The more sophisticated model is not necessarily the better one for specific MM applications and specific tasks of the study. This is true not only because of restrictions of computer resources but due to approximate nature of the MM description. This conclusion is a result of many trials and errors in water simulations. Class I molecular mechanics representations, i.e., simple three-site rigid water models (TIP3P and its modifications) are successfully used in many biomolecular computations in spite of the development of new force fields. This success can be accounted for not by the importance of additional terms but by their implicit compensation by parameter choice for simple model (Class I force fields). More sophisticated force fields (Class II and Class III), together with quantum mechanics, both *ab initio* and semiempirical, computations are useful for detailed description of mutual positions of molecules in water solutions of biomolecules.

Modern Molecular Mechanics Force Fields: Description, Recent Development, and Applications

The MM computations during the first two decades were mainly performed by the researchers using their own force fields and homemade computer programs. Most of the modern publications refer to standard, ready-to-use software with implemented

(or sometimes slightly modified) force field parameters. The number of program and parameter packages is rather great and new or newly modernized software and force fields appear every year. We will present below a short overview of the selected program sets and force fields most frequently used for study of biomolecular systems. Some of them have been already mentioned in previous sections. During last decade, some software packages such as NAMD (Phillips et al. 2005) and TINKER (Ponder 2015) have appeared which give an opportunity to perform computations using various molecular mechanics force fields. Some popular quantum mechanics software packages such as GAUSSIAN and SCHRODINGER supply users with the possibility to use MM force fields as well.

Allinger's Force Fields and Programs

We started description of the modern force field with short characterization of Allinger's parameters and programs as they were the first ones ready for use by any researcher. The first parameter set (MM1) was described in 1973 (Allinger and Sprague 1973), and the MM1 program was released through Quantum Chemistry Program Exchange in 1976. The force field and the program MM1 were followed by a series of improved and extended versions for various classes of organic molecules, MM2, MM3, MM4 (see Allinger et al. 1989, 1996; Lii et al. 2005 for references). These force fields have been parameterized based on experimental data, including the electron diffraction, vibrational spectra, heats of formation, and crystal structures. The calculation results were verified via comparison with high-level *ab initio* quantum chemistry computations, and the parameters were additionally adjusted. The last program versions include rather complex force field equations (including additional terms to harmonic potentials for stretching and bending and cross-terms; thus MM3 and MM4 force field are those of class II) and various calculation options (e.g., prediction of frequencies and intensities of vibrational spectra, calculating the entropy and heat of formation for a molecule at various temperatures). MM4's molecular dynamics option creates a description of the vibrational and conformational motion of molecules as a function of time. The program examines the atoms of a molecule and their environment and decides which atom type is appropriate. The program includes more energy terms and more complicated expressions for some terms than most other popular MM force fields. The force field contains greater number of the constants which are adjusted to extended set of experimental data. The stretching and bending energy terms include higher order contributions, several cross-terms are inserted; the electrostatic contribution includes interactions of bond dipoles.

The MM family of force fields designed by Allinger and coauthors is used till recent time for calculations on small and medium size organic molecules, e.g., steroidal-porphyrin aggregates (Lorecchio et al. 2014). New improvements to MM4 force field were suggested by Allinger (2011) recently. The classic of molecular mechanics, after some examples make the conclusion: "A good (well developed, Class 3) molecular mechanics force field will usually give molecular structures that

are approximately of chemical accuracy. When it fails to do so, there is a good chance that nature knows something that we don't" (Allinger 2011).

Merck Molecular Force Field (MMFF94)

This force field of class II was developed for a broad range of “medium size” molecules primary of importance for drug design. It differs from Allinger's and many other popular force fields in several aspects. The core portion of MMFF94 has primarily been derived from high-quality computational quantum chemistry data (up to MP4SDQ/TZP level of theory) for a wide variety of chemical systems of interest to organic and medical chemists (Halgren 1996, 1999b). Nearly all MMFF94 parameters have been determined in a mutually consistent fashion from the full set of available computational data. The force field reproduces well both computational and experimental data, including experimental bond lengths (0.014 Å rms), bond angles (1.2° rms), vibrational frequencies, conformational energies, and rotational barriers. The mathematical expressions of the force field differ from many other force fields. The expressions for stretching and bending energies (like Allinger's force fields) contain in addition to harmonic terms stretch-bend cross terms and out-of-plane terms. The van der Waals energy is expressed by buffered (7-14) potentials (Eq. 8, the designations are the same as in Eqs. 2, 3, 4, 5, 6, and 7)

$$E_{\text{vdw}} = \epsilon_{ij} \left(1.07 \rho_{ij} / (r_{ij} + 0.07 \rho_{ij})^7 \right) \left((1.12 \rho_{ij}^7 / (r_{ij}^7 + 0.12 \rho_{ij}^7)) - 2 \right) \quad (8)$$

The electrostatic term contains buffering constant as well. The specially formulated combination rules for ϵ_{ij} and ρ_{ij} are used. The MMFF94 version was developed for molecular dynamics simulations, and MMFF94s version (Halgren 1999a) was modified for energy minimization studies (*s* in the title means static). Both versions provide users with nearly the same results for majority of molecules and complexes. Additional parameters of force field and careful adjustment to high-level reference data result in good reproduction of extended set of experimental data on conformational energies of the molecules appeared after force field publication (Halgren 1999b). Nevertheless, Bordner et al. (2003) reveal that MMFF94 values of sublimation energy for a set of compounds are systematically 30–40 % lower than experimental data. It can be explained by “computational nature” of MMFF parameters and by the adjustment to quantum mechanics results without corrections for the experimental data, i.e., nonadditive contributions do not accounted for implicitly.

This force field is used in many recent studies of biologically important molecules, which employ combined experimental and computational approaches. Anti-inflammatory drugs (Pawar et al. 2014; Sawant et al. 2014) and anticholinergic and antiasthmatic drugs (Guo et al. 2014) are among the compounds considered recently with MMFF.

The Force Fields and Programs Designed by Scheraga and Coauthors

As we had already mentioned that the first MM parameter set for the simulation of proteins was described in 1975 by Scheraga and coauthors (Momany et al. 1975); their empirical conformational energy program for peptides (ECEPP) was released through *Quantum Chemistry Program Exchange* in the same year. This force field family was refined continuously till last years; we will briefly describe two versions of Scheraga's force field used for studies of protein three-dimensional structure and protein folding up to recent time, ECEPP/3 (Némethy et al. 1992) and ECEPP-05 (Arnautova et al. 2006). The ECEPP/3 force field in turn is the modified version of ECEPP/2; it contains updated parameters for proline and oxyproline residues. The ECEPP force fields utilize fixed bond lengths and bond angles obtained from a survey of the crystal structures for each of amino acid residue. The peptide energy is calculated and minimized as a function of torsion angles only, and it is a significant advantage for minimization approaches as it drastically reduces the variable space that must be considered. The energy function of the ECEPP/3 force field consists of torsion and nonbonded (Coulomb and van der Waals) terms (Eqs. 4 and 5). The partial charges of atoms are calculated by the molecular orbital CNDO/2 method, the parameters of (6-12) term are adjusted by comparison with crystal data. The (6-12) terms are substituted by (10-12) ones for the interactions of the hydrogen atoms capable to participate in H bonds with H-bond acceptor atoms. In spite of appearance of new improved and augmented versions of ECEPP force field, the ECEPP/3 is used until recently for protein simulations (e.g., McAllister and Floudas 2010).

The ECEPP-05 force field (Arnautova et al. 2006 and references herein) is adjusted to both new experimental data and quantum mechanics results. It has some distinctive features as compared to both ECEPP/3 version and many other popular force fields. The van der Waals term of the energy is modeled by using the "6-exp" potential function (potential of Buckingham)

$$E_{ij} = -A_{ij}r_{ij}^{-6} + B_{ij} \exp(-C_{ij}r_{ij}) \quad (9)$$

where r_{ij} is the distance between atoms i and j ; A_{ij} , B_{ij} , and C_{ij} are parameters of the potential. The combination rules are applied directly to these parameters.

$$A_{ij} = (A_{ii}A_{jj})^{1/2}; B_{ij} = (B_{ii}B_{jj})^{1/2}; C_{ij} = (C_{ii} + C_{jj})/2 \quad (10)$$

The parameters of (6-exp) potentials were derived by using so-called global-optimization-based method consisting of two steps. An initial set of parameters is derived from quantum mechanics interaction energies (at MP2/6-31G* level of ab initio theory) of dimers of selected molecules; at the second step, the initial set is refined to satisfy the following criteria: the parameters should reproduce the observed crystal structures and sublimation enthalpies of related compounds, and

the experimental crystal structure should correspond to the global minimum of the potential energy.

The atomic charges were fitted to reproduce the molecular ab initio electrostatic potential, calculated at HF/6-31G* level; the fitting was carried out using the restrained electrostatic potential (RESP) method taken from the AMBER program. The method was applied to obtain a single set of charges using several conformations of a given molecule (the multiple-conformation-derived charges). An additional point charge (with zero nonbonded parameters) is used to model the lone pair electrons of sp^2 nitrogen. The torsional parameters of the side chains of the 20 naturally occurring amino acids were computed by fitting to rotational energy profiles obtained from ab initio MP2/6-31G** calculations. The peptide backbone torsional parameters were obtained by fitting the energies to the Φ - Ψ energy maps of terminally blocked glycine and alanine amino acids constructed at MP2/6-31G**//HF/6-31G** level of theory. The performance of the force field was evaluated by simulating crystal structures of small peptides.

As the important problems of the protein structure and functions as well as of the natural science as a whole refer to prediction of three-dimensional structure and pathways of folding of proteins, Scheraga and coauthors have modified and augmented the force fields mentioned above by some theoretical approaches and concepts. We will notice a few of them related directly to MM calculations. The all-atom ECEPP05 force field was coupled with an implicit solvent-accessible surface area model (ECEPP05/SA, Arnautova and Scheraga 2008). Recognizing the impossibility of searching enormous conformational space of real protein with an all-atom potential functions, a hierarchical procedure was developed. Its main features are the initial use of a united residue (UNRES) potential functions and an efficient procedure, conformational space annealing (CSA) (Scheraga et al. 2002, and references herein). The protein is first optimized with the low-resolution “coarse-grained” UNRES model (Liwo et al. 1999). In UNRES force field, the backbone is represented as a virtual-bond chain of C_α atoms, and the side chains are depicted as ellipsoids. The interaction sites are the united-atom side chains and the centers of the peptide groups between C_α atoms. The lowest-energy UNRES conformation, as well as a set of distinct low-energy conformations, is then converted to all-atom models. Finally, the all-atom models are refined with the EDMC method, with inclusion of continuum hydration model.

The search for low-energy “native-like” structures includes several minimization and Monte Carlo procedures. The simplest of the MC methods is Monte Carlo with minimization, i.e., a Metropolis Monte Carlo (MCM) algorithm in which every trial state is first energy-minimized before the Boltzmann acceptance criterion is applied. The electrostatically driven Monte Carlo (EDMC) method employs a move set in which individual peptide groups are selected at random and rotated “in place” (i.e., the conformational change is localized to the peptide group as much as possible) so as to optimize the alignment of its dipole moment with the local electric field. The methodology assumes that a protein molecule is driven toward the native structure by the combined action of electrostatic interactions and stochastic conformational

changes associated with thermal movements (Ripoll and Scheraga 1988). The use of the force fields and methods enables Scheraga and coauthors not only to rationalize but to predict some interesting features of protein three-dimensional structure and pathways of folding. “With increasing refinement of the computational procedures, the computed results are coming closer to experimental observations, providing an understanding as to how physics directs the folding process” (Scheraga 2008).

Several recent successful applications of UNRES force to simulations of biological functions of proteins can be mentioned. The mechanism of growth of amyloid β -peptide fibrils (the main component of plaques formed in the human brain during Alzheimer disease) was suggested (Rojas et al. 2011) and supported by simulated 2D ultraviolet spectroscopy (Lam et al. 2013). The mechanism of HSP70 heat-shock protein action was proposed which was based on ATP-induced scissor-like motion of the two sub-domains of the nucleotide-binding domain (NBD) (Golas et al. 2012). The computationally obtained structure of this chaperon was confirmed then by experimental study. Some additional examples of UNRES applications are mentioned in very recently published review of Scheraga (2015). Similar approach was developed by Scheraga and coauthors for nucleic acids interactions. The first paper on coarse-grained models of nucleic acids represents NARES approach (Maciejczyk et al. 2010) to nucleic acid bases. Each base is simulated by a combination of three to five interaction centers of 6-12 van der Waals and permanent dipole interactions. Then, UNRES-2P model was developed (He et al. 2013) representing a polynucleotide chain by a sequence of virtual sugar atoms, located at the geometric center of the sugar ring, linked by virtual bonds with attached united base and united phosphate groups. This oversimplified model reproduces the double helix formation, some its thermodynamics properties (He et al. 2013), and was very recently applied to DNA-protein interactions (Yin et al. 2015).

Force Fields and Programs Developed by Kollman and Coauthors. AMBER

The computer software and force fields referred to as “AMBER” (Assisted Model Building with Energy Refinement) are the most popular of those designed for wide class of biomolecules including proteins and nucleic acids. The first description of the computer program AMBER appeared in 1981 (Weiner and Kollman 1981), and the first detailed specification of force field was published 3 years later (Weiner et al. 1984). The force field contained usual “minimalist” set of energy terms (Eqs. 1, 2, 3, 4, 5, and 6 of this survey), the parameters being adjusted to both experimental and quantum mechanics data. The parameters of their previous publications and those of other authors were used as starting points for the adjustment. Atomic charges are obtained by fitting a partial charge model to electrostatic potentials calculated by ab initio quantum mechanics calculations (with minimal STO 3G basis set). The (10-12) potential was included for description of H-bonding (this term was deleted in new versions of AMBER force fields and software). The force field accounted

for implicit solvent by using distance dependent dielectric constant ($\epsilon = R_{ij}$), united atom approach (i.e., several atoms, e.g., four atoms of a methyl group are considered as one particle) enabled one to perform calculations of sufficiently extended systems (such as fragments of DNA double helix of a few nucleotide pairs or peptides of dozens of amino acid residues). In spite of several deficiencies of the first AMBER force field (e.g., too short distances in stacking complexes due to short equilibrium radii for van der Waals interactions, defects in atomic charges due to very approximate quantum mechanics data), hundreds of papers were published with this force field and the program. In 2005, even after release of several new and improved AMBER force field versions, the paper (Weiner et al. 1984) was the 10th most-cited one in the history of the *Journal of the American Chemical Society* (Case et al. 2005).

Since those publications new versions of both force fields and program sets were released. Cornell et al. (1995) published “second generation” AMBER force field for simulating the structures, conformational energies, and interaction energies of proteins, nucleic acids, and many related organic molecules in condensed phases. The charges were determined using the calculations with 6-31G* basis set and restrained electrostatic potential (RESP) fitting and had been shown to reproduce interaction energies, free energies of solvation, and conformational energies of small molecules to a good degree of accuracy. The new van der Waals parameters were derived to account for liquid properties. Retaining most of torsion parameters of “the previous generation” of force field, the improvement of potentials for peptide backbone was performed for Φ and Ψ dihedrals as well as for the nucleoside χ dihedrals by adding extra Fourier terms. This “second generation” force field is usually referred to as *ff94* (parameter set *parm94*).

There are a number of AMBER parameter sets, with names beginning with “ff” and followed by a two-digit year number, such as *ff99*. We will briefly mention some new features of these force fields. Some versions differ from *ff94* by minor (but important for some problems) modifications. The *ff96* force field differs from *ff94* only in the peptide Φ - Ψ torsional potential, and *parm96* was developed because it was clear that *parm94* overstabilized α helices relative to β sheets. Thus, *parm94* seems adequate for comparing structures with similar secondary structures, but *parm96* seems to more accurately represent the relative stability of α and β secondary structures. The improved version *ff99* (Wang et al. 2000) was developed using the restrained electrostatic potential (RESP) approach to derive the partial charges. The adjustment of atom point charges and torsion constants for the set of 34 molecules to the highest level ab initio model (GVB/LMP2) and experimental data enables authors to reproduce data for 55-molecule set with the better accuracy than MM3, MMFF, and early versions of CHARMM force fields. It was demonstrated that “a well-parameterized harmonic force field with a reliable charge method can describe the structure and intramolecular energies for organic systems very well” (Wang et al. 2000). The authors tried to use the adjusted parameters (*parm99*) in nonadditive models (which include polarization), and obtained “reasonable results using the same parameters derived for the additive model, although additional torsional parameterization is required to achieve the same high level accuracy

as that found using the additive model” (Wang et al. 2000). This force field is widely used during more than 10 years. The other AMBER force field, “a third-generation point-charge all-atom force field for proteins” (referred to as *ff03*) was developed by Duan et al. (2003). The main difference from previous force fields is that all quantum mechanics calculations were done in the condensed phase with continuum solvent models and an effective dielectric constant of $\epsilon = 4$. Initial tests on peptides demonstrated a high degree of similarity between the calculated and the statistically measured Ramachandran maps for Glycine and Alanine peptide fragments. One more AMBER force field referred to as general Amber force field (GAFF) should be mentioned (Wang et al. 2004). GAFF have been designed as an extension of the AMBER force fields to be compatible with existing versions for proteins and nucleic acids and has parameters for most organic molecules of pharmaceutical importance that are composed of H, C, N, O, S, P, and halogens. The parameterization is based on more than 3000 MP2/6-31G* optimizations and 1260 MP4/6-311G(d,p) single-point calculations. Unlike most conventional force fields, parameters for all combinations of atom types are not contained in an exhaustive table but are determined algorithmically for each input molecule based on the bonding topology (which determines the atom types) and the geometry. A completely automated, table-driven procedure (called *antechamber*, a part of the late versions of AMBER program set) was developed to assign atom types, charges, and force field parameters to almost any organic molecule (Wang et al. 2006). New versions of AMBER software work with several force fields. The newest AMBER 11 version (Case et al. 2010) supports CHARMM fixed-charge force fields as well.

The initial AMBER force fields and program packages (due to the state of the theory and computer facilities) were generally aimed at the search for energy minima of separate molecular systems in vacuum. The current AMBER versions are strongly aimed at simulations (molecular dynamics, evaluations of free energy changes) of biomolecules in water solutions. Both explicit solvent models (using TIP3P, TIP4P, TIP5P, and SPC water models as well as models of some organic solvents) and implicit ones are supported. By default, explicit solvation model involves electrostatic interactions handled by a particle-mesh Ewald (PME) procedure, and long-range Lennard-Jones attractions are treated by a continuum model. The PME is a modified form of Ewald summation, a method to efficiently calculate the infinite range Coulomb interaction under periodic boundary conditions, and PME is a modification to accelerate the Ewald reciprocal sum to near linear scaling using the three-dimensional fast Fourier transform (3DFFT).

An accurate description of the aqueous environment is essential for atom-level biomolecular simulations but may become very expensive computationally. An implicit model replaces the discrete water molecules by an infinite continuum medium with some of the dielectric and “hydrophobic” properties of water. The continuum implicit solvent models have several advantages over the explicit water representation, especially in molecular dynamics simulations (e.g., they are often less expensive and generally scale better on parallel machines; they correspond to instantaneous solvent dielectric response; the continuum model corresponds to solvation in an infinite volume of solvent, there are no artifacts of periodic boundary conditions; estimating free energies of solvated structures is much more straight-

forward than with explicit water models). Despite the fact that the methodology represents an approximation at a fundamental level, it has in many cases been successful in calculating various macromolecular properties (Case et al. 2005).

The total energy of a solvated molecule can be written as a sum of molecule's energy in vacuum (gas phase) and the free energy of transferring the molecule from vacuum into solvent, that is, solvation free energy typically assumed as can be decomposed into the electrostatic and nonelectrostatic parts. The above decomposition is the basis of the widely used PB/SA (Poisson-Boltzman solvent accessible surface area) scheme. AMBER follows this way to compute the nonelectrostatic part of the energy. The most time-consuming part of solvation energy estimations is the computation of the electrostatic contribution. AMBER employs one of the most popular biomolecular applications the finite-difference method (FDPB) and analytic generalized Born (GB) method. In the GB model, a molecule in solution is represented as a set of spherical cavities of dielectric constant of 1 with charges at their centers embedded in a polarizable dielectric continuum solvent of a higher dielectric constant.

The AMBER software and force fields are widely used for the simulation of various systems of biological and material science importance. Some researchers develop proper additions and corrections to “standard” parameters for better description of specific systems. We refer to only few examples. Perez et al. (2007) reparameterized the *parm99* force field for nucleic acid simulations improving the representation of the α/γ conformational space. Zgarbova et al. (2011) refined description of glycosidic torsion angles of nucleic acids. Other work of Zgarbova et al. (2013) refers to ϵ and ζ DNA torsions. Very recently, Ogata and Nakamura (2015) suggested the improvement of parameters of the force field for phospholipids. Song et al. (2008) adjusted the AMBER force field to reproduce conformational properties of an oxidative DNA lesion, 2,6-diamino-4-hydroxy-5-formamidopyrimidine. New improved parameters for modified peptide chain were suggested by Grubišić et al. (2013) recently. Some of such improvements have been included into newer versions of AMBER package. AMBER force field can be used via other MM program packages, such as CHARMM, GROMACS, NAMD, and with widely used QM package GAUSSIAN.

The latest version of AMBER program package is AMBER14 distributed with AmberTools 15 (Case et al. 2015). The new versions of AMBER software support the option of allowing part of the system to be described quantum mechanically in an approach known as a hybrid (or coupled potential) QM/MM simulation. Both semiempirical and the density functional theory-based tight-binding (DFTB) Hamiltonian can be used. The system may contain both two nonbonded parts and covalently bonded QM and MM subsystems (Case et al. 2010).

Other Popular Force Fields and MM Software. CHARMM, OPLS, GROMOS

The force fields and program packages listed above are used for nearly the same type of simulations as those considered above. The main differences refer to the

organization of the programs and to relative role of experimental data, and quantum mechanics results in force field parameter adjustment. The first versions of force fields and programs as well as the whole research programs were initiated by such pioneers of MM applications as Karplus (CHARMM, Chemistry at HARvard Molecular Mechanics), Jorgensen (OPLS, Optimized Potentials for Liquid Simulations), Gunsteren and Berendsen (GROMOS, GRONingen MOlecular Simulation package), and GROMACS (GRONingen MACHine for Chemical Simulation). The progress and improvement of the force fields and programs are connected to comparison with the approaches and results of other programs. The parametrization of second generation AMBER force field (Cornell et al. 1995) had been influenced by Jorgensen OPLS potentials for proteins (Jorgensen and Tirado-Rives 1988), whereas these OPLS united atom potentials adopted bond stretch, angle bend, and torsional terms from the AMBER united-atom force field (Weiner et al. 1984). The OPLS parameters for the 20 neutral peptide residues were obtained primarily via Monte Carlo simulations for the 36 organic liquids, and the parametrization for the charged protein residues was performed via comparisons with ab initio results for ion-molecule complexes (6-31G(d) basis set) and on Monte Carlo simulations for hydrated ions (Jorgensen and Tirado-Rives 1988). The OPLS-AA (all atom) force field (Jorgensen et al. 1996) retain most of bond stretch and angle bending parameters from AMBER all-atom force fields, but torsion and nonbonding constants are reparametrized utilizing both experimental and ab initio (RHF/6-31G* level) data. This force field was improved for peptides by means of refitting the key Fourier torsional coefficients using accurate ab initio data (LMP2/cc-pVTZ (-f)/HF/6-31G** level) as the target (Kaminski et al. 2001).

Most of popular program sets utilize various force fields, not only those developed in the research groups of the authors of the program with the same title. The software suite GROMACS was developed by Berendsen group at the University of Groningen, The Netherlands, starting in early 1990s. This fast program for molecular dynamics simulation does not have a force field of its own but is compatible with GROMOS, OPLS, and AMBER force fields. The GROMACS (van der Spoel et al. 2005) was developed and optimized especially for use on PC-clusters. It originated from the Fortran package GROMOS developed by van Gunsteren et al. (1987). The GROMOS force fields are not pure all-atom force fields; aliphatic CH_n groups are treated as united atoms. The recent version, GROMOS05 (Christen et al. 2005), utilize new force field sets (53A5 and 53A6, the first number in the names of a parameter sets refer to the number of atom types) based on extensive reparameterization of the previous GROMOS force field (Oostenbrink et al. 2004). In contrast to the parameterization of most other biomolecular force fields, this parameterization is based primarily on reproducing the free enthalpies of hydration and apolar solvation for a range of organic compounds. Recently, 53A6 force field has been reparametrized by Pol-Fachin et al. (2014) by adjustment to QM data (HF/6-31G* level of theory) for applications to study Glycoproteins.

Very popular for the computations of wide range of macromolecular systems are CHARMM (Chemistry at HARvard Macromolecular Mechanics) program and force field sets. The computations include energy minimization, molecular

dynamics, and Monte Carlo simulations. CHARMM originated at Martin Karplus's group at Harvard University, the first publication on CHARMM program and force field (refer to Brooks et al. 1983). Several versions of both software and force field were released since the first ones. The improvement of the programs and force fields are in progress till now. The mathematical expressions for calculation of CHARMM energy are nearly the same as "minimalist" ones for other force fields (MacKerell 2004). The minor differences with e.g., AMBER or OPLS force fields refer to inclusion of harmonic improper and Urey-Bradley terms. The parameters of various versions of force field are consistently adjusted with emphasis to specific molecules to be applied. The first versions of CHARMM force fields were aimed primarily to protein molecular dynamics simulations. Then, the special adjustments were performed for the simulations of the nucleic acids (Foloppe and MacKerell 2000) and lipids (Klauda et al. 2008). The improvement of parameter for specific molecules continues till recent times, while other parameters can be used from the previous versions. Raman et al. (2010) recently extended the CHARMM additive carbohydrate all-atom force field to enable modeling glycosidic linkages in carbohydrates involving furanoses. Like in most popular modern force fields, both the ab initio quantum mechanics computations and experimental data for solid and liquid states are used in adjustment of the parameters of CHARMM force fields. The distinct feature of this adjustment refers to charge adsorption. Like OPLS force fields (Jorgensen et al. 1996), CHARMM utilize so-called supramolecular approach. The partial atomic charges are adjusted to reproduce ab initio (HF/6-31G* level) minimum interaction energies and geometries of model compounds with water or for model compound dimers (with energy and distance corrections for reproducing the correct experimental densities). On the other hand, the use of ab initio data on small clusters to optimize the van der Waals parameters leads to poor condensed phase properties. This requires that the optimization of these parameters be performed by empirical fitting to reproduce thermodynamics properties from condensed phase simulations, taking into account the relative values of the parameters obtained from high-level ab initio data on interactions of the model compounds with rare gases (Klauda et al. 2008).

Very recently, Vanommeslaeghe and MacKerell (2015) published detailed description of the development of CHARMM method and software. We will mention here only some most important for biological and medical applications recent advancements of both all-atom additive force field (on the way to CHARMM36 version, modern class I model) and polarizable force fields (class II models). Best et al. (2012) presented a combination of refinements important with respect to the equilibrium between helical and extended conformations in protein folding simulations. Peptide backbone (φ , ψ) potential has been refined against experimental solution NMR data. Side-chain χ_1 and χ_2 torsion parameters have been optimized by fitting to quantum-mechanical energy surfaces, followed by additional empirical optimization targeting NMR data. A comprehensive validation of the revised force field indicates that the revised CHARMM 36 parameters represent an improved model for modeling and simulation studies of proteins, including protein folding, assembly, and functionally relevant conformational

changes (Best et al. 2012). The refinement of the CHARMM27 all-atom additive force field for nucleic acids via adjustment of backbone parameters to QM data (MP2/6-31 + G(d) level of theory) followed by trial MD simulations of DNA duplexes enabled Hart et al. (2012) to improve treatment of the BI/BII equilibrium of DNA conformations. Many other specific refinements of the CHARMM additive force fields for these and other classes of biologically important molecules have been mentioned in the recent paper of Vanommeslaeghe and MacKerell (2015).

Along with comprehensive refinements and biomolecular applications of class I CHARMM force fields, polarizable CHARMM models are extensively developed during recent years for more accurate and physically based description of biologically important molecular systems. These force fields use the classical Drude oscillator (i.e., “charge-on-spring”) model. In the model, a virtual negatively charged particle (“Drude particle”) is connected to the core of each polarizable atom by means of a harmonic potential (“spring”). Description of this class II force field is out of scope of this survey, but we will mention some results of proteins and nucleic acid computations with such models. Very recently, Huang et al. (2014) summarized recent efforts to include the explicit treatment of induced electronic polarization in biomolecular force fields. They arrive with conclusion that the inclusion of explicit electronic polarizability leads to significant differences in the physical forces affecting the structure and dynamics of proteins (Huang et al. 2014). The polarizable force field more accurately models peptide-folding cooperativity (Huang and MacKerell 2014). Development of a CHARMM Drude polarizable model for nucleic acids (Baker et al. 2011) requires significant adjustments of the phosphodiester backbone and glycosidic bond dihedral parameters of base-sodium interactions and of selected electrostatic parameters in the model (Savelyev and MacKerell 2014a). The modern polarizable model satisfactorily treats the equilibrium between the A, BI, and BII conformations of DNA (Savelyev and MacKerell 2014b).

It is interesting to mention that multiple adjustments of different additive force fields mentioned above and other force fields result in an increase of similarity in both corresponding parameters values and the results of applications to complex molecular systems (such as protein and nucleic acid fragments), but many differences both in corresponding parameter values and computation results still remain.

Conclusions

The molecular mechanics computation approach to the study of the molecular systems arose a half of century ago and today has become a useful and powerful research tool for many branches of natural science. During this period the number of publications utilizing this approach increased by a few orders of magnitude. This increase continues up to present days, followed by new journals appearance and a flow of MM publications in the journals of rather wide area of coverage. The MM considers the molecular systems via classic Newtonian mechanics approach, i.e., using classical approximation to essentially quantum systems. The MM suggests

representation of the system energy as a sum of terms responsible for interactions of various physical natures.

The minimalist MM approach involves the simple expressions for chemical bond stretching, bond angle bending, torsions about the bonds, and nonbonded van der Waals and Coulomb contributions calculated via additive atom-atom scheme. The force fields (sets of mathematical formulae and numerical parameters) are derived from simple physical considerations followed by adjustment to the experimental data and precise quantum mechanics results. An additivity of the terms responsible for separate atom contributions and for interactions of different physical nature and transferability of the parameters between atoms and molecules of similar structure are the main assumptions of MM approach. Development of the approach results in elaboration of a variety of computer software and force fields for simulating different molecular systems of biological, medical, and industrial importance. Various force fields differ in the complexity of mathematical expressions and in the relative role of experimental and quantum mechanics results used for the parameter adjustment. The development of algorithms for the MM applications (including those for molecular dynamics and Monte Carlo techniques) and the increase of computer power enable researchers to approach such complex problems as predicting the pathways of the biopolymers three-dimensional structure, molecular recognition in various biological processes, assistance in creation of compounds with desirable properties (including drugs and industrial materials). Analysis of the current state of physicochemical study of biological processes suggests that MM simulations based on empirical force fields have an ever-increasing impact on understanding structure and functions of biomolecules. Rather close results obtained with different force fields can be accounted for by existing various solutions of the task of parameter adjustment even to the same target data (really, target data used when elaborating various force fields differ, slightly or considerably). Force fields are internally consistent. Their parameters are highly correlated within a force field. Several force fields have been elaborated recently going beyond the atom-atom approach and the additivity principle. Some force fields utilize a hierarchical procedure for consideration of complex systems.

Bibliography

- Allinger, N. L. (1959). Conformational analysis. III. Applications to some medium ring compounds. *Journal of the American Chemical Society*, 81, 5727.
- Allinger, N. L. (2010). *Molecular structure: Understanding steric and electronic effects from molecular mechanics*. Hoboken: Wiley.
- Allinger, N. L. (2011). Understanding molecular structure from molecular mechanics. *Journal of Computer-Aided Molecular Design*, 25, 295.
- Allinger, N. L., & Sprague, J. T. (1973). Calculation of the structures of hydrocarbons containing delocalized electronic systems by the molecular mechanics method. *Journal of the American Chemical Society*, 95, 3893.
- Allinger, N. L., Tribble, M. T., Miller, M. A., & Wertz, D. W. (1971). Conformational analysis. LXIX. Improved force field for the calculation of the structures and energies of hydrocarbons. *Journal of the American Chemical Society*, 93, 1637.

- Allinger, N. L., Yuh, Y. H., & Lii, J. H. (1989). Molecular mechanics. The MM3 force field for hydrocarbons. *Journal of the American Chemical Society*, *111*, 8551.
- Allinger, N. L., Chen, K., & Lii, J. H. (1996). An improved force field (MM4) for saturated hydrocarbons. *Journal of Computational Chemistry*, *17*, 642.
- Antony, J., & Grimme, S. (2006). Density functional theory including dispersion corrections for intermolecular interactions in a large benchmark set of biologically relevant molecules. *Physical Chemistry Chemical Physics*, *8*, 5287.
- Arnautova, Y. A., & Scheraga, H. A. (2008). Use of decoys to optimize an all-atom force field including hydration. *Biophysical Journal*, *95*, 2434.
- Arnautova, Y. A., Jagielska, A., & Scheraga, H. A. (2006). A new force field (ECEPP-05) for peptides, proteins and organic molecules. *The Journal of Chemical Physics*, *110*, 5025.
- Baker, C. M., Anisimov, V. M., & MacKerell, A. D., Jr. (2011). Development of CHARMM polarizable force field for nucleic acid bases based on the classical Drude oscillator model. *The Journal of Physical Chemistry B*, *115*, 580.
- Bartell, L. S. (1960). On the effects of intramolecular van der Waals forces. *The Journal of Chemical Physics*, *32*, 827.
- Barton, D. H. R. (1948). Interaction between non-bonded atoms, and the structure of cis-decalin. *Journal of the Chemical Society*, 340.
- Barton, D. H. R. (1950). The conformation of the steroid nucleus. *Experientia*, *6*, 316.
- Berendsen, H. J. C., Postma, J. P. M., von Gunstaren, W. F., & Hermans, J. (1981). Interaction models for water in relation to protein hydration. In B. Pullman (Ed.), *Intermolecular forces* (pp. 331–342). Dordrecht: Reidel.
- Berman, H. M., Olson, W. K., Beveridge, D. I., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R., & Schneider, B. (1992). The nucleic acid database. A comprehensive relational database of three-dimensional structures of nucleic acids. *Biophysical Journal*, *63*, 751.
- Berman, H. M., Henrick, K., & Nakamura, H. (2003). Announcing the worldwide protein data bank. *Nature Structural & Molecular Biology*, *10*, 980.
- Bernal, J. D., & Fowler, R. H. (1933). A theory of water and ionic solution, with particular reference to hydrogen and hydroxyl ions. *The Journal of Chemical Physics*, *1*, 515.
- Best, R. B., Zhu, X., Shim, J., Lopes, P. E. M., Mittal, J., Feig, M., & MacKerell, A. D., Jr. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles. *Journal of Chemical Theory and Computation*, *8*, 3257.
- Bordner, A. J., Civasotto, C. N., & Abagyan, R. A. (2003). Direct derivation of van der Waals force field parameters from quantum mechanical interaction energies. *The Journal of Physical Chemistry B*, *107*, 9601.
- Bradley, D. F., Lifson, S., & Honig, B. (1964). Theory of optical and other properties of biopolymers: Applicability and elimination of the first-neighbor and dipole-dipole approximations. In B. Pullman (Ed.), *Electronic aspects of biochemistry*. New York: Academic.
- Brandenburg, J. G., & Grimme, S. (2014). Accurate modeling of organic molecular crystals by dispersion-corrected density functional tight binding (DFTB). *The Journal of Physical Chemistry Letters*, *5*, 1785.
- Brant, D. A., & Flory, P. J. (1965). The configuration of random polypeptide chains. II. Theory. *Journal of the American Chemical Society*, *87*, 2791.
- Brooks, B. R., Brucoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, *4*, 187.
- Case, D. A., Cheatham, T. E., III, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., & Woods, R. J. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, *26*, 1668.
- Case, D. A., Darden, T. A., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S.

- R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., & Kollman, P. A. (2010). *AMBER 11*. San Francisco: University of California.
- Case, D. A., Berryman, J. T., Betz, R. M., Cerutti, D. S., Cheatham, T. E., III, Darden, T., Duke, R. E., Giese, T. J., Gohlke, H., Goetz, A. W., Homeyer, N., Izadi, S., Janowski, P., Kaus, J., Kovalenko, A., Lee, T. S., LeGrand, S., Li, P., Luchko, T., Luo, R., Madej, B., Merz, K. M., Monard, G., Needham, P., Nguyen, H., Nguyen, H. T., Omelyan, I., Onufriev, A., Roe, D. R., Roitberg, A., Salomon-Ferrer, R., Simmerling, C. L., Smith, W., Swails, J., Walker, R. C., Wang, J., Wolf, R. M., Wu, X., York, D. M., & Kollman, P. A. (2015). *AMBER 2015*. San Francisco: University of California.
- Christen, M., Hunenberger, P. H., Bakowies, D., Baron, R., Burgi, R., Geerke, D. P., Heinz, T. N., Kastenholz, M. A., Krautler, V., Oostenbrink, C., Peter, C., Trzesniak, D., & Van Gunsteren, W. F. (2005). The GROMOS software for biomolecular simulation: GROMOS05. *Journal of Computational Chemistry*, 26, 1719.
- Chuprina, V. P., & Poltev, V. I. (1983). Possible conformations of double-helical polynucleotides containing incorrect base-pairs. *Nucleic Acids Research*, 11, 5205.
- Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I., Merz, K., Jr., Ferguson, D., Spellmeyer, D., Fox, T., Caldwell, J., & Kollman, P. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117, 5179.
- Craig, D. P., Mason, R., Pauling, P., & Santry, D. P. (1965). Molecular packing in crystals of the aromatic hydrocarbons. *Proceedings of the Royal Society A*, 286, 98.
- Cramer, C. J. (2004). *Essentials of computational chemistry: Theories and models*. Chichester: Wiley.
- De Santis, P. (1992). Conformational energy calculations of macromolecules. *Current Contents*, 34, 8.
- De Santis, P., Giglio, E., Liquori, A. M., & Ripamonti, A. (1963). Stability of helical conformations of simple linear polymers. *Journal of Polymer Science: Part A*, 1, 1383.
- De Santis, P., Giglio, E., Liquori, A. M., & Ripamonti, A. (1965). Interaction and stability of helical polypeptide chains. *Nature*, 206, 456.
- De Voe, H., & Tinoco, I., Jr. (1962). The stability of helical polynucleotides: Base contributions. *Journal of Molecular Biology*, 4, 500.
- Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., & Kollman, P. (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry*, 24, 1999.
- Eisenberg, D. (2003). The discovery of the α -helix and β -sheet, the principal structural features of proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 11207.
- Engler, E. M., Andose, J. D., & Schleyer, P. R. (1973). Critical evaluation of molecular mechanics. *Journal of the American Chemical Society*, 95, 8005.
- Foloppe, N., & MacKerell, A. D., Jr. (2000). All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *Journal of Computational Chemistry*, 21, 86.
- Gibson, K. D., & Scheraga, H. A. (1967a). Minimization of polypeptide energy. I. Preliminary structures of bovine pancreatic ribonuclease S-peptide. *Proceedings of the National Academy of Sciences of the United States of America*, 58, 420.
- Gibson, K. D., & Scheraga, H. A. (1967b). Minimization of polypeptide energy. II. Preliminary structures of oxytocin, vasopressin, and an octapeptide from ribonuclease. *Proceedings of the National Academy of Sciences of the United States of America*, 58, 1317.
- Golas, E., Maisuradze, G. G., Senet, P., Oldziej, S., Czaplowski, C., Scheraga, H. A., & Liwo, A. (2012). Simulation of the opening and closing of Hsp70 chaperones by coarse-grained molecular dynamics. *Journal of Chemical Theory and Computation*, 8, 1750.

- Gresh, N., Claverie, P., & Pullman, A. (1986). Intermolecular interactions: Elaboration on an additive procedure including an explicit charge-transfer contribution. *International Journal of Quantum Chemistry*, 29, 101.
- Grimme, S. (2004). Accurate description of van der Waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry*, 25, 1463.
- Grimme, S. (2011). Density functional theory with London dispersion corrections. *WIREs Computational Molecular Science*, 1, 211.
- Grubišić, S., Brancato, G., & Barone, V. (2013). An improved AMBER force field for α , α -dialkylated peptides: Intrinsic and solvent-induced conformational preferences of model systems. *Physical Chemistry Chemical Physics*, 15, 17395.
- Guo, X., Wang, Z., Zuo, L., Zhou, Z., Guo, X., & Sun, T. (2014). Quantitative prediction of enantioseparation using β -cyclodextrin derivatives as chiral selectors in capillary electrophoresis. *Analyst*, 139, 6511.
- Halgren, T. A. (1996). Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17, 490.
- Halgren, T. A. (1999a). MMFF VI. MMFF94s option for energy minimization studies. *Journal of Computational Chemistry*, 20, 720.
- Halgren, T. A. (1999b). MMFF VII. Characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. *Journal of Computational Chemistry*, 20, 730.
- Hart, K., Foloppe, N., Baker, C. M., Denning, E. J., Nilsson, L., & MacKerell, A. D., Jr. (2012). Optimization of the CHARMM additive force field for DNA: Improved treatment of the BI/BII conformational equilibrium. *Journal of Chemical Theory and Computation*, 8, 348.
- He, Y., Maciejczyk, M., Oldziej, S., Scheraga, H. A., & Liwo, A. (2013). Mean-field interactions between nucleic-acid-base dipoles can drive the formation of the double helix. *Physical Review Letters*, 110, 098101.
- Hendrickson, J. B. (1961). Molecular geometry. I. Machine computation of the common rings. *Journal of the American Chemical Society*, 83, 4537.
- Hendrickson, J. B. (1962). Molecular geometry. II. Methyl-cyclohexanes and cycloheptanes. *Journal of the American Chemical Society*, 84, 3355.
- Hendrickson, J. B. (1973). Molecular geometry. VIII. Proton magnetic resonance studies of cycloheptane conformations. *Journal of the American Chemical Society*, 95, 494.
- Hill, T. L. (1946). On steric effects. *The Journal of Chemical Physics*, 14, 465.
- Hill, T. L. (1948). Steric effects. I. Van der Waals potential energy curves. *The Journal of Chemical Physics*, 16, 399.
- Huang, J., & MacKerell, A. D., Jr. (2014). Induction of peptide bond dipoles drives cooperative helix formation in the (AAQAA)₃ peptide. *Biophysical Journal*, 107, 991.
- Huang, J., Lopes, P. E. M., Roux, B., & MacKerell, A. D., Jr. (2014). Recent advances in polarizable force fields for macromolecules: Microsecond simulations of proteins using the classical Drude oscillator model. *The Journal of Physical Chemistry Letters*, 5, 3144.
- Jorgensen, W. L., & Tirado-Rives, J. (1988). The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110, 1657.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79, 926.
- Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118, 11225.
- Jurecka, P., Cerny, J., Hobza, P., & Salahub, D. R. (2007). Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of noncovalent complexes compared with ab initio quantum mechanics calculations 80. *Journal of Computational Chemistry*, 28, 555.

- Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., & Jorgensen, W. L. (2001). Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *The Journal of Physical Chemistry B*, *105*, 6474.
- Khutorsky, V. E., & Poltev, V. I. (1976). Conformations of double-helical nucleic acids. *Nature*, *264*, 483.
- Kitaigorodski, A. I. (1959). *Organic chemical crystallography*. New York: Consultants Bureau.
- Kitaigorodsky, A. I. (1973). *Molecular crystals and molecules*. New York: Academic.
- Kitaygorodsky, A. I. (1960). Calculation of conformations of organic molecules. *Tetrahedron*, *9*, 183.
- Kitaygorodsky, A. I. (1961). The interaction curve of non-bonded carbon and hydrogen atoms and its application. *Tetrahedron*, *14*, 230.
- Klauda, J. B., Venable, R. M., MacKerell, A. D., Jr., & Pastor, R. W. (2008). Considerations for lipid force field development. *Current Topics in Membranes*, *60*, 1.
- Lam, A. R., Rodriguez, J. J., Rojas, A., Scheraga, H. A., & Mukamel, S. (2013). Tracking the mechanism of fibril assembly by simulated two-dimensional ultraviolet spectroscopy. *The Journal of Physical Chemistry A*, *117*, 342.
- Langlet, J., Claverie, P., Caron, F., & Boeue, J. C. (1981). Interactions between nucleic acid bases in hydrogen bonded and stacked configurations: The role of the molecular charge distribution. *International Journal of Quantum Chemistry*, *20*, 299.
- Laury, M. L., Wang, L.-P., Pande, V. S., Head-Gordon, T. L., & Ponder, J. W. (2015). Revised parameters for the AMOEBA polarizable atomic multipole water model. *Journal of Physical Chemistry B*. doi:10.1021/jp510896n.
- Leach, A. R. (2001). *Molecular modelling: Principles and applications*. Harlow: Prentice Hall (Pearson Education).
- Leach, S. J., Némethy, G., & Scheraga, H. A. (1966a). Computation of the sterically allowed conformations of peptides. *Biopolymers*, *4*, 369.
- Leach, S. J., Némethy, G., & Scheraga, H. A. (1966b). Intramolecular steric effects and hydrogen bonding in regular conformations of polyamino acids. *Biopolymers*, *4*, 887.
- Levitt, M., & Lifson, S. (1969). Refinement of protein conformations using a macromolecular energy minimization procedure. *Journal of Molecular Biology*, *46*, 269.
- Lewars, E. G. (2011). *Computational chemistry. Introduction to the theory and applications of molecular and quantum mechanics* (2nd ed.). Dordrecht/Heidelberg/London/New York: Springer.
- Lifson, S., & Warshel, A. (1968). Consistent force field for calculations of conformations, vibrational spectra, and enthalpies of cycloalkane and *n*-alkane molecules. *The Journal of Chemical Physics*, *49*, 5116.
- Lii, L.-H., & Allinger, N. L. (1991). The MM3 force field for amides, polypeptides and proteins. *Journal of Computational Chemistry*, *12*, 186.
- Lii, L.-H., Chen, K.-H., Johnson, G. P., French, A. D., & Allinger, N. L. (2005). The external-anomeric torsional effect. *Carbohydrate Research*, *340*, 853.
- Liwo, A., Lee, J., Ripoll, D. R., Pillardy, J., & Scheraga, H. A. (1999). Protein structure prediction by global optimization of a potential energy function. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 5482.
- Lorecchio, C., Venanzi, M., Mazzuca, C., Lettieri, R., Palleschi, A., Thi, T. H. N., Cardová, L., Drasar, P., & Monti, D. (2014). Tuning the chiroptical and morphological properties of steroidal-porphyrin aggregates: A mechanistic, structural, and MM investigation. *Organic & Biomolecular Chemistry*, *12*, 3956.
- Maciejczyk, M., Spasic, A., Liwo, A., & Scheraga, H. A. (2010). Coarse-grained model of nucleic acid bases. *Journal of Computational Chemistry*, *31*, 1644.
- MacKerell, A. D., Jr. (2004). Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, *25*, 1584.
- Mahoney, M. W., & Jorgensen, W. L. (2001). Quantum, intramolecular flexibility, and polarizability effects on the reproduction of the density anomaly of liquid water by simple potential functions. *The Journal of Chemical Physics*, *115*, 10758.

- Mason, R. (1969). The intermolecular potential and structure of crystals of aromatic molecules. *Molecular Crystals and Liquid Crystals*, 9, 3.
- Matsuoka, O., Clementi, E., & Yoshimine, M. (1976). CI study of the water dimer potential surface. *The Journal of Chemical Physics*, 64, 1351.
- McAllister, S. R., & Floudas, C. A. (2010). An improved hybrid global optimization method for protein tertiary structure prediction. *Computational Optimization and Applications*, 45, 377.
- McGuire, R. F., Momany, F. A., & Scheraga, H. A. (1972). Energy parameters in polypeptides. V. An empirical hydrogen bond potential function based on molecular orbital calculations. *The Journal of Physical Chemistry*, 76, 375.
- Momany, F. A., Vanderkooi, G., & Scheraga, H. A. (1968). Determination of intermolecular potentials from crystal data. I. General theory and application to crystalline benzene at several temperatures. *Proceedings of the National Academy of Sciences of the United States of America*, 61, 429.
- Momany, F. A., Carruthers, L. M., & Scheraga, H. A. (1974). Intermolecular potentials from crystal data. IV. Application of empirical potentials to the packing configurations and lattice energies in crystals of amino acids. *The Journal of Physical Chemistry*, 78, 1621.
- Momany, F. A., McGuire, R., Burgess, A., & Scheraga, H. (1975). Energy parameters in polypeptides. VII. Geometric parameters, partial atomic charges, nonbonded interactions, hydrogen bond interactions, and intrinsic torsional potentials for the naturally occurring amino acids. *The Journal of Physical Chemistry*, 79, 2361.
- Nada, H., & van der Eerden, J. P. J. M. (2003). An intermolecular potential model for the simulation of ice and water near the melting point: A six-site model of H₂O. *The Journal of Chemical Physics*, 118, 7401.
- Nash, H. A., & Bradley, D. F. (1966). Calculation of the lowest energy configurations of nucleotide base pairs on the basis of an electrostatic model. *The Journal of Chemical Physics*, 45, 1380.
- Némethy, G., & Scheraga, H. A. (1965). Theoretical determination of sterically allowed conformations of a polypeptide chain by a computer method. *Biopolymers*, 3, 155.
- Némethy, G., Gibson, K. D., Palmer, K. A., Yoon, C. N., Paterlini, G., Zagari, A., Rumsey, S., & Scheraga, H. A. (1992). Energy parameters in polypeptides. 10. Improved geometrical parameters and nonbonded interactions for use in the ECEPP/3 algorithm, with application to proline-containing peptides. *The Journal of Physical Chemistry*, 96, 6472.
- Ogata, K., & Nakamura, S. (2015). Improvement of parameters of the AMBER potential force field for phospholipids for description of thermal phase transitions. *The Journal of Physical Chemistry*. doi:10.1021/acs.jpcc.5b01656.
- Oostenbrink, C., Villa, A., Mark, A. E., & van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, 25, 1656.
- Pauling, L., & Corey, R. B. (1951). The pleated sheet, a new layer configuration of polypeptide chains. *Proceedings of the National Academy of Sciences of the United States of America*, 37, 251.
- Pauling, L., Corey, R. B., & Branson, H. R. (1951). The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences of the United States of America*, 37, 205.
- Pawar, S., Sawant, S., Nerkar, A., & Bhosale, A. (2014). In silico design, synthesis and pharmacological screening of novel 2-(6-substituted benzo [d] thiazol-2-yl) isoindoline-1, 3-diones as potential COX-2 inhibitors for anti-inflammatory activity. *International Journal of Pharmacy and Pharmaceutical Sciences*, 6, 353.
- Perez, A., Marchan, I., Svozil, D., Sponer, J., Cheatham, T. E., Laughton, C. A., & Orozco, M. (2007). Refinement of the AMBER force field for nucleic acids: Improving the description of α/γ conformers. *Biophysical Journal*, 92, 3817.
- Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., & Klaus Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26, 1781.

- Pinilla, C., Irani, A. H., Nicola Seriani, N., & Scandolo, S. (2012). Ab initio parameterization of an all-atom polarizable and dissociable force field for water. *The Journal of Chemical Physics*, *136*, 114511.
- Pol-Fachin, L., Hugo Verli, H., & Lins, R. D. (2014). Extension and validation of the GROMOS 53A6_{glyc} parameter set for glycoproteins. *Journal of Computational Chemistry*, *35*, 2087.
- Polozov, R. V., Poltev, V. I., & Sukhorukov, B. I. (1973). Relation of the interactions of nucleic acid bases to the helical conformations of polynucleotides. *Studia Biophysica*, *40*, 13.
- Poltev, V. I., & Bruskov, V. I. (1978). On molecular mechanisms of nucleic acid synthesis fidelity aspects I. Contribution of base interactions. *Journal of Theoretical Biology*, *70*, 69.
- Poltev, V. I., & Shulyupina, N. V. (1986). Simulation of interactions between nucleic-acid bases by refined atom-atom potential functions. *Journal of Biomolecular Structure & Dynamics*, *3*, 739.
- Poltev, V. I., & Sukhorukov, B. I. (1967). Theoretical examination of the physical nature of the intermolecular interactions determining the conformational state of polynucleotides. *Biophysics (Moscow)*, *12*, 879.
- Poltev, V. I., & Sukhorukov, B. I. (1970). Semiempirical calculations of interaction energy of DNA nitrous bases. *Studia Biophysica*, *24/25*, 179.
- Poltev, V. I., Grokhlina, T. I., & Malenkov, G. G. (1984). Hydration of nucleic-acid bases studied using novel atom-atom potential functions. *Journal of Biomolecular Structure & Dynamics*, *2*, 413.
- Ponder, J. W. (2015). TINKER – Software tools for molecular design. <http://dasher.wustl.edu/tinker/>
- Pullman, A., & Pullman, B. (1968). Aspects of the electronic structure of the purine and pyrimidine bases of the nucleic acids and of their interactions. *Advances in Quantum Chemistry*, *4*, 267.
- Pullman, B., Claverie, P., & Caillet, J. (1966). Van der Waals-London interactions and the configuration of hydrogen-bonded purine and pyrimidine pairs. *Proceedings of National Academy of Sciences of United States of America*, *55*, 904.
- Rae, A. I. M., & Mason, R. (1968). The intermolecular potential and the lattice energy of benzene. *Proceedings of the Royal Society A*, *304*, 487.
- Ramachandran, G. N. (1990). This week's citation classic. *Current Contents*, *10*, 119.
- Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology*, *7*, 95.
- Ramachandran, K. I., Deepa, G., & Namboori, K. (2008). *Computational chemistry and molecular modeling: Principles and applications*. Berlin: Springer.
- Raman, E. P., Guvench, O., & MacKerell, A. D., Jr. (2010). CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *The Journal of Physical Chemistry B*, *114*, 12981.
- Rasse, D., Warme, P. K., & Scheraga, H. A. (1974). Refinement of the X-ray structure of rubredoxin by conformational energy calculations. *Proceedings of the National Academy of Sciences of the United States of America*, *71*, 3736.
- Ren, P., & Ponder, J. W. (2003). Polarizable atomic multipole water model for molecular mechanics simulation. *The Journal of Physical Chemistry B*, *107*, 5933.
- Ren, P., & Ponder, J. W. (2004). Temperature and pressure dependence of the AMOEBA water model. *The Journal of Physical Chemistry B*, *108*, 13427.
- Ren, P., Wu, C., & Ponder, J. W. (2011). Polarizable atomic multipole-based molecular mechanics for organic molecules. *Journal of Chemical Theory and Computation*, *7*, 314.
- Renugopalakrishnan, V., Lakshminarayanan, A. V., & Sasisekharan, V. (1971). Stereochemistry of nucleic acids and polynucleotides III. Electronic charge distribution. *Biopolymers*, *10*, 1159.
- Ripoll, D. R., & Scheraga, H. A. (1988). On the multiple-minima problem in the conformational analysis of polypeptides. II. An electrostatically driven Monte Carlo method – Tests on poly(L-alanine). *Biopolymers*, *27*, 1283.
- Rojas, A. V., Liwo, A., & Scheraga, H. A. (2011). A study of the α -helical intermediate preceding the aggregation of the amino-terminal fragment of the β amyloid peptide (A β 1–28). *The Journal of Physical Chemistry B*, *115*, 12978.

- Savelyev, A., & MacKerell, A. D., Jr. (2014a). Balancing the interactions of ions, water and DNA in 1314 the Drude polarizable force field. *The Journal of Physical Chemistry B*, *118*, 6742.
- Savelyev, A., & MacKerell, A. D., Jr. (2014b). All-atom polarizable force field for DNA based on 1317 the classical Drude oscillator model. *Journal of Computational Chemistry*, *35*, 1219.
- Sawant, R. L., Hardas, D. B., Pawa, K. K., & Shinde, A. K. (2014). QSAR analysis of structurally similar 1, 3, 4-oxadiazole/thiadiazole and 1, 2, 4-triazole derivatives of biphenyl-4-yloxy acetic acid as antiinflammatoryagents. *World Journal of Pharmaceutical Research*, *3*, 1844.
- Scheraga, H. A. (2008). From helix-coil transitions to protein folding. *Biopolymers*, *89*, 479.
- Scheraga, H. A. (2015). My 65 years in protein chemistry. *Quarterly Reviews of Biophysics*, *48*, 117.
- Scheraga, H. A., Pillardy, J., Liwo, A., Lee, J., Czaplowski, C., Ripoll, D. R., Wedemeyer, W. J., & Arnautova, Y. A. (2002). Evolution of physics-based methodology for exploring the conformational energy landscape of proteins. *Journal of Computational Chemistry*, *23*, 28.
- Scott, R. A., & Scheraga, H. A. (1966a). Conformational analysis of macromolecules. III. Helical structures of poly-glycine and poly-L-alanine. *The Journal of Chemical Physics*, *45*, 2091.
- Scott, R. A., & Scheraga, H. A. (1966b). Conformational analysis of macromolecules. II. The rotational isomeric states of the normal hydrocarbons. *Journal of Chemical Physics*, *44*, 3054.
- Shi, Y., Xia, Z., Zhang, J., Best, R., Wu, C., Ponder, J. W., & Ren, P. (2013). Polarizable atomic multipole-based AMOEBA force field for proteins. *Journal of Chemical Theory and Computation*, *9*, 4046.
- Shipman, L. L., Burgess, A. W., & Scheraga, H. A. (1975). A new approach to empirical intermolecular and conformational potential energy functions. I. Description of model and derivation of parameters. *Proceedings of the National Academy of Sciences of the United States of America*, *72*, 543.
- Sippl, M. J., Némethy, G., & Scheraga, H. A. (1984). Intermolecular potentials from crystal data. 6. Determination of empirical potentials for O-H...O=C hydrogen bonds from packing configurations. *Journal of Physics Chemistry*, *88*, 6231.
- Snir, J., Nemenoff, R. A., & Scheraga, H. A. (1978). A revised empirical potential for conformational, intermolecular, and solvation studies. 5. Development and testing of parameters for amides, amino acids and peptides. *The Journal of Physical Chemistry*, *82*, 2527.
- Song, K., Hornak, V., de los Santos, C., Grollman, A. P., & Simmerling, C. (2008). Molecular mechanics parameters for the FapydG DNA lesion. *Journal of Computational Chemistry*, *29*, 17.
- Stillinger, F. H., & Rahman, A. (1974). Improved simulation of liquid water by molecular dynamics. *The Journal of Chemical Physics*, *60*, 1545.
- Tröster, P., Lorenzen, K., & Tavan, P. (2014). Polarizable six-point water models from computational and empirical optimization. *The Journal of Physical Chemistry B*, *118*, 1589.
- Van der Spoel, D., Lindahl, E., Hess, B., Groehof, G., Mark, A. E., & Berendsen, H. J. C. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, *26*, 1701.
- Van Gunsteren, W., Fand, H. J., & Berendsen, C. (1987). *Groningen molecular simulation (GROMOS) library manual*. Groningen: BIOMOS.
- Vanommeslaeghe, K., & MacKerell, A. D., Jr. (2015). CHARMM additive and polarizable force fields for biophysics and computer-aided drug design. *Biochimica et Biophysica Acta, General Subjects*, *1850*, 861.
- Wang, J., Cieplak, P., & Kollman, P. A. (2000). How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry*, *21*, 1049.
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and testing of a general AMBER force field. *Journal of Computational Chemistry*, *25*, 1157.
- Wang, J., Wang, W., Kollman, P. A., & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics & Modeling*, *25*, 247.
- Wang, L.-P., Chen, J., & van Voorhis, T. (2013a). Systematic parametrization of polarizable force fields from quantum chemistry data. *Journal of Chemical Theory and Computation*, *9*, 452.

- Wang, L.-P., Head-Gordon, T., Ponder, J. W., Ren, P., Chodera, J. D., Eastman, P. K., Martinez, T. J., & Pande, V. S. (2013b). Systematic improvement of a classical molecular model of water. *The Journal of Physical Chemistry B*, *117*, 9956.
- Warne, P. K., & Scheraga, H. A. (1974). Refinement of the X-ray structure of lysozyme by complete energy minimization. *Biochemistry*, *13*, 757.
- Warne, P. K., Momany, F. A., Rumball, S. V., & Scheraga, H. A. (1974). Computation of structures of homologous proteins; α -lactalbumin from lysozyme. *Biochemistry*, *13*, 768.
- Warshel, A., & Lifson, S. (1970). Consistent force field calculations. II. Crystal structures, sublimation energies, molecular and lattice vibrations, molecular conformations, and enthalpies of alkanes. *The Journal of Chemical Physics*, *53*, 582.
- Watson, J. D., & Crick, F. H. C. (1953). A structure for deoxyribose nucleic acid. *Nature*, *171*, 737.
- Weiner, P., & Kollman, P. (1981). AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *Journal of Computational Chemistry*, *2*, 287.
- Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., & Weiner, P. (1984). A new force field for molecular mechanical simulation of nucleic acids and proteins. *Journal of the American Chemical Society*, *106*, 765.
- Westheimer, F. H., & Mayer, J. E. (1946). The theory of the racemization of optically active derivatives of diphenyl. *The Journal of Chemical Physics*, *14*, 733.
- Wiberg, K. B. (1965). A scheme for strain energy minimization. Application to the cycloalkanes. *Journal of the American Chemical Society*, *87*, 1070.
- Williams, D. E. (1966). Nonbonded potential parameters derived from crystalline aromatic hydrocarbons. *The Journal of Chemical Physics*, *45*, 3770.
- Williams, D. E. (1967). Nonbonded potential parameters derived from crystalline hydrocarbons. *The Journal of Chemical Physics*, *47*, 4680.
- Williams, D. E., & Weller, R. R. (1983). Lone-pair electronic effects on the calculated ab initio SCF-MO electric potential and the crystal structures of azabenzenes. *Journal of the American Chemical Society*, *105*, 4143.
- Yan, J. F., Momany, F. A., Hoffmann, R., & Scheraga, H. A. (1970). Energy parameters in polypeptides. II. Semiempirical molecular orbital calculations for model peptides. *The Journal of Physical Chemistry*, *74*, 420.
- Yilmazer, N. D., & Korth, M. (2015). Enhanced semiempirical QM methods for biomolecular interactions. *Computational and Structural Biotechnology Journal*, *13*, 169.
- Yin, Y., Sieradzan, A. K., Liwo, A., He, Y., & Scheraga, H. A. (2015). Physics-based potentials for coarse-grained modeling of protein-DNA interactions. *Journal of Chemical Theory and Computation*, *11*, 1792.
- Yu, W., Lopes, P. E. M., Roux, B., & MacKerell, A. D. (2013). Six-site polarizable model of water based on the classical Drude oscillator. *The Journal of Chemical Physics*, *138*, 034508.
- Zgarbova, M., Otyepka, M., Spomer, J., Mladek, A., Banas, P., Cheatham, T. E., & Jurecka, P. (2011). Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *Journal of Chemical Theory and Computation*, *7*, 2886.
- Zgarbova, M., Luque, F. J., Spomer, J., Cheatham, T. E., Otyepka, M., & Jurecka, P. (2013). Toward improved description of DNA backbone: Revisiting epsilon and zeta torsion force field parameters. *Journal of Chemical Theory and Computation*, *9*, 2339.
- Zhurkin, V. B., Poltev, V. I., & Florentiev, V. L. (1980). Atom-atom potential functions for conformational calculations of nucleic-acids. *Molecular Biology (Moscow), English Translation*, *14*, 882.

The Position of the Clamped Nuclei Electronic Hamiltonian in Quantum Mechanics

3

Brian Sutcliffe and R. Guy Woolley

Contents

Introduction	70
The Clamped Nuclei Approximation	70
The Separation of Translational Motion	74
Choosing Electronic and Nuclear Variables in the Translationally Invariant Hamiltonian	76
Which Is the “Correct” Clamped Nuclei Hamiltonian?	86
The Symmetries of the Clamped Nuclei Electronic Hamiltonian	90
Permutational Symmetry	92
Point Groups and Transformations	97
Spin and Point Group Symmetry	111
The Construction of Approximate Eigenfunctions of the Clamped Nuclei Hamiltonian	113
Conclusions	119
Bibliography	120

Abstract

Arguments are advanced to support the view that at present it is not possible to derive molecular structure from the full quantum mechanical Coulomb Hamiltonian associated with a given molecular formula that is customarily regarded as representing the molecule in terms of its constituent electrons and nuclei. However molecular structure may be identified provided that some additional chemically motivated assumptions that lead to the clamped nuclei Hamiltonian are added to the quantum mechanical account.

B. Sutcliffe (✉) • R.G. Woolley
School of Science and Technology, Nottingham Trent University, Nottingham, UK
Service de Chimie Quantique et Photophysique, Université Libre de Bruxelles,
Bruxelles, Belgium
e-mail: bsutclif@ulb.ac.be

Introduction

The traditional specification of a molecule in classical chemistry is in terms of atoms joined by bonds, and this accounts for the central fact of chemistry that the generic molecular formula is associated with the occurrence of isomers. Such an approach does not provide a useful basis for a physical theory since we do not know the general laws of interaction between atoms. Instead a more abstract description in terms of the particle constituents of a molecule, electrons and nuclei, is employed. We shall confine the discussion to the nonrelativistic level of theory; with this proviso the interactions between electrons and nuclei are assumed to be fully specified by Coulomb's law, and this makes possible the explicit formulation of a molecular Hamiltonian. This so-called Coulomb Hamiltonian will be given explicitly (Eq. 1) in the next section; it forms the starting point of the chapter.

We concentrate on two broad themes. It is obvious that the whole collection of isomers supported by a given molecular formula share the same Coulomb Hamiltonian. The first part of the chapter is concerned with how this fundamental fact has been treated in quantum chemistry through the introduction of the *clamped nuclei Hamiltonian*. This involves two crucial assumptions: (1) the nuclei can be treated as fixed ("clamped") classical particles that merely provide a classical external potential for the electrons and (2) formally identical nuclei can be treated as distinguishable. The second part of the chapter discusses in a general way the basic quantum mechanical theory of the clamped nuclei Hamiltonian, concentrating particularly on its symmetry properties.

The Clamped Nuclei Approximation

The conventional nonrelativistic Hamiltonian for a system of N electrons with position variables, \mathbf{x}_i^e , and a set of A nuclei with position variables \mathbf{x}_i^n may be written as

$$\begin{aligned}
 H(\mathbf{x}^n, \mathbf{x}^e) = & -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla^2(\mathbf{x}_i^e) + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{N'} \frac{1}{|\mathbf{x}_i^e - \mathbf{x}_j^e|} - \frac{e^2}{4\pi\epsilon_0} \sum_{i=1}^A \sum_{j=1}^N \frac{Z_i}{|\mathbf{x}_j^e - \mathbf{x}_i^n|} \\
 & - \frac{\hbar^2}{2} \sum_{k=1}^A \frac{\nabla^2(\mathbf{x}_k^n)}{m_k} + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{A'} \frac{Z_i Z_j}{|\mathbf{x}_i^n - \mathbf{x}_j^n|}.
 \end{aligned} \tag{1}$$

This is the *Coulomb Hamiltonian* for the electrons and nuclei specified by a given molecular formula. We use a Schrödinger representation in which the operators are simple time-independent multiplicative operators acting on functions of the coordinate variables ("wavefunctions"). Kato (1951) established that the Coulomb Hamiltonian, H , is essentially self-adjoint.¹ This property, which is stronger than Hermiticity, guarantees that the time evolution

¹The work was completed in 1944 and was actually received by the journal in October 1948.

$$\Psi(t) = \exp\left(\frac{-i\mathbf{H}t}{\hbar}\right)\Psi(0)$$

of a Schrödinger wavefunction is unitary, and so conserves probability. Furthermore the eigenvalues of \mathbf{H} are associated with a complete set of eigenfunctions. This is not necessarily true for operators that are Hermitian but not self-adjoint.² It was pretty obvious to applied mathematicians that the kinetic energy operator alone is indeed self-adjoint because of their classical mechanical experience. It was shown by Stone in the 1930s that multiplicative operators of the kind specified above are also self-adjoint but it was entirely unobvious that the sum of the operators would be self-adjoint because the sum of the operators is defined only on the intersection of their domains.

What Kato showed was that for a range of potentials including Coulomb ones, and for any function f in the domain \mathcal{D}_0 of the full kinetic energy operator \mathbf{T}_0 , the domain of the full problem \mathcal{D}_V contains \mathcal{D}_0 and there are two constants a and b such that

$$\|\mathbf{V}f\| \leq a\|\mathbf{T}_0f\| + b\|f\|,$$

where a can be taken as small as is liked. This result is often summarized by saying that the Coulomb potential is small compared to the kinetic energy. Given this result he then proved that the usual operator is indeed, for all practical purposes, self-adjoint and is bounded from below. Why worry about this? Well if the operator is not self-adjoint it could support solutions interpretable as a particle falling into a singularity or getting to infinity in a finite time and these are unacceptable as physical solutions. Such pathologies occur in, for example, the classical mechanics of three bodies in a Coulomb field. The practical significance of Kato's proof is the guarantee that such unphysical solutions will not arise from solving the quantum mechanical eigenvalue problem for the Coulomb Hamiltonian.

It is easily established that the Coulomb Hamiltonian is invariant under the coordinate transformations that correspond to uniform translations, rotation-reflections, and permutations of particles with identical masses and charges. Because of the symmetry of the Coulomb Hamiltonian its eigenfunctions will be basis functions for irreducible representations (irreps) of the translation group in three dimensions, the orthogonal group in three dimensions, and for the various symmetric groups corresponding to the sets of identical particles.

Quantum mechanical molecular structure calculations are most commonly attempted by first clamping the nuclei at fixed positions and then performing electronic structure calculations treating the clamped nuclei as providing a potential field for the electronic motion. With the nuclei clamped at a particular fixed

²An elementary example is afforded by the momentum operator $\widehat{p} = -i\hbar d/dq$, which is Hermitian on an appropriately defined class of L^2 functions $\phi(q)$; for these functions it is self-adjoint on $-\infty \leq q \leq +\infty$ but this property is lost if either of the ∞ limits is replaced by any finite value a – see, for example, Thirring (1981).

geometry specified by the constant vectors $\mathbf{x}_i^n = \mathbf{a}_i$, $i = 1, 2, \dots, A$, this modified Hamiltonian takes the form

$$\begin{aligned} H^{\text{cn}}(\mathbf{a}, \mathbf{x}^e) = & -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla^2(\mathbf{x}_i^e) - \frac{e^2}{4\pi\epsilon_0} \sum_{i=1}^A \sum_{j=1}^N \frac{Z_i}{|\mathbf{x}_j^e - \mathbf{a}_i|} + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{N'} \frac{1}{|\mathbf{x}_i^e - \mathbf{x}_j^e|} \\ & + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{A'} \frac{Z_i Z_j}{|\mathbf{a}_i - \mathbf{a}_j|}. \end{aligned} \quad (2)$$

It is customary to incorporate the nuclear repulsion energy into Eq. 2; the nuclear repulsion term is merely an additive constant and so does not affect the form of the electronic wavefunctions. Its inclusion modifies the spectrum of the clamped nuclei Hamiltonian only trivially by changing the origin of the energy. The eigenvalue equation for the clamped nuclei Hamiltonian is then

$$H^{\text{cn}}(\mathbf{a}, \mathbf{x}^e) \psi_p^{\text{cn}}(\mathbf{a}, \mathbf{x}^e) = E_p^{\text{cn}}(\mathbf{a}) \psi_p^{\text{cn}}(\mathbf{a}, \mathbf{x}^e), \quad (3)$$

in which the eigenvalues (“electronic energies”) have a parametric dependence on the constant nuclear position vectors $\{\mathbf{a}_i\}$.

It is sometimes asserted that the clamped nuclei Hamiltonian can be obtained from the Coulomb Hamiltonian by letting the nuclear masses increase without limit. The Hamiltonian that would result if this were done would be

$$\begin{aligned} H^{\text{nn}}(\mathbf{x}^n, \mathbf{x}^e) = & -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla^2(\mathbf{x}_i^e) - \frac{e^2}{4\pi\epsilon_0} \sum_{i=1}^A \sum_{j=1}^N \frac{Z_i}{|\mathbf{x}_j^e - \mathbf{x}_i^n|} + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{N'} \frac{1}{|\mathbf{x}_i^e - \mathbf{x}_j^e|} \\ & + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{A'} \frac{Z_i Z_j}{|\mathbf{x}_i^n - \mathbf{x}_j^n|}, \end{aligned} \quad (4)$$

with the formal Schrödinger equation, by analogy with Eq. 3,

$$H^{\text{nn}}(\mathbf{x}^n, \mathbf{x}^e) \psi_p^{\text{nn}}(\mathbf{x}^n, \mathbf{x}^e) = E_p^{\text{nn}}(\mathbf{x}^n) \psi_p^{\text{nn}}(\mathbf{x}^n, \mathbf{x}^e). \quad (5)$$

Given that the Coulomb Hamiltonian has eigenstates such that

$$H(\mathbf{x}^n, \mathbf{x}^e) \psi(\mathbf{x}^n, \mathbf{x}^e) = E \psi(\mathbf{x}^n, \mathbf{x}^e), \quad (6)$$

if the solutions of Eq. 5 were well defined, it would seem that the eigenstates in Eq. 6 could be expanded as a sum of products of the form

$$\psi(\mathbf{x}^n, \mathbf{x}^e) = \sum_p \Phi_p(\mathbf{x}^n) \psi_p^{\text{nn}}(\mathbf{x}^n, \mathbf{x}^e), \quad (7)$$

where the $\{\Phi\}$ play the role of “nuclear wavefunctions.”

In the Hamiltonian (Eq. 4), the nuclear variables are free and not constant and there are no nuclear kinetic energy operators to dominate the potential operators involving these free nuclear variables. The Hamiltonian thus specified cannot be self-adjoint in the Kato sense. The Hamiltonian can be made self-adjoint by clamping the nuclei because the electronic kinetic energy operators can dominate the potential operators which involve only electronic variables. The Hamiltonian (Eq. 2) is thus a proper one and the solutions Eq. 3 are a complete set. But since the Hamiltonian (Eq. 4) is not self-adjoint it is not at all clear that the hoped for eigensolutions of Eq. 5 form a complete set suitable for the expansion (Eq. 7). However that may be, it was observed more than 30 years ago (Woolley and Sutcliffe 1977) that the arguments for an expansion (Eq. 7) are quite formal because the Coulomb Hamiltonian has a completely continuous spectrum arising from the possibility of uniform translational motion and so its solutions cannot be properly approximated by a sum of this kind. This means too that the arguments of Born and Oppenheimer (1927), and of Born and Huang (1955) for their later approach to representations of this kind, are also quite formal.

As a basis for the Born–Oppenheimer and the Born approach, it is commonly assumed that it is possible to construct an analytic potential function $V(\mathbf{x}^n)$ such that

$$E_p^{\text{cn}}(\mathbf{a}) = V(\mathbf{a}), \text{ for some } p \text{ and for all } \mathbf{a} \quad (8)$$

and that this potential forms an adequate starting point for a discussion of the nuclear motion part of the full problem. Examination of the form of Eq. 2 makes it clear, however, that $E_p^{\text{cn}}(\mathbf{a})$ takes the same value for all choices of \mathbf{a} that differs from a given choice merely by a uniform translation. Similarly it remains unchanged if the \mathbf{a} differ only by a constant orthogonal transformation. Thus any potential formed according to Eq. 8 will have some variables under any change of which no change in the potential will be described. In the context of calculations of molecular spectra, these variables are often referred to as *redundant* ones. It is also the case that $E_p^{\text{cn}}(\mathbf{a})$ is invariant under the permutation of any nuclei with the same charge (nuclear mass does not enter into Eq. 3). This means that the potential in Eq. 8 will have the same value for all geometries that can be obtained from a given geometry by means of a permutation of nuclei with the same charge. Should the potential have any minima at all, it always has as many as there are permutations of the nuclei with the same charge. This would seem to make the assumption of a single isolated minimum in the potential, which is essential to the usual account of the Born–Oppenheimer approximation, a rather too restrictive one for comfort, except perhaps in the case of the diatomic system.

It is thus not at all clear to precisely which question the clamped nuclei Hamiltonian provides the answer and a further discussion of the properties of the Coulomb Hamiltonian is required before the clamped nuclei problem can be put into an appropriate form for yielding a potential. There are two main ways in which such a discussion can be attempted. If it is desired to stay with the Coulomb Hamiltonian in its laboratory-fixed form then the solutions must be expressed in coherent state

(wave-packet) form to allow for their continuum nature. If the solutions are required as L^2 -normalizable wavefunctions, then the translational motion must be separated from the Coulomb Hamiltonian and the solutions of the remaining translationally invariant part must be sought. It is in this second approach that it is easiest to make contact with the standard arguments and this will be considered in the following section.

The Separation of Translational Motion

All that is needed to remove the center-of-mass motion from the molecular Coulomb Hamiltonian is a linear point transformation symbolized by

$$(\mathbf{t}\xi) = \mathbf{x}\mathbf{V}. \quad (9)$$

In Eq. 9, \mathbf{t} is a $3 \times N_T - 1$ matrix ($N_T = N + A$) and ξ is a 3×1 matrix, so that the combined (bracketed) matrix on the left of Eq. 9 is $3 \times N_T$. \mathbf{V} is an $N_T \times N_T$ matrix which, from the structure of the left side of Eq. 9, has a special last column whose elements are

$$V_{iN_T} = M_T^{-1}m_i, \quad M_T = \sum_{i=1}^{N_T} m_i. \quad (10)$$

Hence ξ is the standard center-of-mass coordinate

$$\xi = M_T^{-1} \sum_{i=1}^{N_T} m_i \mathbf{x}_i. \quad (11)$$

As the coordinates \mathbf{t}_j , $j = 1, 2, \dots, N_T - 1$ are to be translationally invariant we require the condition,

$$\sum_{i=1}^{N_T} V_{ij} = 0, \quad j = 1, 2, \dots, N_T - 1 \quad (12)$$

on each remaining column of \mathbf{V} and it is easy to see that Eq. 12 forces $\mathbf{t}_j \rightarrow \mathbf{t}_j$ as $\mathbf{x}_i \rightarrow \mathbf{x}_i + \mathbf{a}$, all i .

The \mathbf{t}_i are independent if the inverse transformation

$$\mathbf{x} = (\mathbf{t}\xi) \mathbf{V}^{-1} \quad (13)$$

exists. The structure of the right side of Eq. 13 shows that the bottom row of \mathbf{V}^{-1} is special and, without loss of generality, its elements may be required to be

$$(\mathbf{V}^{-1})_{N_T i} = 1, \quad i = 1, 2, \dots, N_T. \quad (14)$$

The inverse requirement on the remainder of \mathbf{V}^{-1} implies that

$$\sum_{i=1}^{N_T} (\mathbf{V}^{-1})_{ji} m_i = 0, \quad j = 1, 2, \dots, N_T - 1. \quad (15)$$

The Coulomb Hamiltonian (Eq. 1) in the new coordinates becomes

$$\begin{aligned} \mathbf{H}(\mathbf{t}, \xi) &= -\frac{\hbar^2}{2} \sum_{i=1}^{N_T-1} \frac{1}{\mu_{ii}} \nabla^2(\mathbf{t}_i) - \frac{\hbar^2}{2} \sum_{i,j=1}^{N_T-1'} \frac{1}{\mu_{ij}} \vec{\nabla}(\mathbf{t}_i) \cdot \vec{\nabla}(\mathbf{t}_j) + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{N_T'} \frac{Z_i Z_j}{r_{ij}(\mathbf{t})} \\ &\quad - \frac{\hbar^2}{2M_T} \nabla^2(\xi) \\ &= \mathbf{H}'(\mathbf{t}) - \frac{\hbar^2}{2M_T} \nabla^2(\xi). \end{aligned} \quad (16)$$

Here the positive constants $1/\mu_{ij}$ are given by:

$$\frac{1}{\mu_{ij}} = \sum_{k=1}^{N_T} m_k^{-1} V_{ki} V_{kj}, \quad i, j = 1, 2, \dots, N_T - 1. \quad (17)$$

The operator r_{ij} is the interparticle distance operator expressed as a function of the \mathbf{t}_i . Thus

$$r_{ij}(t) = \left(\sum_{\alpha} \left(\sum_{k=1}^{N_T-1} \left((\mathbf{V}^{-1})_{kj} - (\mathbf{V}^{-1})_{ki} \right) t_{\alpha k} \right)^2 \right)^{1/2}. \quad (18)$$

In Eq. 16, the $\vec{\nabla}(\mathbf{t}_i)$ are gradient operators expressed in the Cartesian components of the \mathbf{t}_i and the last term represents the center-of-mass kinetic energy operator. Since the center-of-mass coordinate does not enter the potential energy term, the center-of-mass motion may be separated off completely so that the eigenfunctions of \mathbf{H} are of the form

$$T(\xi) \Psi(\mathbf{t}), \quad (19)$$

where $\Psi(\mathbf{t})$ is a wavefunction for the Hamiltonian $\mathbf{H}'(\mathbf{t})$, Eq. 16, which will be referred to as the *translationally invariant* Hamiltonian. The eigenfunctions of this Hamiltonian will be basis functions for irreps of the orthogonal group in three dimensions and for the various symmetric groups of the sets of identical particles.

It should be emphasized that different choices of \mathbf{V} are unitarily equivalent so that the spectrum of the translationally invariant Hamiltonian is independent of the

particular form chosen for \mathbf{V} , provided that it is consistent with Eqs. 10 and 12. In particular it is perfectly possible to put the kinetic energy operator into diagonal form by choosing an orthogonal matrix \mathbf{U} that diagonalizes the positive definite symmetric matrix of dimension $N_T - 1$ formed from the $1/\mu_{ij}$ and then replacing elements of the originally chosen \mathbf{V} according to

$$V_{ij} \rightarrow \sum_{k=1}^{N_T-1} V_{ik} U_{kj}, \quad j = 1, 2, \dots, N_T-1.$$

As can be seen from Eq. 18, the practical problem with any choice of \mathbf{V} is the complicated form given to the potential operator.

Choosing Electronic and Nuclear Variables in the Translationally Invariant Hamiltonian

In order to identify the electrons, let the translationally invariant electronic coordinates be chosen with respect to the center-of-nuclear mass

$$\mathbf{t}_i^e = \mathbf{x}_i^e - \mathbf{X}, \quad \mathbf{X} = M^{-1} \sum_{i=1}^A m_i \mathbf{x}_i^n, \quad M = \sum_{i=1}^A m_i.$$

In the case of the atom $A = 1$ and the origin is the nucleus. Other coordinate choices are possible, but this is the only choice that avoids a term in the kinetic energy operator coupling the electronic and nuclear variables and which allows the electronic part of the potential to be written in terms of the electronic variables and the clamped nuclei positions (see Mohallem and Tostes 2002; Sutcliffe 2000).

There is no need to specify the proposed $A - 1$ translationally invariant nuclear variables \mathbf{t}^n other than to say that they are expressed entirely in terms of the laboratory nuclear coordinates by means of a matrix \mathbf{V}^n exactly like \mathbf{V} in Eq. 9, but with side A and with M in place of M_T and \mathbf{X} in place of ξ . It is also sometimes useful to define a set of redundant Cartesian coordinates

$$\bar{\mathbf{x}}_i^n = \mathbf{x}_i^n - \mathbf{X}, \quad i = 1, 2, \dots, A, \quad \text{so that} \quad \sum_{i=1}^A m_i \bar{\mathbf{x}}_i^n = 0. \quad (20)$$

Of course the laboratory nuclear variable \mathbf{x}_i^n cannot be completely written in terms of the $A - 1$ translationally invariant coordinates arising from the nuclei, but in the electron-nucleus attraction and in the nuclear repulsion terms the center-of-nuclear mass \mathbf{X} cancels out. For ease of writing \mathbf{x}_i^n will continue to be used in those terms but it should be remembered that the nuclear potentials are functions of the translationally invariant coordinates defined by the nuclear coordinates.

On making this choice of electronic coordinates the electronic part of Eq. 16 is

$$\begin{aligned}
H'(\mathbf{x}^n, \mathbf{t}^e) = & -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla^2(\mathbf{t}_i^e) - \frac{\hbar^2}{2M} \sum_{i,j=1}^N \vec{\nabla}(\mathbf{t}_i^e) \cdot \vec{\nabla}(\mathbf{t}_j^e) - \frac{e^2}{4\pi\epsilon_0} \sum_{i=1}^A \sum_{j=1}^N \frac{Z_i}{|\mathbf{t}_j^e - \mathbf{x}_i^n|} \\
& + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{N'} \frac{1}{|\mathbf{t}_i^e - \mathbf{t}_j^e|} + \sum_{i,j=1}^{A'} \frac{Z_i Z_j}{|\mathbf{x}_i^n - \mathbf{x}_j^n|}.
\end{aligned} \tag{21}$$

This electronic Hamiltonian is translationally invariant and would yield the usual form were the nuclear masses to increase without limit. It has been noted (Kutzelnigg 2007) that to take (Eq. 21) as the electronic Hamiltonian is inconsistent with a consideration of the solution to the full problem being expressed in a series in terms of powers of the inverse total nuclear mass, since this Hamiltonian already contains a term involving the inverse of this mass to the first power. There is, however, no need to consider this term at the first stage of development of a solution to the full problem and it can be included at the point where terms of similar magnitude are considered. The remaining part of Eq. 21 is then exactly the same as the clamped nuclei form. The clamped nuclei form can be deployed consistently in an account of solutions to the full problem only if a uniform translational factor is included in the full solution. In the work of Nakai et al. (2005) (see also Sutcliffe 2005) the translational motion of the center-of-mass is subtracted to yield a Schrödinger eigenvalue problem from which the translational part of any continuous spectrum has been removed. Of course the spectrum of the resulting operator can have a continuous spectral range, as can the translationally invariant form itself, for reasons quite other than translational motion.

The nuclear part of Eq. 16 involves only kinetic energy operators and has the form:

$$K^n(\mathbf{t}^n) = -\frac{\hbar^2}{2} \sum_{i,j=1}^{A-1} \frac{1}{\mu_{ij}^n} \vec{\nabla}(\mathbf{t}_i^n) \cdot \vec{\nabla}(\mathbf{t}_j^n), \tag{22}$$

with the inverse mass matrix defined as a special case of Eq. 17 involving only the original nuclear variables.

Both (Eqs. 21 and 22) are invariant under any orthogonal transformation of both the electronic and nuclear variables. If the nuclei are clamped in Eq. 21 then invariance remains only under those orthogonal transformations of the electronic variables that can be reexpressed as changes in the positions of nuclei with identical charges while maintaining the same nuclear geometry. The form (Eq. 21) remains invariant under all permutations of the electronic variables and is invariant under permutations of the variables of those nuclei with the same charge. Thus if an electronic energy minimum is found at some clamped nuclei geometry there will be as many minima as there are permutations of identically charged nuclei. The kinetic energy operator (Eq. 22) is invariant under all orthogonal transformations of the nuclear variables and under all permutations of the variables of nuclei with the same mass.

The splitting of the translationally invariant Hamiltonian $H'(\mathbf{t})$ into two parts breaks its symmetry, since each part exhibits only a sub-symmetry of the full problem. If wavefunctions derived from approximate solutions to Eq. 21 are to be used to construct solutions to the full problem (Eq. 6) utilizing (Eq. 22) care will be needed to couple the sub-symmetries to yield solutions with full symmetry.

Atoms

For the atom there is no nuclear kinetic energy part and, denoting the nuclear mass by m_n , the full Hamiltonian is simply the electronic Hamiltonian

$$\begin{aligned} H'^e(\mathbf{t}^e) = & -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla^2(\mathbf{t}_i^e) - \frac{\hbar^2}{2m_n} \sum_{i,j=1}^N \vec{\nabla}(\mathbf{t}_i^e) \cdot \vec{\nabla}(\mathbf{t}_j^e) - \frac{e^2}{4\pi\epsilon_0} \sum_{j=1}^N \frac{Z_i}{|\mathbf{t}_j^e|} \\ & + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{N'} \frac{1}{|\mathbf{t}_i^e - \mathbf{t}_j^e|}. \end{aligned} \quad (23)$$

The electronic problem for the atom (Eq. 23) has exactly the same form and symmetry as the full problem and meets the requirements for Kato self-adjointness, for there is a kinetic energy operator in all of the variables that are used to specify the potential terms. This would continue to be the case were the nuclear mass to increase without limit.

The atom is sometimes used as an illustration when considering the original form of the Born–Oppenheimer approximation (as in Deshpande and Mahanty 1969), but the only aspect of the approximation that can be thus illustrated is the translational motion part and that is easily considered in first order by treating the second term in Eq. 23 as a perturbation to the solution obtained using an infinite nuclear mass. The inclusion of this term in this way is analogous to making the usual *diagonal Born–Oppenheimer* correction and it can be made exactly in the case of any one-electron atom (see Handy and Lee 1996). As noted in Hinze et al. (1998) it is usually made approximately simply by including the diagonal part of the mass polarization term (the second term in Eq. 23 above) to produce an electronic reduced mass

$$\frac{1}{\mu_e} = \frac{1}{m_n} + \frac{1}{m}$$

in place of $1/m$.

The Hamiltonian (Eq. 23) maintains full symmetry and is invariant under electronic permutations and under rotation-reflections of the electronic coordinates. Trial functions are usually constructed from atomic orbitals and from their spin-orbitals. Permutational antisymmetry is achieved by forming Slater determinants from the spin-orbitals. Rotational symmetry is usually realized by vector coupling

of orbitals that form bases for representations of the rotation group $SO(3)$. Spin-eigenfunctions too are achieved by vector coupling.³

Molecules

Even after separating the translational motion, for a molecule there is always at least one nuclear variable in the kinetic energy part of the operator, and self-adjointness cannot be achieved if such terms are neglected while the potential terms involving the nuclear variables are included except by clamping the nuclei. The treatment of molecules is, thus, technically much more difficult than is the treatment of atoms.

Although the discussion that follows is, for the most part, quite general, explicit consideration is confined to the diatomic case in order to avoid overburdening the exposition with details. However here too, there are certain technical features which simplify the diatomic case and which cannot be transferred to the polyatomic case so care will be taken in the following discussion not to make the diatomic the general case. For a system with two nuclei the natural nuclear coordinate is the internuclear vector which will be denoted here simply as \mathbf{t} . When needed to express the electron-nuclei attraction terms, \mathbf{x}_i^n is simply of the form $\alpha_i \mathbf{t}$ where α_i is a signed ratio of the nuclear mass to the total nuclear mass. In the case of a homonuclear system $\alpha_i = \pm \frac{1}{2}$. The di-nuclear electronic Hamiltonian is

$$\begin{aligned} H^{e}(\mathbf{t}^e) = & -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla^2(\mathbf{t}_i^e) - \frac{\hbar^2}{2(m_1+m_2)} \sum_{i,j=1}^N \vec{\nabla}(\mathbf{t}_i^e) \cdot \vec{\nabla}(\mathbf{t}_j^e) \\ & - \frac{e^2}{4\pi\epsilon_0} \sum_{j=1}^N \left(\frac{Z_1}{|\mathbf{t}_j^e + \alpha_1 \mathbf{t}|} + \frac{Z_2}{|\mathbf{t}_j^e + \alpha_2 \mathbf{t}|} \right) \\ & + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{N'} \frac{1}{|\mathbf{t}_i^e - \mathbf{t}_j^e|} + \frac{Z_1 Z_2}{R}, \quad R = |\mathbf{t}|, \end{aligned} \quad (24)$$

while the nuclear kinetic energy part is:

$$-\frac{\hbar^2}{2} \left(\frac{1}{m_1} + \frac{1}{m_2} \right) \nabla^2(\mathbf{t}) \equiv -\frac{\hbar^2}{2\mu} \nabla^2(\mathbf{t}). \quad (25)$$

The electronic part is not self-adjoint in the manner prescribed by Kato because it contains no kinetic energy terms involving the nuclear variable which would dominate the potential energy terms. The full Hamiltonian would not be Kato self-adjoint if both nuclear masses were to increase without limit either. It is seen from

³Some specifics of the implementation of permutational and rotational symmetry in quantum mechanics are discussed in section "The Symmetries of the Clamped Nuclei Electronic Hamiltonian."

Eq. 25, however, that if only one nuclear mass increases without limit then the kinetic energy term in the nuclear variable remains in the full problem and so the Hamiltonian remains self-adjoint in the Kato sense.

The di-nuclear case has been considered numerically by Frolov (1999) in a study of the hydrogen molecular ion. In extremely accurate calculations on the discrete states of this system, he investigated what happened when first one and then two nuclear masses are increased without limit. He showed that when one mass increased without limit, any discrete spectrum persisted but when two masses were allowed to increase without limit, the Hamiltonian ceased to be well-defined and this failure led to what he called *adiabatic divergence* in attempts to compute discrete eigenstates. This sort of behavior would certainly be anticipated from the present discussion.

Irrespective of any choices made for the nuclear masses, the electronic Hamiltonian (Eq. 24) becomes self-adjoint in the Kato sense if the nuclei are clamped for then the nuclear variables in the potential terms become constants and the only variables are the electronic ones. So the clamped nuclei potential is dominated by the electronic kinetic energy. Thus the usual practice of clamped nuclei electronic structure calculations is a consistent one.

Writing the variable \mathbf{t} in spherical polar coordinates, R , β , and α where $t_z = R \cos \beta$, were the clamped nuclei Hamiltonian to be used to define a potential it is easily seen that for $\mathbf{t} = \mathbf{a}$, $R = a$ then

$$E^{\text{cn}}(\mathbf{a}) = V(a), \quad (26)$$

so that the potential would have the form $V(R)$. But the potential is not just a curve, it is a series of spherical shells of rotation swept out by the curve by all choices of β and α . It is thus a genuine central-field potential. If the internuclear distance is fixed but \mathbf{a} allowed to rotate or invert, then $E^{\text{cn}}(\mathbf{a})$ is a sphere of constant energy as swept out by the variables β and α at radius a . If \mathbf{a} is oriented so as to define a z -axis then $E^{\text{cn}}(\mathbf{a})$ will take the same value at $+a_z$ and $-a_z$ so that if there is a minimum at $+a_z$ there will be another at $-a_z$. The electronic Hamiltonian is not invariant under inversion of the nuclear variables alone unless the two nuclei have identical charges in which case inversion and permutation will have identical effects. In differential geometry terms, the potential is homeomorphic to S^2 .

The Hamiltonian (Eq. 24) is invariant under all rotations of the electronic coordinates about the internuclear axis and all reflections in a plane containing the internuclear axis. The electronic states can be labeled by a quantum number m which takes the values $0, \pm 1, \pm 2, \dots$ corresponding to the eigenvalues of the z -component of the electronic angular momentum about the internuclear axis.

It is easily seen that the potential will tend to increase without limit as $R \rightarrow 0$ but the behavior as $R \rightarrow \infty$ presents a problem. To see this consider the asymptotic behavior of the electron-nucleus potential terms in the case of the one-electron homonuclear di-hydrogen molecule ion. The electronic coordinate is

$$\mathbf{t}^e = \mathbf{x} - \frac{1}{2}(\mathbf{x}_1^n + \mathbf{x}_2^n), \quad (27)$$

where \mathbf{x} is the laboratory coordinate of the electron. As the internuclear distance becomes very large, the nuclear repulsion term becomes very small and one would expect the trial wavefunction to approach the wavefunction for a one-electron ion corresponding to one of the atoms. Thus one might expect the lowest energy wavefunction to be of the form

$$Ne^{-cr}, \quad \mathbf{r} = \mathbf{x} - \mathbf{x}_1^n, \quad r = |\mathbf{r}|,$$

for instance. However working in the chosen coordinate set

$$\mathbf{r} = \mathbf{t}^e - \frac{1}{2}\mathbf{t},$$

so that the expected asymptotic electronic solution could be expressed only in terms of both the electronic and nuclear variables. This does not, of course, mean that the potential cannot approach the required value. It simply means that it cannot do so in any calculation in which the trial functions are confined to electronic functions whose variable origin is at the center-of-nuclear mass.

This sort of difficulty is a general one and obviously not confined simply to one-electron diatomic molecules. It would clearly be unwise to attempt to approximate solutions for molecules at energies close to their dissociation limits in terms of electronic coordinates with the origin at the center-of-nuclear mass. A trial function for the general case of the Born–Huang form

$$\psi(\mathbf{t}^n, \mathbf{t}^e) = \sum_p \Phi_p(\mathbf{t}^n) \psi_p^{\text{nn}}(\mathbf{t}^n, \mathbf{t}^e), \quad (28)$$

where the \mathbf{t}^e have an origin at the center-of-nuclear mass, could, therefore, approximate only a limited region of the spectrum of the full problem.

This difficulty cannot be got round by working in the laboratory frame. The solution to the full problem would be defined in terms of a three-dimensional subspace expressed in terms of a translation variable and a $3(N_T - 1)$ -dimensional subspace expressible in terms of translationally invariant variables. Translationally invariant variables must involve at least a pair of variables and so there must be at least one such variable which involves a laboratory frame electron and a laboratory frame nuclear variable. All this can be easily illustrated by considering the exact ground-state wavefunction of the hydrogen atom, as is seen in Kutzelnigg (2007).

This point is developed in more detail by Hunter (1981) in a paper considering to what extent a separation of electronic and nuclear motion would be possible if the exact solution to the Schrödinger equation for the Coulomb Hamiltonian was actually known. Were the exact solution known Hunter argues (1975) that it could be written in the form

$$\psi(\mathbf{t}^n, \mathbf{t}^e) = \chi(\mathbf{t}^n) \phi(\mathbf{t}^n, \mathbf{t}^e), \quad (29)$$

with a nuclear wavefunction defined by means of

$$|\chi(\mathbf{t}^n)|^2 = \int \psi(\mathbf{t}^n, \mathbf{t}^e)^* \psi(\mathbf{t}^n, \mathbf{t}^e) d\mathbf{t}^e.$$

Then, providing this function has no nodes,⁴ an “exact” electronic wavefunction could be constructed as

$$\phi(\mathbf{t}^n, \mathbf{t}^e) = \frac{\psi(\mathbf{t}^n, \mathbf{t}^e)}{\chi(\mathbf{t}^n)}, \quad (30)$$

if the normalization choice

$$\int \phi(\mathbf{t}^n, \mathbf{t}^e)^* \phi(\mathbf{t}^n, \mathbf{t}^e) d\mathbf{t}^e = 1$$

is made. In fact it is possible (Hunter 1981) to show that χ must be nodeless even though the usual approximate nuclear wavefunctions for vibrationally excited states do have nodes. The electronic wavefunction (Eq. 30) is, therefore, properly defined and a potential energy surface could be defined in terms of it as

$$U(\mathbf{t}^n) = \int \phi(\mathbf{t}^n, \mathbf{t}^e)^* H'(\mathbf{t}^n, \mathbf{t}^e) \phi(\mathbf{t}^n, \mathbf{t}^e) d\mathbf{t}^e. \quad (31)$$

Although no exact solutions to the full problem are known for a molecule, some extremely good approximate solutions for excited vibrational states of H_2 have been computed and Czub and Wolniewicz (1978) took such an accurate approximation for an excited vibrational state in the $J = 0$ rotational state of H_2 and computed $U(R)$. They found strong spikes in the potential at close to two positions at which the usual wavefunction would have nodes. To quote Czub and Wolniewicz (1978):

This destroys completely the concept of a single internuclear potential in diatomic molecules because it is not possible to introduce on the basis of non-adiabatic potentials a single, approximate, mean potential that would describe well more than one vibrational level.

It is obvious that in the case of rotations the situation is even more complex.

Bright Wilson suggested (1979) that using the clamped nuclei Hamiltonian instead of the full one in Eq. 31 to define the potential might avoid the spikes but Hunter (1981) showed why this was unlikely to be the case and Cassam-Chenai (2006) repeated the work of Czub and Wolniewicz using an electronic Hamiltonian and showed that exactly the same spiky behavior occurred. However Cassam-Chenai showed, as Hunter had anticipated, that if one simply ignored the spikes, the potential was almost exactly the same as would be obtained by deploying the electronic Hamiltonian in the usual way.

⁴A similar requirement must be placed on the denominator in Eq. 12 of Kutzelnigg (2007) for the equation to provide a secure definition.

Although the spiky nature of an “exact” potential has been demonstrated explicitly only for $J = 0$ states of a small diatomic molecule, there is no reason to suppose that its occurrence is not general. Matters would be further complicated by rotational motion. Thus the demonstrably smooth potentials generated by solving an electronic problem cannot be approximations to any exact form but are simply computationally useful intermediates in a solution to the full problem. It would therefore seem unwise to assign too much weight to them in explaining chemical structure.

In the standard approach to solving the nuclear motion part of the diatomic problem, the potential $V(R)$ is specified and the nuclear motion Hamiltonian becomes

$$-\frac{\hbar^2}{2\mu}\nabla^2(\mathbf{t}) + V(R). \quad (32)$$

Expressing this Hamiltonian in spherical polar coordinates one obtains the usual form

$$-\frac{\hbar^2}{2\mu}\left(\frac{1}{R^2}\frac{\partial}{\partial R}R^2\frac{\partial}{\partial R}\right) + \frac{1}{2\mu R^2}L^2 + V(R), \quad (33)$$

where L is the operator for the angular momentum of the nuclear motion. The angular part of the solution is known analytically and the solution of the nuclear motion problem can be reduced to one in the single variable R .

The eigensolutions to this problem are quite naturally eigenfunctions of the nuclear angular momentum and can easily be chosen with the required permutational symmetry. But things are not quite so clear for the electronic part of the problem because one does not in practice have a form which is explicit in the nuclear variables as it is computed only at fixed nuclear geometries. It is easy to achieve the correct permutational symmetry for the electronic part of the function at each and every nuclear geometry, but it would be not at all easy to make each function an eigenfunction of the electronic angular momentum.

To try to deal with the rotational motion it is possible to reformulate the diatomic problem to exhibit explicitly the angular symmetry of the Hamiltonian. As shown in Kołos and Wolniewicz (1963) and, in a somewhat more general way in Sutcliffe (2007), it is possible to define an internal coordinate system by a transformation that makes the internuclear vector \mathbf{t} the z -axis in a right-handed coordinate system and in this system the electronic Hamiltonian (Eq. 24) becomes

$$-\frac{\hbar^2}{2m}\sum_{i=1}^N\nabla^2(\mathbf{r}_i) - \frac{\hbar^2}{2(m_1 + m_2)}\sum_{i,j=1}^N\vec{\nabla}(\mathbf{r}_i)\cdot\vec{\nabla}(\mathbf{r}_j) + V(R, \mathbf{r}), \quad (34)$$

with

$$V(R, \mathbf{r}) = \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{N'} \frac{1}{r_{ij}} + \frac{e^2}{4\pi\epsilon_0} \frac{Z_1 Z_2}{R} - \frac{e^2}{4\pi\epsilon_0} \sum_{j=1}^N \left(\frac{Z_1}{r_{j1}(R)} + \frac{Z_2}{r_{j2}(R)} \right), \quad (35)$$

where the \mathbf{r}_i are the electronic variables expressed in the transformed system. In this formulation the vector \mathbf{t} orients freely and ‘‘clamping the nuclei’’ comes down to simply choosing $R = a$. A clamped nuclei solution of Eq. 34 would lead to a clamped nuclei energy

$$E^{\text{cn}}(a) = V(a), \quad (36)$$

rather than the form given by Eq. 26. Thus any minima in $V(a)$ would not be duplicated by the requirement of rotational invariance. Inversion is achieved by means of

$$\beta \rightarrow \pi - \beta, \quad \alpha \rightarrow \pi + \alpha$$

and thus involves just the angular part of the formulation. An identical operation achieves the nuclear permutation. The electronic Hamiltonian is invariant under neither operation unless the nuclei have identical charge.

The nuclear kinetic energy operator (Eq. 25) becomes

$$-\frac{\hbar^2}{2\mu R^2} \frac{\partial}{\partial R} R^2 \frac{\partial}{\partial R} + \frac{1}{2\mu R^2} \mathbf{D}_1(\alpha, \beta, \mathbf{r}), \quad (37)$$

with

$$\mathbf{D}_1(\alpha, \beta, \mathbf{r}) = \left[(\mathbf{J}_x - \mathbf{l}_x)^2 + (\mathbf{J}_y - \mathbf{l}_y)^2 + \frac{\hbar}{i} \cot \beta (\mathbf{J}_y - \mathbf{l}_y) \right],$$

where the electronic angular momentum is

$$\mathbf{l} = \sum_{i=1}^N \mathbf{l}(i) = \frac{\hbar}{i} \sum_{i=1}^N \mathbf{r}_i \times \frac{\partial}{\partial \mathbf{r}_i}, \quad (38)$$

and the total angular momentum operator is denoted \mathbf{J} and involves both the electronic and nuclear variables in such a way that $\mathbf{J}_z = \mathbf{l}_z$. The Jacobian for this transformation is

$$R^2 \sin \beta.$$

It is seen that in this formulation any solution of the clamped nuclei form of the electronic Hamiltonian (Eq. 34) will give rise to a potential which is simply a curve and not a surface of rotation. However the angular part of the nuclear kinetic energy

operator now involves the electronic angular momentum so that the electronic motion and the overall rotational motion are coupled.

The electronic wavefunction in this case has the same axial symmetry as in the previous case and is characterized by the quantum number m . However the states can no longer be regarded as occurring in degenerate pairs for $m > 0$ since l_z is the z -component of the total angular momentum so m can take integer values lying between J and $-J$, where J is the total angular momentum. The coupling of the electronic and nuclear angular momenta lifts the m degeneracy as the Hamiltonian becomes a system of $2J + 1$ coupled partial differential equations.

In the first of the two possible ways of looking at the diatomic, one remains in the Cartesian product space $R^3 \times R^{3N}$ and it is thus necessary to give some explicit consideration to the angular properties of solutions to the electronic part of the problem. If the usual approach were taken to approximating solutions to the nuclear motion Hamiltonian using sums of products of electronic and nuclear parts, a typical term in the sum used as trial function for the form (Eq. 33) would be

$$\phi_{pm}(\mathbf{r}^e, R)^L \Phi_{pm}(R) \Theta_{Lm}(\beta, \alpha), \quad (39)$$

where p denotes an electronic state and L the nuclear angular momentum quantum number. There are no operators in the nuclear motion part of the problem which explicitly couple the electronic and nuclear motions. It is thus possible to represent for any electronic state, any number of rotational states specified by values of L , without considering any coupling. L in Eq. 24 is not the total angular momentum operator and so a description of rotational motion given in these terms yields only an approximate quantum number.

If one transforms to the manifold $R_+ \times S^2 \times R^{3N}$ then one can consider explicitly the rotational coupling of electronic and angular motions. The fact that the transformation is to a manifold rather than a vector space means that any operator built using coordinates defined on the manifold will be well defined only where the Jacobian for the transformation does not vanish. This does not cause great problems here because the only places where the Jacobian vanishes are when $R = 0$ and where $\beta = 0$ and $\beta = \pi$. The region around $R = 0$ is inaccessible because of the nuclear repulsion term and the exact angular wavefunctions take care of the problem with β .

It would not be consistent to use only a single term in a product approximation for a trial for the form (Eq. 37) with a potential $V(R)$ except in considering a $J = 0$ state. Here the minimum consistent product approximation is

$$\sum_{m=-J}^J \phi_{pm}(\mathbf{r}, R)^J \Phi_{pm}(R) \Theta_{Jm}(\beta, \alpha). \quad (40)$$

It is only for $J = 0$ states that the forms (Eqs. 39 and 40) are the same.

Of course the angular momentum coupling in Eq. 40 implies a coupling of different electronic states, between Σ and Π states or Σ and Δ states, for example. To allow explicitly for that possibility (Eq. 40) should really be extended to

$$\sum_{m=-J}^J \sum_{pm}^J \phi_{pm}(\mathbf{r}, R)^J \Phi_{pm}(R) \Theta_{Jm}(\beta, \alpha), \quad (41)$$

where the electronic state is denoted pm to indicate that the state must have quantum number m . Thus for $J = 2$ one would need at least five electronic states.

Which Is the “Correct” Clamped Nuclei Hamiltonian?

There is clearly a choice between the form (Eqs. 24 and 34) and although in the clamped nuclei approximation both would yield the same energies for any chosen internuclear separation a , the resulting energy would be a potential for two quite different situations.

To generalize from the diatomic case, if the usual approach were taken to approximating solutions to the nuclear motion Hamiltonian using sums of products of electronic and nuclear parts a typical term in the sum used as trial function for the form (Eq. 33) would be

$$\phi_p(\mathbf{t}^e, \mathbf{t}^n) \Phi_p(\mathbf{t}^n), \quad (42)$$

where p denotes an electronic state. The solutions are on the Cartesian product space $R^{3A-3} \times R^{3N}$. There is again no explicit coupling of the nuclear motion and electronic motions and it is thus possible to represent for any electronic state, any number of rotational states. It is only in the diatomic case that the nuclear angular momentum can be realized explicitly as part of the nuclear kinetic energy so it is not generally possible to choose Φ directly as an eigenfunction of the nuclear angular momentum, neither is it possible to choose ϕ directly as an eigenfunction of the electronic angular momentum. ϕ as usually computed belongs to the totally symmetric representation of the symmetric group of each set of nuclei with identical charges. Φ could then be a basis function for an irreducible representation of the symmetric group for each set of particles with identical masses if the permutational symmetry were properly considered in solving the nuclear motion problem.

Clamped nuclei calculations are usually undertaken so as to yield a potential that involves no redundant coordinates. Thus a translationally invariant electronic Hamiltonian like (Eq. 24) would actually generate a more general potential than this. A clamped nuclei potential is, therefore, more properly associated with the electronic Hamiltonian after the separation of rotational motion like (Eq. 34) than with the merely translationally invariant one. With this choice again, the minimum consistent product approximation is

$$\sum_{m=-J}^J \phi_{pm}(\mathbf{r}, \mathbf{R})^J \Phi_{pm}(\mathbf{R}) |JMm\rangle, \quad (43)$$

where \mathbf{R} represents the $3A - 6$ internal coordinates invariant under all orthogonal transformations of the \mathbf{t}^n and $|J M m\rangle$ is an angular momentum eigenfunction. The general solutions are on the manifold $R^{3A-6} \times S^3 \times R^{3N}$ though for triatomic molecules the internal coordinate part of the manifold is confined to $R_+ \times R^2$ because the three nuclear positions define a plane. It is only in the diatomic case that the electronic variables play a direct part in the specification of the angular momentum eigenfunctions and where there is therefore only one internal coordinate. However the Coriolis coupling terms in the angular part of the Hamiltonian contain terms in the electronic angular momentum so coupling of the electronic motion to the angular motion would still be anticipated, see sections V and VI of Sutcliffe (2000). Although the angular momentum coupling could imply a coupling of different electronic states, as it certainly does in a diatomic, it is not obviously implied in the general case. To achieve permutational symmetry in the nuclear motion part of the wavefunction would in the general case be very tricky. The nuclei are identified in the process of defining a body-fixed frame to describe the rotational motion, even if they are identical. If only a subset of a set of identical nuclei were used in such a definition, some permutation of the nuclear variables would induce a change in the definition of the body-fixed frame and thus spoil the rotational separation. Thus permutations of identical nuclei are considered usually only if such permutations correspond to point group operations which leave the body-fixing choices invariant.

If one considers the clamped nuclei Hamiltonian as providing input for the full Hamiltonian in which the rotational motion is made explicit, the basic nuclear motion problem should be treated as a $2J + 1$ dimensional problem. If this is done then the translational and rotational symmetries of the full problem are properly dealt with. However the solutions are not generally basis functions for irreps of the symmetric groups of sets of identical nuclei except for such subgroups as constitute the point groups used in frame fixing. This restriction of the permutations is usually assumed justified by appealing to the properties of the potential surface. The idea here was introduced by Longuet-Higgins (1963) and is widely used in interpreting molecular spectra.

As noted earlier, the original attempts to justify the Born–Oppenheimer and the Born approaches from the full Coulomb Hamiltonian lack rigorous mathematical foundations. So far there have been no attempts to make the foundations of the Born approach mathematically secure. However the coherent states approach has been used to give mathematically rigorous accounts of surface crossings and a review of this work can be found in Hagedorn and Joye (2007). It seems very unlikely that it would be possible to provide a secure foundation for the Born approach in anything like the manner in which it is usually presented.

The Born–Oppenheimer approximation, whose validity depends on there being a deep enough localized potential well in the electronic energy, has, however, been extensively treated. The mathematical approaches depend upon the theory of fiber bundles and the electronic Hamiltonian in these approaches is defined in terms of a fiber bundle. It is central to these approaches, however, that the fiber bundle should be trivial, that is that the base manifold and the basis for the fibers be

describable as a direct product of Cartesian spaces. This is obviously possible with the decomposition choice made for Eq. 42 but not obviously so in the choice made for Eq. 43.

The Born–Oppenheimer approach has been put on a rigorous foundation for diatomics with solutions of the form (Eq. 39) in work which is described in a helpful context in Combes and Seiler (1980). For solutions like (Eq. 40), it is possible that more than one vector (coordinate) space can be constructed on it because the transformation is to a manifold. In fact two coordinate spaces are possible on S^2 a trivial one and a twisted one, the latter associated with the possibility of an electronic wavefunction with a Berry phase and the “twisted” solutions are accounted for in Herrin and Howland (1997).

A mathematically satisfactory account of the Born–Oppenheimer approximation for polyatomics in an approach based on Eq. 43 has not yet been provided but it has proved possible to provide one based on Eq. 42 (see Klein et al. 1992). Because the nuclear kinetic energy operator in the space R^{3A-3} cannot be expressed in terms of the nuclear angular momentum, it is not possible in this formulation to separate the rotational motion from the other internal motions. This work also considers the possibility that there are two minima in the potential as indeed there would be because of inversion symmetry if the potential minimum were at other than a planar geometry. It does not, however, consider the possibility of such multiple minima as might be induced by permutational symmetry. It might be possible to extend the two minima arguments to the multiple minima case and perhaps provide a mathematically secure account of the Longuet-Higgins approach to ignoring some of the inconvenient permutations. This has not so far been attempted.

For a secure account to be given in terms of the separation (Eq. 43), which is what is really required if one is to use the clamped nuclei electronic Hamiltonian, it would be necessary to consider more than one coordinate space. On the manifold S^3 at least two coordinate spaces are required to span the whole manifold. The internal coordinates within any coordinate space are such that it is possible to construct two distinct molecular geometries at the same internal coordinate specification, so that a potential expressed in the internal coordinates cannot be analytic everywhere (Collins and Parsons 1993). It would therefore seem to be a very tricky job. But even if it were to be accomplished it seems very unlikely that a multiple minima argument could be constructed to account for point group symmetry in this context. It is possible to show (see section IV of Sutcliffe (2000)) that in the usual Eckart form of the Hamiltonian for nuclear motion, permutations can be such as to cause the body-fixed frame definition to fail completely.

If it is wished to perform a clamped nuclei calculation on a molecule containing three or more nuclei, avoiding translations and rigid rotations, it is necessary to fix the values of six of the $3A$ nuclear variables. In practice this is usually done by choosing one nucleus, x_1^n , at the origin, one nucleus, x_2^n , defining an axis and a third, x_3^n , defining a plane. Every possible geometry of the molecule, except where the three particles become collinear, can be specified with this choice but not every component of the $3A$ variables will appear in the clamped nuclei

electronic energy as six of them have been chosen to be zero. This means that even though the clamped nuclei electronic energy can be specified in terms of a molecular geometry in which the positions of A points can be given, the energy itself is a function of only the relative positions of a subset of the nuclei. Thus performing clamped nuclei calculations will not make possible the expression of the electronic energy in anything other than internal coordinates and the electronic energy when expressed in any set of internal coordinates cannot be analytic everywhere. Thus there cannot really be a “global” potential energy surface. In any case, as has been seen, the potential, even locally, cannot be regarded as an approximation to anything in particular and thus should be treated simply as a convenient peg on which to hang further calculations. From this perspective, the further calculations should properly be ones in which the electronic Hamiltonian results from the full Hamiltonian in which the rotational motion has been made explicit. Such Hamiltonians have only a local validity and can be defined only where the Jacobian for the transformation to the rotational variables does not vanish. However at present there is no satisfactory account of how nuclear permutational symmetry should be treated from this perspective, neither is there any secure mathematical justification of the Born–Oppenheimer approximation or of the Born approach.

Naturally any extension of the trial wavefunction for the full Coulomb Hamiltonian problem from a single term to a many term form must be welcomed as an advance; it is, however, simply a technical advance and it might prove premature to load that technical advance with too much physical import.

At present it is not possible to place properly the clamped nuclei electronic Hamiltonian in the context of the full problem, including nuclear motion. However if the nuclei were treated as distinguishable particles, even when formally identical, then some of difficulties that arise from the consideration of nuclear permutations would not occur. But it would still be necessary to be able to justify the choice of subsets of permutations among identical particles when such seem to be required to explain experimental results. A particular difficulty arises here for it is not possible to distinguish between isomers nor is it possible to specify a molecular geometry, unless it is possible to distinguish between formally identical particles.

But regarding the nuclei as distinguishable would not avoid the difficulty of constructing total angular momentum eigenfunctions from the nuclear and electronic parts. Such treatment of the nuclei would not make the traditional demonstrations of the Born–Oppenheimer or the Born approximations mathematically sound either. However it would ensure that the mathematically sound presentations of the Born–Oppenheimer approximation mentioned earlier need no further extension to include permutations of identical nuclei. There is, unfortunately, little good to be said, from a mathematical point of view, of the traditional Born argument. This is troubling because the Born approach is assumed to provide the basis for the consideration of chemical reactions on and between potential energy surfaces. However it is clear that the clamped nuclei electronic Hamiltonian can be usefully deployed in nuclear motion calculations if the nuclei are considered identifiable and in the next sections this Hamiltonian will be considered in detail.

The Symmetries of the Clamped Nuclei Electronic Hamiltonian

Before considering the symmetry under permutations of identical particles it is necessary first to say a little about the spin of particles. Each particle is specified not only by space variables but also by spin variables. These have not been considered so far because there are no spin operators in the Hamiltonians discussed in the previous sections. Nevertheless spin is, indirectly, very important in the construction of approximate wavefunctions.

Spin ideas are usually developed in terms of a vector operator S with three components S_α , $\alpha = x, y, z$ and functions $\Theta_{S, M_s, k}^N$ such that

$$S^2 \Theta_{S, M_s, k}^N = \hbar^2 S(S+1) \Theta_{S, M_s, k}^N, \quad S^2 = S_x^2 + S_y^2 + S_z^2,$$

and

$$S_z \Theta_{S, M_s, k}^N = \hbar^2 M \Theta_{S, M_s, k}^N,$$

with,

$$S = 0, 1, 2, \dots, \quad -S \leq M_s \leq S. \quad (44)$$

The $\Theta_{S, M_s, k}^N$ are called spin eigenfunctions and the k index denotes a particular member of a possible set. In the simple case of a single electron, $S = \frac{1}{2}$, $M_s = \pm \frac{1}{2}$, and $k = 1$ so usually $\Theta_{\frac{1}{2}, \frac{1}{2}, 1}^1$ is written as α and $\Theta_{\frac{1}{2}, -\frac{1}{2}, 1}^1$ as β .

It can be shown that under the operations of the full rotation-reflection group in three dimensions, O(3):

$$(S_x S_y S_z) \rightarrow (S_x S_y S_z) |\mathbf{R}| \mathbf{R}, \quad (45)$$

where \mathbf{R} is the matrix representation of the rotation inversion and is thus an orthogonal matrix with determinant $|\mathbf{R}|$ which is $+1$ for a rotation and -1 for a reflection. Clearly S^2 is an invariant operator. Thus, its eigenfunctions provide a basis for irreducible representations of O(3) and in general we label these representations by the S value to which they correspond, using k to distinguish between different basis functions for the same representation. Obviously the dimension of these is $2S + 1$ since the $2S + 1$ functions with $M_s = S, S - 1, \dots, -S$ are degenerate in the sense of having the same S eigenvalue.

It is useful to treat the spin functions as having variables s_i though these can take only discrete values and from now on to label the Cartesian space variables as \mathbf{r}_i so that \mathbf{x}_i can be used to denote the totality of variables. To construct a many-particle function to describe space and spin for a collection of N identical particles it is apparently simply necessary to form the products:

$$\Psi_n(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \Psi_n(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) \Theta_{S, M_s, k}^N(s_1, s_2, \dots, s_N) \quad (46)$$

at will, and the generalization to groups of sets of identical particles is obvious. However there is a snag. It turns out that there is an extra rule that must be obeyed. The Hamiltonian is invariant under the operators that permute the variable designations (space and spin) of sets of identical particles. For any one set, these operators form a group and the full invariance group of H is the direct product of the groups for all the sets of identical particles. This means that the eigenfunctions of H are a basis for the irreps of this group. The extra rule is related to what irreps can actually occur and it is usually called the *Pauli principle*. If the set of identical particles individually have spin $S = 0, 1, 2, 3$, etc., then the only irreps that can arise are the totally symmetric ones. Thus every wavefunction must be invariant under permutations that interchange the variable designations of particles with integer spins (*bosons*). If the particles individually have spin $S = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}$ etc., then the only irreps are the antisymmetric ones. Thus every wavefunction must change sign if the permutation is odd or be invariant if the permutation is even for permutations that interchange the variable designations of particles with odd half-integer spin (*fermions*).

What this means is that not every function of the form (Eq. 46) actually corresponds to a physical state, but it can be shown that in any given problem there are enough eigenfunctions of S^2 with any given M_s value (i.e., k is sufficiently large) to form a basis for irreps of the permutation group of N particles. Similarly it can be shown that there are enough space functions to provide irreps of the same group. It is then possible to derive rules for combinations of the space and spin functions such that the resulting function obeys the Pauli principle.

Since in this and the following sections only the clamped nuclei Hamiltonian will be considered, it will be convenient to simplify the notation somewhat and to drop the superscripts on the electronic variables because, from now on, these will be the only variables to be considered. Thus Eq. 2 is rewritten as

$$\begin{aligned} H(\mathbf{a}, \mathbf{r}) = & -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla^2(\mathbf{r}_i) - \frac{e^2}{4\pi\epsilon_0} \sum_{j=1}^A \sum_{i=1}^N \frac{Z_j}{|\mathbf{r}_i - \mathbf{a}_j|} + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{N'} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\ & + \frac{e^2}{8\pi\epsilon_0} \sum_{i,j=1}^{A'} \frac{Z_i Z_j}{|\mathbf{a}_i - \mathbf{a}_j|}, \end{aligned} \quad (47)$$

while Eq. 3 is rewritten as

$$H(\mathbf{a}, \mathbf{r}) \psi_p(\mathbf{a}, \mathbf{r}) = E_p(\mathbf{a}) \psi_p(\mathbf{a}, \mathbf{r}). \quad (48)$$

For later purposes it is useful to group the first two terms by writing

$$h(i) = -\frac{\hbar^2}{2m} \nabla^2(\mathbf{r}_i) - \frac{e^2}{4\pi\epsilon_0} \sum_{j=1}^A \frac{Z_j}{|\mathbf{r}_i - \mathbf{a}_j|} \equiv k(i) + V(i), \quad (49)$$

and to express the third term with

$$g(i, j) = \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (50)$$

The electronic part of the Hamiltonian for any given set of nuclear positions may then be written as

$$\mathbf{H} = \sum_{i=1}^N \mathbf{h}(i) + \frac{1}{2} \sum_{i,j=1}^{N'} g(i, j). \quad (51)$$

The operator \mathbf{h} is self-adjoint and a mathematically simple extension of the hydrogen atom operator. It therefore has some discrete one-particle eigenfunctions

$$\mathbf{h}(i)\phi(\mathbf{r}_i) = \epsilon\phi(\mathbf{r}_i). \quad (52)$$

Such one-particle functions are called *orbitals* and the associated energies ϵ are *orbital energies*. If there is only one nucleus, they are called *atomic orbitals* (AOs) and in the many nuclei case, *molecular orbitals* (MOs). The orbitals may be extended to become *spin-orbitals* by multiplying them with an α or β spin function and the resulting spin-orbitals are therefore denoted as $\phi(\mathbf{x}_i)$ to indicate the inclusion of spin. Clearly every orbital can generate two spin-orbitals.

It is usual in the case of electrons to work with functions that obey the Pauli principle from the start. This is possible when using an orbital basis through the use of Slater determinants of spin-orbitals which, by definition, change sign under odd permutations but are invariant under even permutations as is required for fermions. These will be explicitly considered later; first the permutational symmetry will be considered more generally.

Permutational Symmetry

The clamped nuclei Hamiltonian is invariant under the permutation of electrons so that its eigenfunctions must be basis functions for irreducible representations (irreps) of the symmetric group of degree N . However, because of the Pauli principle, not all irreducible representations of this group are allowed.

As noted above for electrons there are only two kinds of spin functions so that in any product of spin functions for N spin variables, only a limited set of permutations of the spin variables will lead to distinct product functions. Since there are no spin terms in the electronic Hamiltonian, the operators for total spin commute with the Hamiltonian and so its eigenfunctions must be spin eigenfunctions as well as basis functions for an irrep of the symmetric group. To see how these two requirements can be combined consider permutations rather generally.

The abstract permutation operator on N objects $1, 2, 3, \dots, N$ is always denoted by

$$P = \begin{pmatrix} 1 & 2 & 3 & \dots & N \\ i_1 & i_2 & i_3 & \dots & i_N \end{pmatrix}, \quad (53)$$

where the i_r are the set of objects $1, 2, 3, \dots, N$ at most in some new order. The effect of the operator on the original set of objects is

$$P(1, 2, 3, \dots, N) = (i_1 i_2 \dots i_N). \quad (54)$$

Clearly the product of any two permutations is a permutation, the products are associative, and there is an identity permutation, and every permutation possesses a unique inverse so the permutations form a group as required. The order of the group is obviously $N!$, because there are $N!$ distinct permutations of N objects. This is the symmetric group of degree N , usually denoted as S_N .

A permutation like

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \quad \text{is often written as } (1\ 2\ 3), \quad (55)$$

on the interpretation that 2 replaces 1, 3 replaces 2, and 1 replaces 3. The right-hand expression in Eq. 55 is said to be in a *cycle* form and is a cycle of *length* 3. Any permutation can be decomposed uniquely into cycles of disjoint elements, thus for example:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 4 & 5 & 1 & 3 & 6 \end{pmatrix} = (1\ 2\ 4)(3\ 5)(6), \quad (56)$$

where the order in which the individual cycles are written down does not matter and where, of course, each cycle is invariant under cyclic permutations of its elements. A cycle of length two is called a *transposition* and is an involutory operation, that is:

$$(i_1 i_2) = (i_1 i_2)^{-1}. \quad (57)$$

A cycle can always be written as a product of transpositions, thus

$$(1\ 2\ 3) = (1\ 2)(2\ 3), \quad (58)$$

where the transposition on the right operates first, but the decomposition is obviously not unique nor are the transpositions always of disjoint elements. However the *number* of transpositions into which a cycle can be decomposed is obviously unique and this number is said to be the *parity* of the permutation, *even* if the number is even and *odd* if the number is odd. It follows, therefore, that we can classify the permutation operators themselves as either even or odd. Parity is usually indicated as $(-1)^p$ where p is the number of transpositions. The product of two

even permutations is an even permutation so that the even permutations constitute a subgroup (counting the identity permutation as even) of the symmetric group of order $N!$. This subgroup is called the *alternating* group. It is obvious then that the set of all distinct transpositions are the generators of the symmetric group.

It is reasonably easy to show that all permutations in the same class have the same cycle structure.⁵ Thus if the possible cycle structures are known, then the number of classes and hence the number of irreducible representations is known. But the number of cycle structures is obviously the number of possible ways that there is of dividing N objects in integer partitions. Thus for four objects the possible cycle structures are

$$\begin{aligned}
 & (i_1 i_2 i_3 i_4), \\
 & (i_1 i_2 i_3) (i_4), \\
 & (i_1 i_2) (i_3 i_4), \\
 & (i_1 i_2) (i_3) (i_4), \\
 & (i_1) (i_2) (i_3) (i_4).
 \end{aligned} \tag{59}$$

If there are m cycles in any partition, the cycle lengths are usually written symbolically as λ_i , where

$$\lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_m = N, \tag{60}$$

and by convention

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_m \geq 0. \tag{61}$$

Thus in Eq. 59 above in the first decomposition there is only one cycle and $\lambda_1 = 4$, in the second there are only two cycles $\lambda_1 = 3$ and $\lambda_2 = 1$. It is a custom to denote the totality of λ for any partition by the vector $[\lambda] \equiv 4[\lambda_1, \lambda_2 \dots \lambda_m]$ and to indicate r repeats of a cycle of given length λ by λ^r . Thus in Eq. 59 above the partitions should be denoted as $[4]$, $[3\ 1]$, $[2^2]$, $[2\ 1^2]$, and $[1^4]$.

The notation (Eq. 59) is not especially well adapted for present purposes and instead a notation invented by Young, the so-called Young diagram, is used. The Young diagrams equivalent to Eq. 59 are

⁵This means that the permutation and its inverse are always in the same class. A group with this property is said to be an *ambivalent* group.

$$\begin{array}{ccccc}
 \begin{array}{|c|c|c|c|} \hline \square & \square & \square & \square \\ \hline \end{array} &
 \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & & \\ \hline \end{array} &
 \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} &
 \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \\ \hline \square & \\ \hline \end{array} &
 \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array} \\
 [4] & [3\ 1] & [2^2] & [2\ 1^2] & [1^4]
 \end{array} \tag{62}$$

where the boxes in each row stand for a cycle and different cycles form different rows in the diagram.

Since there are as many irreducible representations of a group as it has classes and there is one Young diagram for each class, Young diagrams can obviously be used to label each of the different irreducible representations. The dimensionality of the irreducible representation must pretty clearly be associated with the number of ways it is possible to put the objects 1, 2,... N into the boxes in the Young diagram, in a unique way. In fact it turns out that if the numbers 1, 2,... N are put into the boxes in a diagram so that when they are read across any row they are increasing and when they are read down any column they are also increasing, the totality of all such assignments is the dimension of the irreducible representation.

Thus the first diagram in Eq. 62 can only yield

$$\begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 4 \\ \hline \end{array} \tag{63}$$

A Young diagram with numbers in it like (Eq. 63) is called a *Young tableau* and it is seen at once that there is only one tableau for the last Young diagram in Eq. 62, namely:

$$\begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline 3 \\ \hline 4 \\ \hline \end{array} \tag{64}$$

so that both [4] and [1⁴] correspond to one-dimensional irreducible representations.

The representation [3 1] yields

$$\begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & & \\ \hline \end{array} \quad
 \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline 3 & & \\ \hline \end{array} \quad
 \begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & & \\ \hline \end{array} \tag{65}$$

and in these circumstances it is clearly nice to have some sort of standard order; the one often adopted is to read the tableau across each row as if it were an ordinary number and then to arrange the tableaux in ascending order of these numbers, but other orders are used. Thus the numbers in Eq. 65 are

$$1342 \quad 1243 \quad 1234 \tag{66}$$

so that the standard order for the tableaux would be

$$\begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & & \\ \hline \end{array}
 \quad
 \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline 3 & & \\ \hline \end{array}
 \quad
 \begin{array}{|c|c|c|} \hline 1 & 3 & 4 \\ \hline 2 & & \\ \hline \end{array}
 \tag{67}$$

This representation is therefore three-dimensional.

One can obviously get the tableaux for $[2 \ 1^2]$ simply by transposing the tableaux for $[3 \ 1]$ and a little reflection will show that this is generally the case, simply because of the structure of the Young diagrams. The representations $[2 \ 1^2]$ and $[3 \ 1]$ are said to be *conjugate* representations (or sometimes *adjoint*, or *dual* or *associated*). In the present case the representation $[2^2]$ is self-conjugate. Generally if the representation is denoted by $[\lambda]$ then the conjugate representation is denoted by $[\tilde{\lambda}]$, and of course both representations are of the same dimension.

Since there is obviously freedom in which is associated with which, the choice is made in the two one-dimensional cases that $[N]$ is associated with the symmetric representation and $[1^N]$ with the antisymmetric representation. Thus if one has a representation $[\lambda]$, then the conjugate representation $[\tilde{\lambda}]$ is just $[\lambda] \times [1^N]$.

The $[\lambda]$ symbols are also used to denote classes from time to time and in that context the correspondences established between Eqs. 62 and 59 are appropriate. Thus $[1^4]$ represents the last partition in Eq. 59 which is clearly the class that just contains the unit operator, $[2 \ 1^2]$ stands for the class containing all operators involving only one transposition and so on.

Now it can be shown that if one can find a basis for a representation $[\lambda]$ for the symmetric group and another basis which is a representation for $[\tilde{\lambda}]$ then the direct product of these two bases is a basis for the antisymmetric representation $[1^N]$. So to construct antisymmetric wavefunctions from space-spin products, as in constructing electronic wavefunctions, one can do so as

$$\Psi^{[1^N]} = \sum_i \Phi_i^{[\tilde{\lambda}]}(\mathbf{r}) \Theta_{S, M_S, i}^{[\lambda]}(s), \tag{68}$$

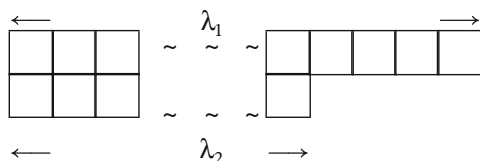
where the sum goes over all functions in the basis for the irrep and N identical particles are assumed involved. They can have any spins at all and although in the expression above they are assumed to be spin eigenfunctions this is obviously not necessary just as long as a suitable set of spin products is used. However if they are spin- $\frac{1}{2}$ particles then the only possible representations that can be constructed in the spin space are of the form $[\lambda_1, \lambda_2]$. No other irreps are possible. Furthermore in the case of N electrons properly coupled to a spin state S then

$$\lambda_1 + \lambda_2 = N, \quad \lambda_1 - \lambda_2 = 2S, \tag{69}$$

so that

$$\lambda_1 = \frac{N}{2} + \mathcal{S}, \quad \lambda_2 = \frac{N}{2} - \mathcal{S}, \quad (70)$$

where



Thus any allowed space wave functions, such as those formed from orbital products, must form a basis for $[\tilde{\lambda}]$.

This result arises basically because there are only *two* kinds of spin functions for spin- $\frac{1}{2}$ particles. Thus lots of permutations that *could* arise do not arise in practice because they are the same if they involve only, say, α spin particles. However it can also be shown that if the spin basis consists of spin eigenfunctions then the reps provided of \mathcal{S}_N are actually irreps and that they are the same irreps for any value of M_S . That is, there are *exactly* as many spin eigenfunctions with the same \mathcal{S} value as the dimension of the irrep $[\lambda_1, \lambda_2]$ of \mathcal{S}_N . This is an enormous blessing and makes life much easier in electronic structure calculations than it might otherwise be. For fermions of spin greater than $\frac{1}{2}$ it is not the case and it makes, for example, the problem of calculating statistical weights in molecular spectroscopy, or nuclear spin states in nuclear physics, much more complicated.

Point Groups and Transformations

In the Cartesian coordinate system, the Laplacian operator for the i th electron is written as

$$\nabla^2(i) = \frac{\partial^2}{\partial r_{xi}^2} + \frac{\partial^2}{\partial r_{yi}^2} + \frac{\partial^2}{\partial r_{zi}^2}, \quad (71)$$

and the inter-electron distance as

$$r_{ij} = \left(\sum_{\alpha} (r_{\alpha j} - r_{\alpha i})^2 \right)^{1/2}, \quad (72)$$

where here and hereafter sums over α run over x , y , and z . The electron-nucleus distance will be written as

$$r_{ai} = \left(\sum_{\alpha} (r_{ai} - a_{\alpha})^2 \right)^{1/2}, \quad (73)$$

where the α th coordinate of the nucleus a is written as a_{α} to emphasize that it is a pure number and not a variable, since the nuclear positions are assumed to be fixed.

It is easy to show that

$$r_{ij}^2 = (\mathbf{r}_j - \mathbf{r}_i)^T (\mathbf{r}_j - \mathbf{r}_i), \quad (74)$$

and

$$r_{ai}^2 = (\mathbf{r}_i - \mathbf{a})^T (\mathbf{r}_i - \mathbf{a}), \quad (75)$$

where the superscript T denotes the matrix transpose.

Now let it be supposed that \mathbf{r} is subject to positive rotation about the z -axis through an angle ϕ to yield the vector \mathbf{r}' . Positive rotations are in the direction of a right-handed screw-twist when the point of the screw is moving in the axis direction. It is then easy to show that the new and old vectors are related by an orthogonal matrix

$$\mathbf{r}' = \mathbf{R}(\phi) \mathbf{r}, \quad (76)$$

with

$$\mathbf{R}(\phi) = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (77)$$

In this section we shall be concerned only with transformations like (Eq. 77), that is, with *point transformations* and, in the case of the clamped nuclei Hamiltonian, consideration may be confined to *finite point groups*, specifically the crystallographic point groups. However from time to time it will be convenient to think a bit more generally and to recognize, for example, that the full rotation-reflection group in three dimensions is a point group, though an *infinite* one, and that some Hamiltonians are invariant under the operators of this group.

Now it should be understood quite clearly as to what happens in a transformation of variables in a fixed basis. It is supposed that *every* point in space goes over into its *image* according to the appropriate transformation rule

$$\mathbf{r}' = \mathbf{R}\mathbf{r}, \quad (78)$$

where \mathbf{R} is a constant orthogonal matrix.⁶

⁶By "constant" here is meant simply that the elements of the matrix are not themselves dependent on the variables.

Thus imagining the situation appropriate to the Schrödinger equation with Hamiltonian (Eq. 47), the assertion of a transformation of variables realized by \mathbf{R} would imply

$$\mathbf{r}'_i = \mathbf{R}\mathbf{r}_i, \quad i = 1, 2, \dots, N. \quad (79)$$

However, in the fixed nuclei approximation, it would *not* imply that

$$\mathbf{a}' = \mathbf{R}\mathbf{a}, \quad \mathbf{a} = \text{any nucleus.} \quad (80)$$

This is not for any sinister reason but simply because of the definition chosen of the clamped nuclei approximation. In this approximation the nuclear positions are triplets of numbers which take particular but fixed values once a particular basis is chosen.⁷

It is now very easy to show that the Laplacian and the inter-electronic distance are invariant under any constant orthogonal transformation of variables but that in general the electron-nucleus distances are not invariant. It is perhaps sensible to show how this comes about for it illustrates one or two points and makes the idea of *invariance* a bit clearer.

First the Laplacian. Using the chain rule for differentiation

$$\frac{\partial}{\partial r_{\alpha i}} = \sum_{\beta} \frac{\partial r'_{\beta i}}{\partial r_{\alpha i}} \frac{\partial}{\partial r'_{\beta i}}, \quad (81)$$

since the transformation does not “mix” \mathbf{r}_i and \mathbf{r}_j so that

$$\frac{\partial}{\partial r_{\alpha i}} = \sum_{\beta} R_{\beta\alpha} \frac{\partial}{\partial r'_{\beta i}}, \quad (82)$$

and hence

$$\frac{\partial}{\partial r_{\alpha i}^2} = \sum_{\gamma\beta} R_{\gamma\beta} R_{\beta\alpha} \frac{\partial^2}{\partial r'_{\gamma i} \partial r'_{\beta i}}, \quad (83)$$

but

$$\nabla^2(i) = \sum_{\alpha} \frac{\partial^2}{\partial r_{\alpha i}^2},$$

⁷It is sometimes convenient to think of the nuclear positions as *defining* a particular embedding for the basis vectors or coordinate frame.

and

$$\sum_{\alpha} R_{\gamma\alpha} R_{\beta\alpha} = \delta_{\gamma\beta} \quad (\mathbf{R}^T \mathbf{R} = E_3), \quad (84)$$

therefore

$$\nabla^2(i) = \sum_{\gamma\beta} \delta_{\gamma\beta} \frac{\partial^2}{\partial r'_{\gamma i} \partial r'_{\beta i}} = \sum_{\beta} \frac{\partial^2}{\partial r'_{\beta i}{}^2} = \nabla^2(i'). \quad (85)$$

But clearly if

$$\nabla^2(i) f(\mathbf{r}_i) = g(\mathbf{r}_i), \quad (86)$$

then

$$\nabla^2(i') f(\mathbf{r}'_i) = g(\mathbf{r}'_i), \quad (87)$$

so that $\nabla^2(i)$ is an invariant operator under the transformation since its effect on an arbitrary function is unchanged (except for the variable names which, in this context, are irrelevant). Notice that if ∇^2 had been set up in the primed system as $\nabla^2(i')$ the same argument would have shown that $\nabla^2(i') = \nabla^2(i)$, as required.

Next r_{ij} . From Eq. 78 it follows that

$$\mathbf{r}'_i = \mathbf{R} \mathbf{r}_i \quad \text{and} \quad \mathbf{r}_i = \mathbf{R}^T \mathbf{r}'_i. \quad (88)$$

Combining the form (Eqs. 74 with 88) we get

$$\begin{aligned} r_{ij}^2 &= \left(\mathbf{R}^T (\mathbf{r}'_j - \mathbf{r}'_i) \right)^T \left(\mathbf{R}^T (\mathbf{r}'_j - \mathbf{r}'_i) \right) \\ &= (\mathbf{r}'_j - \mathbf{r}'_i)^T \mathbf{R} \mathbf{R}^T (\mathbf{r}'_j - \mathbf{r}'_i) \\ &= (\mathbf{r}'_j - \mathbf{r}'_i)^T (\mathbf{r}'_j - \mathbf{r}'_i), \end{aligned}$$

so that

$$r_{ij}^2 = r'_{ij}{}^2, \quad (89)$$

showing that r_{ij}^2 and hence r_{ij} is invariant in the same sense as above.

For the electron-nucleus distance one gets

$$\begin{aligned} r_{ai}^2 &= (\mathbf{R}^T \mathbf{r}'_i - \mathbf{a})^T (\mathbf{R}^T \mathbf{r}'_i - \mathbf{a}) \\ &= (\mathbf{R}^T (\mathbf{r}'_i - \mathbf{R} \mathbf{a}))^T (\mathbf{R}^T (\mathbf{r}'_i - \mathbf{R} \mathbf{a})), \end{aligned}$$

so that

$$r_{ai}^2 = (\mathbf{r}'_i - \mathbf{R}\mathbf{a})^T (\mathbf{r}'_i - \mathbf{R}\mathbf{a}), \quad (90)$$

and this in general is obviously not $r_{ai}^{\prime 2}$. However, if the transformation \mathbf{R} is such that

$$\mathbf{b} = \mathbf{R}\mathbf{a} \quad \text{and} \quad \mathbf{a} = \mathbf{R}^T \mathbf{b}, \quad (91)$$

where \mathbf{b} is the position of another nucleus in the problem with the same charge as the nucleus whose position is \mathbf{a} , then it is clear that even though individual terms in the electron-nucleus attraction operator are *not* invariant, nevertheless the operator as a whole *is* invariant.

The generalization of this result is obvious. If the nuclear positions are divided into sets corresponding to nuclei of the same charge, the clamped nuclei Hamiltonian is invariant under all operators represented by orthogonal matrices \mathbf{R} such that if

$$\mathbf{b} = \mathbf{R}\mathbf{a}, \quad (92)$$

\mathbf{b} is to be found in the same set as \mathbf{a} was found.⁸

To look at it another way, were the coordinates of a set of equivalent nuclei to be written as a row matrix of the coordinates $\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_p$ (so that the “row” is actually a $3 \times p$ matrix), the effect of an invariance preserving operator (a *symmetry* operator) could be symbolized in the partitioned matrix form

$$(\mathbf{b}_1 \mathbf{b}_2 \dots \mathbf{b}_p) = (\mathbf{a}_1 \mathbf{a}_2 \dots \mathbf{a}_p) \mathbf{P}, \quad (93)$$

where \mathbf{P} is a $p \times p$ permutation matrix, that is, one of the form

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 1 & 0 & \dots & 0 \end{pmatrix}, \quad (94)$$

where the permutation corresponds to the effective rearrangement of the nuclei in the electron-nucleus attraction operator. Permutation matrices have just one nonzero entry in any row or column and that entry is always 1 so that such matrices are always orthogonal matrices. Their determinant can, however, be ± 1 depending on the *parity* of the permutation.

⁸Notice that in this approximation the *mass* of the nucleus is of no consequence, only the charge matters.

Thus, to all intents and purposes, when considering the symmetry of a molecule to be represented by a clamped nuclei Hamiltonian, all we need to look at is those orthogonal operators that let equivalent nuclear positions carry a permutation representation of the operator in question. Any such operator is a symmetry operator of the problem. We thus never need to think about the electrons and it is usual in elementary books to start from the point reached here, as if the operations were carried out on the nuclear positions. From here one can go on to show, in the usual way, that the collection of symmetry operators form a group and that the group is a *point group* essentially corresponding to the geometrical figure formed by the equivalent nuclei in the nuclear framework.

Functions can always be visualized in terms of a fixed Cartesian basis and under the change (assumed orthogonal) $\mathbf{r}' = \mathbf{R}\mathbf{r}$ or $\mathbf{r} = \mathbf{R}^T\mathbf{r}'$ the change *induced* in a function $f(\mathbf{r})$ is

$$\mathbf{O}_R f(\mathbf{r}) = \overline{f}(\mathbf{r}) \equiv f(\mathbf{R}^T\mathbf{r}), \quad (95)$$

meaning that $\overline{f}(\mathbf{r})$ is constructed from $f(\mathbf{r})$ by substituting $(\mathbf{R}^T\mathbf{r})_\alpha$ wherever it occurs in the prescription for $f(\mathbf{r})$. \mathbf{O}_R symbolizes the operation in function space.

Considering the transformation of operators again, assuming that one specifies the operator form in a fixed basis, consider an operator \mathbf{A} (assumed linear and Hermitian) such that

$$\mathbf{A}f(\mathbf{r}) = g(\mathbf{r}), \quad (96)$$

and by formal operator algebra we can rewrite this to exhibit the effect of the operator \mathbf{O}_R as

$$\begin{aligned} \mathbf{O}_R \mathbf{A} f(\mathbf{r}) &= \mathbf{O}_R g(\mathbf{r}), \\ \mathbf{O}_R \mathbf{A} f(\mathbf{r}) &= \mathbf{O}_R \mathbf{A} \mathbf{O}_R^{-1} \mathbf{O}_R f(\mathbf{r}) = \mathbf{O}_R g(\mathbf{r}), \\ \mathbf{O}_R \mathbf{A} f(\mathbf{r}) &= \mathbf{O}_R \mathbf{A} \mathbf{O}_R^{-1} \overline{f}(\mathbf{r}) = \overline{g}(\mathbf{r}), \end{aligned} \quad (97)$$

so that we can say that if R induces \mathbf{O}_R then it induces a change in the operators acting on the original function space such that

$$\overline{\mathbf{A}} = \mathbf{O}_R \mathbf{A} \mathbf{O}_R^{-1}, \quad (98)$$

where the functions on which the operators act are implicit in the formalism.

Furthermore if the induced change is such that $\mathbf{A} = \overline{\mathbf{A}}$, that is, \mathbf{A} is an invariant operator, then

$$\mathbf{A} = \mathbf{O}_R \mathbf{A} \mathbf{O}_R^{-1}, \quad (99)$$

or

$$\mathbf{O}_R \mathbf{A} = \mathbf{A} \mathbf{O}_R,$$

or

$$\mathbf{O}_R \mathbf{A} - \mathbf{A} \mathbf{O}_R \equiv [\mathbf{O}_R \mathbf{A}] = 0,$$

and we say that the \mathbf{O}_R commute with \mathbf{A} . We can thus characterize the symmetry operators in the function space that constitutes the domain of \mathbf{A} as those operators which commute with \mathbf{A} . In particular we are interested in the case when \mathbf{A} is the Hamiltonian for the problem.

Now, in practice, (Eq. 98) is not of much use for actually checking on the transformation properties of operators since it does not actually provide a specific rule for the construction of $\bar{\mathbf{A}}$. However as seen before, the actual variable names in an operator relation do not matter, so that if we regard $\mathbf{R}^T \mathbf{r}$ as just a new variable name, say \mathbf{y} , then it follows immediately from Eqs. 96 and 97 that $\bar{\mathbf{A}}$ is just the operator \mathbf{A} made up in the same basis as \mathbf{A} but with $(\mathbf{R}^T \mathbf{r})_\alpha$ for every occurrence of r_α in the operator. This prescription is usable in an obvious way for multiplicative terms in the operator and, for derivative terms, a little thought shows that the replacement

$$\frac{\partial}{\partial r_\alpha} \rightarrow \sum_\gamma (\mathbf{R}^T)_{\alpha\gamma} \frac{\partial}{\partial r_\gamma} = \sum_\gamma R_{\gamma\alpha} \frac{\partial}{\partial r_\gamma} \quad (100)$$

is the appropriate one for the new operator set up on the original basis. (Compare with Eqs. 81 and 82.)

Thus we see that the momentum operator, for example, is not generally an invariant operator since from (Eq. 100)

$$\begin{aligned} \bar{p}_\alpha &= \frac{\hbar}{i} \sum_\gamma R_{\gamma\alpha} \frac{\partial}{\partial r_\gamma} = \sum_\gamma R_{\gamma\alpha} p_\gamma, \\ \bar{p}_\alpha &= (\mathbf{R}^T \mathbf{p})_\alpha, \end{aligned} \quad (101)$$

and (Eq. 101) is the equation that relates the elements of the transformed operator to those of the untransformed operator on a fixed basis. In fact it is easy to see from (Eq. 101) that in this case the operators actually provide a basis for the representation of R by \mathbf{R} as

$$\bar{\mathbf{p}} = \mathbf{p} \mathbf{R}, \quad (102)$$

where the operators are written as a row matrix. The transformation of operators is a matter to which we shall return later, but for the moment let it simply be noted that the arguments presented earlier in this section show that the clamped nuclei

Hamiltonian H is an invariant operator if R is a point group operation and we shall now look at the behavior of its eigenfunctions under O_R .

Any set of degenerate eigenfunctions of H carry a matrix representation of O_R . Thus if we denote the set of degenerate eigenfunctions as f_1, f_2, \dots, f_n and treat these as a row matrix f , then we can write

$$O_R f = f O_R, \quad (103)$$

and we say that O_R provides a matrix representation of O_R in the basis f_i so that a set of such functions in a function space play exactly the same role as do the unit vectors in a coordinate space and are, therefore, sometimes called *basis functions* of the representation.

The representations provided in the basis of degenerate eigenfunctions are usually irreducible and can be chosen to be unitary matrices (in fact usually orthogonal matrices if the functions are real functions). In practice, what one is usually faced with is a collection of functions which have arbitrary (but known) transformation properties and what one actually wants to do is to adapt these functions so that they actually transform like the true eigenfunctions of the problem. This can be done by means of the group theoretical projection operator.

If the unitary matrix irreducible representation of the point group is known and denoted by $\mathbf{D}^{(\mu)}(R)$, where μ labels the irreducible representation, then the projection operator is

$$P_{ij}^{(\mu)} = \frac{n_\mu}{g} \sum_R \left(\mathbf{D}^{(\mu)}(R) \right)_{ij}^* O_R, \quad (104)$$

where g is the order of the group and n_μ the dimension of $\mathbf{D}^{(\mu)}(R)$.⁹ This projection operator is such that

$$P_{ij}^{(\mu)} f = f_i^{(\mu)} \quad \text{or a null result,} \quad (105)$$

where f is an arbitrary function and $f_i^{(\mu)}$ is the i th function in the row matrix $f^{(\mu)}$ such that

$$O_R f^{(\mu)} = f^{(\mu)} \mathbf{D}^{(\mu)}(R). \quad (106)$$

The null result occurs in Eq. 105 if the arbitrary function has no component that lies in the subspace defined by the $f^{(\mu)}$.

In practice one often makes do with the character projection operator based on the characters $\chi^{(\mu)}$ of the irreducible representation,

⁹The operator in this form is clearly only possible for finite groups but similar operators are constructible for most infinite groups of interest.

$$\chi^{(\mu)}(R) = \sum_i \left(D^{(\mu)}(R) \right)_{ii}^*, \quad (107)$$

so that the character projection operator $\mathbf{P}^{(\mu)}$ is

$$\mathbf{P}^{(\mu)} = \frac{\eta^{(\mu)}}{g} \sum_R \left(\chi^{(\mu)}(R) \right)^* \mathbf{O}_R, \quad (108)$$

and produces from an arbitrary function (if not a null result) then a function which is said to *belong* to the μ th irreducible representation but which in general transforms like a linear combination of the basis functions for $\mathbf{D}^{(\mu)}$.

In practice, therefore, everything depends on knowing the representations of the group in question (and the irreducible representations are usually tabulated) and on knowing how the chosen functions actually transform under the operations of the group.

It is helpful at this stage to think of constructing MOs as a linear combination of atomic orbitals (LCAOs) and as a preliminary to this the idea of an MO and of an AO need to be developed a bit further.

It might be thought that orbitals could be best determined by solving the relevant one-electron problem specified by \mathbf{h} , (Eq. 49), for each molecule. This is not so easily done and if a wavefunction is constructed as a determinant of spin-orbitals so constructed it yields a very poor approximation to the electronic energy. This is because the electron interaction term is large and has not been considered in constructing the orbital. Thus to get useful orbitals the fact that the electrons interact strongly must somehow be incorporated into the single particle potential through some sort of effective field process. This was first attempted for atoms in the 1920s by W. Hartree and D. R. Hartree (1927; 1936) but without considering antisymmetry. Their method was developed by Fock (1930) and by Slater (1930) in a properly antisymmetric form. Molecules resisted this approach because it was extremely difficult to solve the one-electron problem with a potential that was not centrosymmetric. However since it was possible to get atomic orbitals, two approaches using them developed, one at the hands of Pauling (1939) using atomic orbitals directly and the other at the hands of Mulliken (1931) by writing molecular orbitals as a linear combination of atomic orbitals. The Pauling approach was called the Valence Bond (VB) method and the Mulliken one the Linear Combination of Atomic Orbitals Molecular Orbital (LCAO MO) method. Although it was the VB method that was used in the first successful molecular electronic structure calculation, on H_2 by Heitler and London (1927), it was the LCAO MO method that became standard in quantum chemical calculations following the work of Roothaan (1951) in the 1950s. Although these methods will be considered in great detail elsewhere in this volume, their basis can best be considered here.

The most familiar atomic orbitals are those for the hydrogen atom which are usually designated according to their angular parts as s for $l=0$, p for $l=1$, d for $l=2$, and so on where there are $2l+1$ orbitals of each type. For calculational

purposes Slater proposed that the radial part of the orbital, while still centered upon the nucleus, should be of the form

$$Nr^{n-1} \exp -\xi r, \quad n = 1, 2, \dots, \quad (109)$$

where N is a normalizing factor, and ξ may be chosen at will or used as a variational parameter.

It is customary to use a real form for the angular parts of the orbital so that, for example, the three p orbitals are specified as

$$x \exp -\xi r, \quad y \exp -\xi r, \quad z \exp -\xi r. \quad (110)$$

The most widespread atomic orbital form in use at present is the Gaussian which has the radial part

$$Nr^{n-1} \exp -\xi r^2, \quad n = 1, 2, \dots. \quad (111)$$

When it is not necessary to specify atomic orbitals in detail they will be denoted $\eta_i(\mathbf{r})$.

In elementary books it is often assumed that you can “eye-ball” the transformation properties from the “pictures” of the orbitals stuck on the nuclei. (Such pictures usually look a bit like twisted sausages and doughnuts.) Such an approach is a bit risky and knowing how to do this sort of thing algebraically avoids getting into the nasty mess that often happens if one relies on the pictures. Indeed, it is clear that it is quite impossible to draw a picture of an orbital. The sausages and doughnuts, therefore, must be at most projection representations in three (or two) dimensions and it is not clear that the transformation of the picture in three (or two) dimensions really corresponds correctly to the induced transformation of the orbital in the function space.

If the atomic orbital $\eta(\mathbf{r})$ is situated at the origin of coordinates and the origin is the invariant point of the point group operators, then any orbital constructed about this origin simply transforms under \mathbf{O}_R to a new functional form in the space

$$\mathbf{O}_R \eta(\mathbf{r}) = \bar{\eta}(\mathbf{r}) = \eta(\mathbf{R}^T \mathbf{r}). \quad (112)$$

If, however, the orbital is developed about another origin, say \mathbf{a} , so that the functional form is

$$\eta((\mathbf{r} - \mathbf{a})), \quad (113)$$

then the situation is a little more involved. The coordinate point \mathbf{a} is not a variable of the problem, but a parameter in the function construction. It does not change under the transformation so that

$$\begin{aligned} \mathbf{O}_R \eta((\mathbf{r} - \mathbf{a})) &= \eta((\mathbf{R}^T \mathbf{r} - \mathbf{a})) \\ &= \eta(\mathbf{R}^T (\mathbf{r} - \mathbf{Ra})), \end{aligned} \quad (114)$$

but clearly \mathbf{Ra} is some other fixed parameter point in the problem \mathbf{b} , say, so that

$$\begin{aligned} \eta(\mathbf{R}^T (\mathbf{r} - \mathbf{Ra})) &= \eta(\mathbf{R}^T (\mathbf{r} - \mathbf{b})) \\ &= \eta(\mathbf{R}^T \mathbf{r}_b). \end{aligned} \quad (115)$$

Effectively the origin of the function is changed, but about that new origin, the transformed function is constructed by the usual rules in terms of the variable $\mathbf{r}_b = (\mathbf{r} - \mathbf{b})$. Clearly if $\eta(\mathbf{R}^T \mathbf{r})$ is $\bar{\eta}(\mathbf{r})$ then $\eta(\mathbf{R}^T \mathbf{r}_b)$ will be $\bar{\eta}(\mathbf{r}_b)$ so that if one has worked out the transformation properties of an orbital at the origin ($a = 0$) then the functional form of the transformed orbital carries over immediately to an orbital at arbitrary point \mathbf{a} given that $\mathbf{b} = \mathbf{Ra}$ is known.

Thus if one chooses a set of AOs which are identical but on different centers then they carry a permutation representation of the operation, much like (Eq. 94) but with appropriate transformation matrices in place of the 0 s and 1 s there. So if, for example, one had a set of identical s-orbitals with one on each center, the transformation matrices under any point group operation would just be 1, so that the matrix would remain just as in Eq. 94. But if one had a set of three equivalent p-functions on each center the 0 s would become 3×3 null matrices and the 1 s would become those three-dimensional transformation matrices which represent the operation on a set of three equivalent p-functions at the origin and so on.

Once these transformation properties have been determined it is an easy matter to apply the projection operator either to construct functions that transform like irreducible representations or like a “special” representation.

Thus in practice we can always get one-particle symmetry functions relatively easily, but the symmetry problem is clearly more involved when it comes to considering many-particle functions.

The paradigm for an approximate many-particle function is a product of one-particle functions (*orbitals*)

$$\Phi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N) = \phi_1(\mathbf{r}_1) \phi_2(\mathbf{r}_2) \dots \phi_N(\mathbf{r}_N). \quad (116)$$

In thinking of approximate solutions to the clamped nuclei problem it is natural to think of the ϕ_r as molecular-orbitals and the \mathbf{r}_i as the electron variables, but the discussion that follows is quite general and it need not even be the case that the variables are those of identical particles. Now from what has been said before it can always be supposed that the orbitals are adapted to particular irreducible representations of the invariance group of the Hamiltonian and that one desires an approximate solution which is also adapted to a particular irreducible representation. The variable-space for the product function on the right-hand side of Eq. 116 is the *direct sum* of each of the variable-spaces for the particles, that is

$\mathbf{r}_1 \oplus \mathbf{r}_2 \oplus \mathbf{r}_3 \oplus \dots \oplus \mathbf{r}_N$. The functions ϕ_r built on each of these variable-spaces individually constitute a separate function space for each variable and the total function space is, therefore, the *direct product* of all these function spaces. If, for the moment, we restrict each of these function spaces to sufficient functions to carry irreducible representations of the invariance group, then the direct product function space is just

$$\phi^{(1)} \otimes \phi^{(2)} \otimes \dots \otimes \phi^{(N)}, \quad (117)$$

where the row matrix $\phi^{(1)}$ contains ϕ_1 and is a basis for the representation $\mathbf{D}^{(1)}$, of dimension n_1 , and so on, so that the total dimension of the space is $n_1 \times n_2 \times \dots \times n_N = P$.

The symmetry operation \mathbf{R} has an orthogonal representative \mathbf{R} in each of the variable-spaces and it has a representative in the total variable-space which is the direct sum of N repeats of \mathbf{R} .

Thus the effect of \mathbf{O}_R on the product function (Eq. 117) is to produce

$$\phi^{(1)}\mathbf{D}^{(1)}(R) \otimes \phi^{(2)}\mathbf{D}^{(2)}(R) \dots \phi^{(N)}\mathbf{D}^{(N)}(R), \quad (118)$$

which, by the ordinary rules of manipulation for direct products, is

$$\phi^{(1)} \otimes \phi^{(2)} \otimes \dots \otimes \phi^{(N)} \mathbf{D}^{(1)}(R) \otimes \mathbf{D}^{(2)}(R) \dots \otimes \mathbf{D}^{(N)}(R). \quad (119)$$

The elements of the row matrix arising from the direct product of the orbitals can obviously be written (with suitable reordering if necessary) as a row matrix of functions like (Eq. 116) but involving such other partner orbitals as are necessary.

Now if it happens that *all* the original orbitals belong to *one*-dimensional irreducible representations of the invariance group, then the matrix of direct products of representations in Eq. 119 is just a scalar, generally speaking ± 1 , so that, in this case, Φ obviously belongs to an irreducible representation of the invariance group of the problem. Otherwise the matrix of direct products of representations generally constitutes a *reducible* representation of the invariance group and in this case the product function (Eq. 116) and its partners do not, generally, carry an irreducible representation of the point group and must be suitably adapted to do so.

There are clearly two possible ways in which this might be done. One could start off from Φ and project out the required components, or one could recognize that there is a matrix relationship between the unadapted set and the adapted set and try to determine the matrix elements from group theoretical arguments. Both approaches have their uses but since the second approach has not yet been considered, it is appropriate to speak of it now.

If, for the moment, we consider the direct product of just two irreducible representations, then it follows at once, from the character orthogonality theorem, that we can write down the number of times that any irreducible representation occurs in the reduction of the direct product representation. If we are thinking of representations μ and ν in the direct product and σ as the resulting representation,

let us call this number $(\mu\nu\sigma)$. It is zero or an integer and of course it is known once the character table for all the irreducible representations of a group are known. This means that we can find a transformation \mathbf{U} such that

$$\mathbf{U}^\dagger \mathbf{D}^{(\mu)} \otimes \mathbf{D}^{(\nu)} \mathbf{U} = \sum_{\sigma} (\mu\nu\sigma) \mathbf{D}^{(\sigma)}, \quad (120)$$

where a direct sum is implied on the right-hand side, which runs over all the representations in the group. In Eq. 120 it is assumed that the representations $\mathbf{D}^{(\mu)}$, etc., are unitary so that from the definition of the direct product, $\mathbf{D}^{(\mu)} \otimes \mathbf{D}^{(\nu)}$ is unitary and can, therefore, be reduced to irreducible form by means of a unitary transformation.

Now if we denote by $\phi^{(\sigma)}$ the row of n_{μ} basis functions for $\mathbf{D}^{(\mu)}$, then the set of $n_{\sigma} = n_{\mu}n_{\nu}$ product functions $\phi^{(\mu)} \otimes \phi^{(\nu)}$ are a basis for the direct product representations and clearly the set of product functions $\phi_s^{(\sigma)}$, $s = 1, 2, \dots, n$ defined by

$$\phi^{(\sigma)} = \phi^{(\mu)} \otimes \phi^{(\nu)} \mathbf{U} \quad (121)$$

are a basis for the reduced representation on the right-hand side of Eq. 120. It is usual to rewrite (Eq. 121) in the form

$$\phi_s^{(\sigma)} = \sum_{i,j} \phi_i^{(\mu)} \phi_j^{(\nu)} (\mu i, \nu j | \sigma s), \quad (122)$$

and refer to the elements of the unitary matrix in Eq. 122 as *vector-coupling* or *Clebsch–Gordan* coefficients.

Care is needed here, because if $(\mu\nu\sigma)$ is zero then clearly (Eq. 122) makes no sense, but this can be taken care of at the formal level by defining the vector-coupling coefficient as zero if $(\mu\nu\sigma)$ is zero. There is also obviously some ambiguity if $(\mu\nu\sigma)$ is greater than one and it will then be necessary to define some standard ordering. But nevertheless, the ideas are clear enough. The point is that if the irreducible representations of the groups are known then it is always possible to discover the vector coupling coefficients and hence to synthesize product functions that carry irreducible representations. This process is often called vector-coupling and obviously having coupled a pair of functions one can then couple another one to that pair and so on. This technique finds its principal use in dealing with the construction of eigenfunctions of the rotation group, but it is a quite general procedure.

There are many different notations for the vector-coupling coefficients and great care must be exercised in determining just which convention a particular book is using, but it is perhaps appropriate to note that there is a common, pretty standard, notation in terms of the *Wigner 3-j* symbol defined, in our particular notation, so that (with the Wigner symbol as the object in brackets on the right-hand side)

$$(\mu i, \nu j | \sigma s)^* = [\sigma]^{1/2} \begin{pmatrix} \mu & \nu & \sigma \\ i & j & s \end{pmatrix}, \quad (123)$$

where $[\sigma]$ is the dimensionality of the σ th irreducible representation. There are, obviously, different sets of 3- j symbols for every group, but the ones most commonly tabulated are those for the three-dimensional rotation group; they are also available for the crystallographic point groups. Whether one chooses synthetic or projective methods to construct many-particle symmetry functions is a matter of choice in any given problem.

Having considered the transformation properties of functions, consider how these properties may be used in the construction of matrix elements of operators and, from this, consider again how operators themselves transform.

Consider the matrix elements of an arbitrary operator \mathbf{A} between two functions. We symbolize the matrix element A_{12} by

$$A_{12} = \int \phi_1^*(\mathbf{x}) \mathbf{A} \phi_2(\mathbf{x}) d\mathbf{x}, \quad (124)$$

where the integral is a definite one over all space and over all coordinates and where $d\mathbf{x}$ symbolizes the appropriate volume element. Since the variables of integration are dummy we may also write

$$A_{12} = \int \phi_1^*(\mathbf{y}) \mathbf{A}' \phi_2(\mathbf{y}) d\mathbf{y}, \quad (125)$$

where \mathbf{A}' symbolizes \mathbf{A} made up with \mathbf{y} just as \mathbf{A} was made up with \mathbf{x} . But now let it be supposed that \mathbf{y} corresponds to a variable change induced by a transformation so that $\mathbf{y} = \mathbf{R}^T \mathbf{x}$. Then

$$A_{12} = \int \bar{\phi}_1^*(\mathbf{x}) \bar{\mathbf{A}} \bar{\phi}_2(\mathbf{x}) |\mathbf{R}| d\mathbf{x}, \quad (126)$$

where $\bar{\mathbf{A}}$ is the transformed operator as in Eq. 98 and $\bar{\phi}_i$ the transformed function as in Eq. 95. $|\mathbf{R}|$ is the Jacobian of the transformation and since the matrix \mathbf{R} is orthogonal, $|\mathbf{R}|$ is ± 1 . If it is -1 then it can be shown that the signs of the limits change also so that the integral is unchanged and thus $|\mathbf{R}|$ can always be treated as $+1$. Thus (Eq. 126) may be rewritten as

$$A_{12} = \int \bar{\phi}_1^*(\mathbf{x}) \bar{\mathbf{A}} \bar{\phi}_2(\mathbf{x}) d\mathbf{x}. \quad (127)$$

Now assume for a moment that \mathbf{A} is \mathbf{H} then of course $\bar{\mathbf{A}}$ is also \mathbf{H} if \mathbf{R} is a symmetry operator of the problem; thus it is the case that

$$\int \phi_1^* \mathbf{H} \phi_2 d\mathbf{x} = \int \bar{\phi}_1^* \mathbf{H} \bar{\phi}_2 d\mathbf{x}. \quad (128)$$

But suppose it turned out that

$$\bar{\phi}_1 \equiv \mathbf{O}_R \phi_1 = -\phi_1, \quad (129)$$

and

$$\bar{\phi}_2 \equiv \mathbf{O}_R \phi_2 = \phi_2,$$

then one would have proved that

$$\int \phi_1^* \mathbf{H} \phi_2 \, d\mathbf{x} = - \int \phi_1^* \mathbf{H} \phi_2 \, d\mathbf{x}, \quad (130)$$

and clearly the only way for this to happen is for the integral to vanish. This is a familiar result as is its generalization, namely, that if ϕ_1 is a basis function for a particular irreducible representation and ϕ_2 is one for another irreducible representation, then the matrix element vanishes unless the direct product of these two representations is reducible to a direct sum of representations which includes the totally symmetric (unit) representation. It should be remembered that although the integral certainly vanishes if the totally symmetric representation is absent, even if this representation is present, the integral could also vanish, for reasons other than the symmetry considered. Thus in the context of the selection rules based on the matrix elements one can say that a particular transition is *forbidden* by symmetry, one cannot say that it is *allowed* by symmetry.

Spin and Point Group Symmetry

So far spin has not been considered explicitly. It is usual to ignore it in discussions of point group symmetry because the total spin operator is invariant under all point group transformations and the axis orientation is defined by the fixed frame choice so there is a fixed internal choice for the z -axis. Thus it is regarded as sufficient to require that the spatial part of any trial function has the correct point group symmetry and then to form properly antisymmetric functions from the spatial parts and the spin eigenfunctions in the manner outlined earlier. If orbitals are used to construct the spatial part, then it is usual simply to extend the symmetry orbitals to become symmetry spin-orbitals. Returning to the functional form used in Eq. 116 this would become

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_N) = \mathbf{A} \phi_1(\mathbf{x}_1) \phi_2(\mathbf{x}_2) \dots \phi_N(\mathbf{x}_N), \quad (131)$$

where \mathbf{A} is the projection operator for the antisymmetric representation of the symmetric group of the electrons. It is usually just called the *antisymmetrizer*:

$$\mathbf{A} = \frac{1}{N!} \sum_p \in p \mathbf{O}_p,$$

where \mathbf{O}_p is the permutation operator and \mathbf{O}_p is the parity of the permutation. \mathbf{O}_p operates on the electronic variables and should thus be associated with

the inverse of a particular permutation. However since the symmetric group is ambivalent and the representation of interest is one-dimensional this is a distinction without a difference. In the case above where the initial function is an orbital product, the operator can equally well be treated as if it operated on the orbital index.

The space parts of the spin-orbitals ϕ_i and ϕ_j can be the same if the spin parts are different. If any two spin-orbitals are the same the projected function simply vanishes. This vanishing is the basis of what is usually called the ‘‘Pauli exclusion principle.’’ The function (Eq. 131) is clearly a determinant of spin-orbitals with the spin-orbital index designating a row (column) and the electron numbering designating a column (row). This was first recognized by Slater and so such determinants are called *Slater determinants* and often denoted by the shorthand

$$M |\phi_1(\mathbf{x}_1) \phi_2(\mathbf{x}_2) \dots \phi_N(\mathbf{x}_N)|,$$

where M is a normalizing factor.

Such Slater determinants are not themselves always spin eigenfunctions but they are eigenfunctions of S_z . They can be either coupled or projected to yield spin eigenfunctions for a particular M_s value and the resulting functions can then be further projected by the step-up operator

$$S_+ = S_x + iS_y$$

to produce functions with $M_s + 1$ or the step-down operator

$$S_- = S_x - iS_y$$

to produce functions with $M_s - 1$.

The presence of spin in the trial function does not modify the previous point group symmetry discussion. However the requirements of spin and permutational symmetry may mean that a trial function is constrained so that it might not be possible to have a trial function with a particular point group symmetry and a specified spin symmetry. As an example consider a determinant of doubly occupied orbitals for four electrons symbolized as

$$M |\phi_1(\mathbf{r}_1) \alpha(s_1) \phi_1(\mathbf{r}_2) \beta(s_2) \phi_2(\mathbf{r}_3) \alpha(s_3) \phi_2(\mathbf{r}_4) \beta(s_4)|.$$

This determinant is a spin eigenfunction with $S = 0$. If the orbitals are point group symmetry orbitals, the only possible symmetry of the many particle function is that generated by the direct product of each orbital symmetry with itself followed by the direct product of the resulting symmetries. If the orbitals belong to one-dimensional irreps then the many particle function must belong to the totally symmetric representation of the point group. It is not possible to represent other symmetries with a function of this form.

The Construction of Approximate Eigenfunctions of the Clamped Nuclei Hamiltonian

The basic mathematical tool used in the construction of approximate eigenfunctions is the variation theorem. This theorem asserts that for any square integrable function Φ which is of a definite and allowed symmetry for the Hamiltonian, the quotient

$$\frac{\int \Phi^* H \Phi d\mathbf{x}}{\int \Phi^* \Phi d\mathbf{x}} = E \geq E_f, \quad (132)$$

where E_f is the lowest energy of the particular symmetry. The function Φ is chosen to contain parameters which can be varied to minimize E to make it as close as possible to E_f . The function is usually scaled so that the denominator in the quotient is unity and the function with this scaling is said to be *normalized to unity* or just to be *normalized*. If Φ is chosen as a linear combination of functions with the coefficients as the parameters then minimizing E leads to a secular problem the roots of which are upper bounds not only to the lowest states of the particular symmetry, but to excited states of increasing energy. At this stage the secular problem can be approximated and approximations to the eigenfunctions can be developed by perturbation theory.

It is usual to construct approximate solutions to the electronic Schrödinger equation in terms of functions composed of spin-orbitals $\phi_r(\mathbf{x})$. In general it is supposed that we have available a set of m ($\geq N$) orbitals $\phi(\mathbf{r})$ and to each of these we attach an $\alpha(s)$ or $\beta(s)$ spin factor function to form the spin-orbitals. When we write \mathbf{x} as a variable designation we shall in future mean space *and* spin variables collectively. We can write formally

$$(\phi_1 \dots \phi_{2m}) = (\alpha\beta) \times (\phi_1 \dots \phi_m) \equiv \phi = \sigma \times \phi, \quad (133)$$

where \times denotes the standard direct (or Kronecker) product of matrices.

The N -electron functions Φ_k are composed of Slater determinants of N spin-orbitals or perhaps as a fixed linear combination of a number of such determinants. The matrix elements of H between such functions can be written in terms of one- and two-electron integrals over the spin-orbitals thus

$$H_{kl} \equiv \langle \Phi_k | H | \Phi_l \rangle = \sum_{rs} {}^{kl} Q_{rs}^1 \langle \phi_r | h | \phi_s \rangle + \frac{1}{2} \sum_{rstu} {}^{kl} Q_{rs,tu}^1 \langle \phi_r \phi_s | g | \phi_t \phi_u \rangle, \quad (134)$$

where the sums go over all spin-orbitals and where ${}^{kl} Q_{rs}^1$ and the ${}^{kl} Q_{rs,tu}^2$ are simple coefficients sometimes called *projective reduction coefficients* or *coupling coefficients* or alternatively one- and two-particle *density matrix elements*. These coefficients will of course be zero if any of the spin-orbitals shown in their indices do not occur in the pair of functions considered. The notation for the one and two-electron integrals is conventional,

$$\langle \phi_r | h | \phi_s \rangle = \int \phi_r^*(\mathbf{r}) h \phi_s(\mathbf{r}) d\mathbf{r}, \quad (135)$$

and

$$\langle \phi_r \phi_s | g | \phi_t \phi_u \rangle = \int \phi_r^*(\mathbf{r}_1) \phi_s^*(\mathbf{r}_2) g(\mathbf{r}_1, \mathbf{r}_2) \phi_t(\mathbf{r}_1) \phi_u(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2. \quad (136)$$

Similarly the overlap matrix elements can be written as

$$M_{kl} = \langle \Phi_k | \Phi_l \rangle = N^{-1} \sum_{rs}^{kl} Q_{rs}^1 \langle \phi_r | \phi_s \rangle. \quad (137)$$

Now let it be supposed that from a given set of spin-orbitals all possible functions Φ_k are constructed and that the resulting set has M members. The best wavefunctions that could be obtained from such a set would be a linear combination of these M terms with coefficients determined using the variation theorem. Use of the variation theorem on such a set leads to the familiar secular problem

$$H - EM | = 0 \quad (138)$$

whose solutions give M eigenvalues $E_1 \leq E_2 \leq E_3 \dots \leq E_M$ each of which is an upper bound to the first, second, third, and so on exact solution of the electronic Schrödinger problem.

In the special case where the spin-orbitals are orthonormal and the trial functions are Slater determinants the expressions for the projective reduction coefficients are both simple and limited, given by Slater's rules to be discussed in detail in later chapters in this work. With such a choice there are Hamiltonian matrix elements between functions that differ from each other only in two or fewer orbitals and $M_{kl} = \delta_{kl}$. The expressions for these coefficients when the orbitals are not orthogonal involve the overlap integrals S_{ij} between all the orbitals in the functions and there is no limitation on orbital differences between the functions and M_{kl} is not the unit matrix. Every electronic permutation must be considered in their evaluation. For non-orthogonal orbitals it is thus much more difficult to consider systems with more than a few electrons and, because atomic orbitals on different centers are not orthogonal, this difficulty has hindered the development of VB theory in a quantitative manner until very recently. An account of modern VB developments forms a later part of this handbook. Usually LCAO MOs are developed so as to be orthogonal so that given

$$\phi_i = \sum_{p=1}^m \eta_p C_{pi} \quad i = 1, 2, 3, \dots, m, \quad (139)$$

to evaluate the two-electron integral $\langle \phi_i \phi_j | g | \phi_k \phi_l \rangle$ involves evaluating all the two-electron integrals $\langle \eta_r \eta_s | g | \eta_t \eta_u \rangle$ and summing. To transform all the two-electron integrals would seem at first sight to involve m^8 multiplications and additions and thus put it beyond computational possibility for anything other than a small number of orbitals. However it actually needs operations only of the order of m^4 , which is bad enough, but can be dealt with.

Although it proved possible to develop the LCAO method with Slater orbitals for diatomic molecules, the evaluation of two-electron integrals when the atomic orbitals were centered on three or four distinct nuclei proved very difficult and is not, even now, commonly undertaken. However it proved possible by the use of Gaussian functions, as pioneered by McWeeny (1950) and developed by Boys (1950) in the 1950s to evaluate the two-electron integrals swiftly and to high precision. Tests on atoms and diatoms showed that although many more Gaussian functions were required to obtain results comparable with a given set of Slater orbitals, the computational time involved was not much longer. Results using Gaussian functions are now regarded as standard for polyatomic systems.

For ease of exposition we shall for the moment assume that the orbitals ϕ_r are orthonormal so that the spin-orbitals are also orthonormal. If the Φ_k are then chosen as normalized Slater determinants or proper linear combinations of them, they too may be chosen to be an orthonormal set so that $M_{kl} = \delta_{kl}$. The secular problem then simplifies to

$$|H - E 1_M| = 0, \quad (140)$$

and this is the form that is typically taken in a CI calculation with orthonormal Mos ϕ_r and orthonormal configuration functions Φ_k , while Eq. 138 is the form typically taken in a VB calculation where, in general, neither the orbitals nor the structure functions are orthonormal.

Let us consider what happens if we subject the spin functions to the transformations

$$\sigma \rightarrow \sigma \mathbf{U}(2), \quad (141)$$

and the orbitals to the transformations

$$\phi \rightarrow \phi \mathbf{U}(m), \quad (142)$$

where $\mathbf{U}(n)$ is an $n \times n$ unitary matrix. In this case the spin-orbitals change

$$\begin{aligned} \phi \rightarrow \phi' &= \sigma \mathbf{U}(2) \times \phi \mathbf{U}(m) \\ &= (\sigma \times \phi) (\mathbf{U}(2) \times \mathbf{U}(m)), \end{aligned} \quad (143)$$

and it is easy to show that the ϕ' form an orthonormal set still.

If we now make up the M functions Φ_k but using the spin-orbitals ϕ'_k rather than the ϕ_k , it is easy to show that the ϕ'_k continue to form an orthonormal set because the ϕ_r are orthonormal. Since the set Φ_k is complete, this means that the ϕ'_k are at most a unitary transformation of the ϕ_k , that is:

$$\Phi'_k = \sum_{n=1}^M \Phi_n U_{nk}(M), \quad (144)$$

where $\mathbf{U}(M)$ is an $M \times M$ unitary matrix. This in turn implies that the matrix \mathbf{H} is transformed as

$$\mathbf{H} \rightarrow \mathbf{H}' = \mathbf{U}^\dagger(M) \mathbf{H} \mathbf{U}(M). \quad (145)$$

The new secular problem is, therefore,

$$|\mathbf{H}' - E \mathbf{1}_M| = 0, \quad (146)$$

which can be rewritten as

$$|\mathbf{U}^\dagger(M) (H - E \mathbf{1}_M) \mathbf{U}(M)| = 0, \quad (147)$$

and using the property

$$|\mathbf{AB}| = |\mathbf{A}| |\mathbf{B}|$$

of determinants of square matrix products it follows at once that the secular problem (Eq. 147) is identical with the secular problem (Eq. 140). Thus the energies are completely unchanged by this unitary transformation.

All possible unitary matrices of a given dimension, say n , form a group under matrix multiplication, a group that is usually written as $U(n)$ and thus we can say that the secular problem is invariant under the operations that constitute the group $U(2) \times U(m)$.

It is wrong to think of this group as a transformation group in the same way as a point group. There is no coordinate transformation that corresponds to the unitary transformation on the orbital space. In this case the transformation on the function space is all that there is, the unitary matrices do not “correspond” to anything though of course, one can assert that each matrix formally represents a particular unitary operator if one wishes. If one does think in this way then we can regard the result that has just been shown as a very weak form of a fundamental result due to Wigner who showed that unitary invariance is a very deep invariance in any quantum mechanics based on square integrable eigenfunctions.

At one level the result just shown is quite useless because for m reasonably large M is astronomical. One can refine the result a little to make M less big, for one

can show that if one starts off with a set of determinants each containing $N_\alpha\alpha$ -spin-orbitals and $N_\beta\beta$ -spin-orbitals ($N_\alpha + N_\beta = N$) so that $M_s = (N_\alpha - N_\beta)/2$ for every determinant, then it is sufficient to include only transformed functions with the same M_s value. Further one can show that if one starts off with functions which are fixed linear combinations of determinants which are spin eigenfunctions (i.e., have a particular expectation value \mathbf{S} say, of \mathbf{S}^2), then again one only need consider the transformed functions that preserve \mathbf{S} . But to give some idea of how big the M can be, the following figures are relevant. For a system where the number of electrons, N , is equal to 10 (as in H_2O) and the number of orbitals, m , is equal to 30, the total number of determinants M is 7.54×10^{10} . If the process is restricted to those functions for which $M_s = 0$, M drops to 2.03×10^{10} and if the process is further restricted to those functions for which $S = 0$ then M becomes 4.04×10^9 .

Clearly it is not possible to compute all the matrix elements and then to solve the secular problem. However it is possible to utilize the group theoretical results in the design of computational methods that yield tractable problems. The key to the connection between matrix element evaluation and the unitary group $\mathbf{U}(m)$ is because the Hamiltonian can be expressed in terms of the generators of this group. The spatial parts of the trial functions can be chosen to provide a basis for irreducible representations of $U(m)$ corresponding to a particular choice of S and of M_s for spin symmetry by the use of *Weyl tableaux*. The required Young diagram for the particular choice of S and of M_s is constructed and then the conjugate diagram written down. Into this diagram the orbital indices are put in such a way that across any row the indices are nondecreasing and down any column they are strictly increasing. The resulting tableaux are the Weyl tableaux and each tableau labels a basis function for an irreducible representation of the group $U(m)$ of dimension equal to the total number of tableaux. Thus for a three electron problem using four orbitals, choosing the Young diagram [2 1] would yield 20 distinct Weyl tableaux. Rather convenient rules may then be specified for matrix elements between such functions. This approach will be considered more fully later in this work.

There is, however, one particularly simple and commonly useful case that is appropriate to consider here. That is the case where the trial wavefunction is simply a single Slater determinant of doubly occupied orbitals and it is a commonly useful form because many molecules have an even number of electrons and their electronic ground states are, in many cases, totally symmetric singlet states. The expected value of the Hamiltonian for such a trial function is

$$E = 2 \sum_r \langle \phi_r | \mathbf{h} | \phi_r \rangle + \sum_{rs} (2 \langle \phi_r \phi_s | \mathbf{g} | \phi_r \phi_s \rangle - \langle \phi_r \phi_s | \mathbf{g} | \phi_s \phi_r \rangle), \quad (148)$$

where the sum is over orbitals, the spin having been integrated out leaving only the spatial integrations to be performed. Using the form (Eq. 139) for the orbitals E may be written as

$$E = 2tr\mathbf{hR} + tr\mathbf{GR},$$

where

$$\begin{aligned}
 h_{ij} &= \langle \eta_i | \mathbf{h} | \eta_j \rangle, \\
 R_{ij} &= \sum_r C_{ir} C_{jr}, \\
 G_{ij} &= 2J_{ij} - K_{ij}, \quad J_{ij} = \sum_{kl} R_{kl} \langle \eta_i \eta_k | g | \eta_j \phi_l \rangle, \quad K_{ij} = \sum_{kl} R_{kl} \langle \eta_i \eta_k | g | \eta_l \phi_j \rangle.
 \end{aligned}$$

It can then be shown that the elements of \mathbf{C} that minimize the energy can be obtained by solving the generalized eigenvalue problem

$$\mathbf{h}^F \mathbf{C} = \mathbf{s} \mathbf{C} \epsilon,$$

where

$$\mathbf{h}^F = \mathbf{h} + \mathbf{G}, \quad S_{ij} = \langle \eta_i | \eta_j \rangle,$$

and ϵ is a diagonal matrix with the orbital energies ϵ_r along the diagonal.

Because the construction of \mathbf{h}^F involves a knowledge of \mathbf{C} the construction of eigenvalues must start with a first guess at \mathbf{C} too, obtain an improved guess, and be continued until no further improvement occurs. There are a number of ways in which this may be done but a popular way is simply to solve the generalized eigenvalue problem iteratively. This scheme has given the name self-consistent field (SCF) to the method generally. The contributions of Hartree and of Fock to its development are sometimes acknowledged by denoting it the HF SCF method.

The SCF method is, at present, the most widely used method in computational chemistry and forms a central feature of all the commonly available computer packages for computational chemistry. Provided that the basis of AOs is full enough, the resulting functions yield energies that are often sufficiently good to enable meaningful distinctions to be made between molecules with different nuclear geometries and so to aid the interpretation of spectroscopic results and reaction mechanisms. However the basis of AOs cannot be made too large because of the number of electron repulsion integrals that must be computed and the difficulties of storing and manipulating these.

It is seen, however, that the distinctive feature in the calculation of the electronic energy is the potential term $V(i)$, all the other terms being of precisely the same form whatever the geometry of a given molecule and between molecules. The potential depends only upon the coordinates of a single particle and thus can be realized in terms of one-particle integrals. This is usually expressed by saying that the potential depends only upon the one-particle density functional. It was shown by Hohenberg and Kohn (1964) that the electronic energy depended on the density functional in a unique way and that, if the density functional was known, then so was the energy. This result has led to the development of variants on the SCF method which involve approximations to the density functional form and require many fewer integral

evaluations than does the direct use of an LCAO MO SCF method. Thus much larger molecules can be tackled in this approach than would otherwise be possible.

Conclusions

It seems fair to say that if one treats the nuclei as distinguishable particles and takes the sum of the electronic energy obtained as an eigenvalue of the clamped nuclei Hamiltonian and the classical nuclear repulsion energy as a potential, the geometrical structure of the minimum in such a potential can be identified as the equilibrium molecular geometry. This can be done without appeal to the chemical bond, and without recognition of functional groups. Certain aspects of the transition state theory of chemical reactions can be rationalized in terms of structures anticipated using the potential at geometries away from the equilibrium one. If the occurrence of polar molecules is considered to be an aspect of molecular structure, polar molecules can be recognized from fixed nuclei electronic structure calculations by means of a dipole moment calculated as the sum of the electronic dipole and the classical nuclear dipole. And there are many other ways in which quantum mechanics formulated with the requirement that the nuclei may be treated as distinguishable clamped particles has been effectively used to illuminate chemical behavior.

However, even if this approach is regarded as giving a satisfactory account of chemical structure, it still remains to justify by full quantum mechanical means the treatment of the nuclei that it involves. But at present such a justification still eludes us. It may be in the future that the multiple well approach to nuclear permutational symmetry will be shown to be properly founded and thus the eigenvalues of the molecular Hamiltonian¹⁰ to be just those anticipated from the previous approach; even so one will still be left with eigenfunctions which exhibit full permutation and rotation-inversion symmetry and it seems impossible to anticipate anything at all like classical chemical structure from these using the standard quantum mechanical machinery.

An obvious objection to this discussion is that the full symmetries of the Coulomb Hamiltonian we have been discussing are manifested necessarily by its stationary states, whereas chemists are concerned with time-dependent states that may exhibit less symmetry. The time-dependent Schrödinger equation simply describes unitary time evolution of a prior state and will not change symmetries; any given initial state is the result of some previous time evolution so what is required is a quantum mechanical theory of initial states that are consistent with chemistry; about that nothing is known. The most that it seems possible to say is that chemical structure can be teased out in terms of a chosen ansatz (the imposition of fixed, distinguishable nuclei) if one has a good idea of what one is looking for. It seems

¹⁰The Coulomb Hamiltonian for the electrons and nuclei specified by the molecular formula.

unlikely that one would ever guess that it was there in the full Coulomb Hamiltonian for a molecular formula unless one had decided on its presence in advance.

Bibliography

- Born, M., & Huang, K. (1955). *Dynamical theory of crystal lattices*. Oxford: Oxford University Press.
- Born, M., & Oppenheimer, J. R. (1927). Zur Quantentheorie der molekeln. *Annalen der Physik*, *84*, 457.
- Boys, S. F. (1950). Electronic wave functions. I. A general method of calculation for the stationary states of any molecular system. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *200*, 542.
- Cassam-Chenai, P. (2006). On non-adiabatic potential energy surfaces. *Chemical Physics Letters*, *420*, 354.
- Collins, M. A., & Parsons, D. F. (1993). Implications of rotation-inversion-permutation invariance for analytic molecular potential energy surfaces. *The Journal of Chemical Physics*, *99*, 6756.
- Combes, J. M., & Seiler, R. (1980). Spectral properties of atomic and molecular systems. In R. G. Woolley (Ed.), *Quantum dynamics of molecules. NATO ASI B57* (p. 435). New York: Plenum.
- Czub, J., & Wolniewicz, L. (1978). On the non-adiabatic potentials in diatomic molecules. *Molecular Physics*, *36*, 1301.
- Deshpande, V., & Mahanty, J. (1969). Born–Oppenheimer treatment of the hydrogen atom. *American Journal of Physics*, *37*, 823.
- Fock, V. (1930). Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Zeitschrift für Physik*, *61*, 126.
- Frolov, A. M. (1999). Bound-state calculations of Coulomb three-body systems. *Physical Review A*, *59*, 4270.
- Hagedorn, G., & Joye, A. (2007). Mathematical analysis of Born–Oppenheimer approximations. In F. Gesztesy, P. Deift, C. Galvez, P. Perry, & W. Schlag (Eds.), *Spectral theory and mathematical physics: A festschrift in honor of Barry Simon's 60th birthday* (p. 203). London: Oxford University Press.
- Handy, N. C., & Lee, A. M. (1996). The adiabatic approximation. *Chemical Physics Letters*, *252*, 425.
- Hartree, D. R. (1927). The wave mechanics of an atom with a non-Coulomb central field. Part I. Theory and methods. *Mathematical Proceedings of the Cambridge Philosophical Society*, *24*, 89.
- Hartree, D. R., & Hartree, W. (1936). Self-consistent field, with exchange, for beryllium. II. The (2s)(2p) ³P and ¹P excited states. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *154*, 588.
- Heitler, W., & London, F. (1927). Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik. *Zeitschrift für Physik*, *44*, 455.
- Herrin, J., & Howland, J. S. (1997). The Born–Oppenheimer approximation: Straight-up and with a twist. *Reviews in Mathematical Physics*, *9*, 467.
- Hinze, J., Alijah, A., & Wolniewicz, L. (1998). Understanding the adiabatic approximation. The accurate data of H₂ transferred to H⁺₃. *Polish Journal of Chemistry*, *72*, 1293.
- Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. *Physical Review*, *136*, 864.
- Hunter, G. (1975). Conditional probability amplitudes in wave mechanics. *International Journal of Quantum Chemistry*, *9*, 237.
- Hunter, G. (1981). Nodeless wave functions and spiky potentials. *International Journal of Quantum Chemistry*, *19*, 755.
- Kato, T. (1951). On the existence of solutions of the helium wave equation. *Transactions of the American Mathematical Society*, *70*, 212.

- Klein, M., Martinez, A., Seiler, R., & Wang, X. P. (1992). On the Born-Oppenheimer expansion for polyatomic molecules. *Communications in Mathematical Physics*, *143*, 607.
- Kolos, W., & Wolniewicz, L. (1963). Nonadiabatic theory for diatomic molecules and its application to the hydrogen molecule. *Reviews of Modern Physics*, *35*, 473.
- Kutzelnigg, W. (2007). Which masses are vibrating or rotating in a molecule? *Molecular Physics*, *105*, 2627.
- Longuet-Higgins, H. C. (1963). The symmetry groups of non-rigid molecules. *Molecular Physics*, *6*, 445.
- McWeeny, R. (1950). Gaussian approximations to wave functions. *Nature*, *166*, 21.
- Mohallem, J. R., & Tostes, J. G. (2002). The adiabatic approximation to exotic leptonic molecules: Further analysis and a nonlinear equation for conditional amplitudes. *Journal of Molecular Structure: Theochem*, *580*, 27.
- Mulliken, R. S. (1931). Bonding power of electrons and theory of valence. *Chemical Reviews*, *9*, 347.
- Nakai, H., Hoshino, M., Miyamoto, K., & Hyodo, S. (2005). Elimination of translational and rotational motions in nuclear orbital plus molecular orbital theory. *The Journal of Chemical Physics*, *122*, 164101.
- Pauling, L. (1939). *The nature of the chemical bond*. Ithaca: Cornell University Press.
- Roothaan, C. C. J. (1951). New developments in molecular orbital theory. *Reviews of Modern Physics*, *23*, 69.
- Slater, J. C. (1930). Note on Hartree's method. *Physical Review*, *35*, 210.
- Sutcliffe, B. T. (2000). The decoupling of electronic and nuclear motions in the isolated molecule Schrödinger Hamiltonian. *Advances in Chemical Physics*, *114*, 97.
- Sutcliffe, B. T. (2005). Comment on "Elimination of translational and rotational motions in nuclear orbital plus molecular orbital theory". *The Journal of Chemical Physics*, *123*, 237101.
- Sutcliffe, B. T. (2007). The separation of electronic and nuclear motion in the diatomic molecule. *Theoretical Chemistry Accounts*, *118*, 563.
- Thirring, W. (1981). *Quantum mechanics of atoms and molecules. A course in mathematical physics* (Vol. 3, E. M. Harrell, Trans.). Berlin: Springer.
- Wilson, E. B. (1979). On the definition of molecular structure in quantum mechanics. *International Journal of Quantum Chemistry*, *13*, 5.
- Woolley, R. G., & Sutcliffe, B. T. (1977). Molecular structure and the Born-Oppenheimer approximation. *Chemical Physics Letters*, *45*, 393.

Remarks on Wave Function Theory and Methods

4

Dariusz Kędziera and Anna Kaczmarek-Kedziera

Contents

Introduction: What and Why?	124
Quantum Mechanics for Dummies	125
On the Way to Quantum Chemistry	133
Variational Principle: An Indicator	136
Perturbation Calculus: The Art of Estimation	137
One-Electron Approximation: Describe One and Say Something About All	142
Hartree–Fock Method: It Is Not That Sophisticated	147
Møller–Plesset Perturbation Theory: HF Is Just the Beginning	152
Beyond the HF Wave Function	159
Coupled Cluster Approximation: The Operator Strikes Back	165
Conclusions	170
Bibliography	171

Abstract

Methods of computational chemistry seem to often be simply a melange of undecipherable acronyms. Frequently, the ability to characterize methods with respect to their quality and applied approximations or to ascribe the proper methodology to the physicochemical property of interest is sufficient to perform research. However, it is worth knowing the fundamental ideas underlying the computational techniques so that one may exploit the approximations intention-

D. Kędziera (✉)

Department of Chemistry and Photochemistry of Polymers, Faculty of Chemistry, Nicolaus Copernicus University, Toruń, Poland
e-mail: teodar@chem.umk.pl

A. Kaczmarek-Kedziera

Faculty of Chemistry, Nicolaus Copernicus University, Toruń, Poland
e-mail: teoadk@chem.umk.pl

ally and efficiently. This chapter is an introduction to quantum chemistry methods based on the wave function search in one-electron approximation.

Introduction: What and Why?

Quantum chemistry is a branch of science originating from quantum mechanics that focuses on investigations of chemical systems. The mathematical roots of quantum chemistry allow it to be treated as a methodology for solving an eigenvalue equation for operators or – even simpler – finding solutions for some differential equations. We cannot totally escape this way of thinking, since this is how things really are. However, a chemist will comprehend quantum chemistry more as a helping tool in experimental work, supporting description of chemical reactions, a tool that plays a role similar to a spectrophotometer or a chromatographic column, a tool that can provide information about the system under consideration, and a powerful tool, the popularization of which was achieved thanks to the fast progress of computer power and the hard work of people who made the transformation from pure theory to computer programs possible. Their efforts were appreciated – in 1998, the Nobel Prize was awarded to Walter Kohn “for his development of the density functional theory” and John Pople “for his development of computational methods.”

From the point of view of the experimentalist, the apparatus of quantum chemistry can be perceived similarly as the NMR spectrometer. One knows that the quality of the obtained NMR spectrum depends not only on the magnetic field of the magnet but also on the signal-processing capabilities. To successfully use NMR spectroscopy in experimental work, detailed knowledge about the technology of production and preparation of the magnets and electronic equipment is not a requisite. It is enough to keep in mind that with the given frequency one gets corresponding accuracy and information. All the rest is simply skill in sample preparation and expertise in spectrum interpretation. For effective usage of computational techniques of quantum chemistry, one must be aware of applied approximation to tune the accuracy of calculations and possess knowledge of the physicochemical phenomenon one wants to describe.

The aim of the present chapter is to provide a gentle introduction to basic quantum chemistry methods – the methods of solving the electronic Schrödinger equation. The chapter is intended for people starting their adventure with computational chemistry and wanting it to become the tool, not the aim itself.

When discussing quantum chemistry, we cannot totally avoid quantum mechanics. However, let us use another comparison: Traveling abroad it is good to know some basic expressions in the local language of the country you go to. It makes life easier and gives pleasure in interpersonal contacts. Still, no one expects a tourist to speak a language as fluently as a native speaker. Therefore, to efficiently apply computational techniques in experimental research, one has to learn some basic quantum mechanical terms that will help during the journey through the remainder of this chapter.

Quantum Mechanics for Dummies

We will begin with the basic terms of quantum mechanics. In this section, they will be introduced in an intuitive manner, to enable understanding of the next sections' content, even by beginners.

We consider a system of N electrons in the field produced by the potentials arising from nuclei (the nuclei are not treated as particles consisting of nucleons but just as a point source of the electrostatic potential). We are interested only in one particular case:

- The probability of finding the electrons of the considered system on the infinite distance from the nuclei is equal to zero. In other words, we want our system to constitute the whole entirety, not breaking into separate and independent parts (this would be the case of the interaction of two electrons with no attraction – two negative charges would repel each other to infinity).
- The energies of this system constitute the discrete spectrum.
- We want to know only the lowest value of energy (the wider approach can be found in the next volume of the present book).

With such limitations, we do not need to consider all of the different general cases and can simply concentrate on the bound-state chemistry.

A central notion in quantum chemistry is a wave function. This is a function characterizing a state of the system. Therefore, it depends on the variables that are adequate for the given system. This means that the wave function has to depend, at least, on spatial coordinates describing motions of the particles in the investigated system. Moreover, the wave function depends on so-called spin variables (spin is an additional degree of freedom included a posteriori in nonrelativistic quantum mechanics). This spin dependency can be built into the wave function by introducing a spin function. For instance, for the electron with a label 1, its wave function depends on the spatial coordinates x_1, y_1 and z_1 and is multiplied by the spin function $\alpha(\sigma_1)$ or $\beta(\sigma_1)$, where σ_1 is a spin variable. The spin functions must fulfill the following requirements:

$$\int \alpha^*(\sigma_1) \alpha(\sigma_1) d\sigma_1 = \int \beta^*(\sigma_1) \beta(\sigma_1) d\sigma_1 = 1 \quad (1)$$

$$\int \alpha^*(\sigma_1) \beta(\sigma_1) d\sigma_1 = \int \beta^*(\sigma_1) \alpha(\sigma_1) d\sigma_1 = 0, \quad (2)$$

where the integration is carried out over the spin variable, which can be treated as the integration variable only. Such a construction may seem to be somehow unnatural; however, it is a convenient way of ascribing spins to the electrons without dealing with its origins.

In general, the wave function must depend on time to reproduce information about the time evolution of the system. However, since we are interested only in the ground state of the system, we can neglect the time dependency. Considering the bound state is equivalent to imposing the condition of the square integrability of the wave function. The integral over all variables in the full range must exist:

$$\iint \dots \int f^* f d\tau = q, \quad (3)$$

where q is a finite real number. An asterisk under the integral denotes the complex conjugate; it comes from the fact that the wave function can be complex in general. The square-integrability condition ensures that the wave function vanishes for the infinite values of all spatial variables and, therefore, that our molecule is kept together. In the above expression, the integration intervals and the integration variables are not stated explicitly. For the investigated N -electron system, the wave function depends on the $3N$ spatial variables (for each particle i , we have x_i , y_i , and z_i coordinates) and additionally N spin variables (σ_i for the particle i):

$$f = f(x_1, y_1, z_1, \sigma_1, x_2, y_2, z_2, \sigma_2, \dots, x_N, y_N, z_N, \sigma_N). \quad (4)$$

The volume element in this $4N$ -dimensional space is

$$d\tau = dV \cdot d\sigma, \quad (5)$$

where the spatial part can be written as

$$dV = dx_1 dy_1 dz_1 dx_2 dy_2 dz_2 \dots dx_N dy_N dz_N, \quad (6)$$

and the spin part is

$$d\sigma = d\sigma_1 d\sigma_2 \dots d\sigma_N. \quad (7)$$

The spatial variables change from $-\infty$ to ∞ and spin variables can take allowed values. One can see that writing all of the integrals, variables, and volume elements explicitly takes time and a lot of paper, even for relatively small systems. Therefore, one usually keeps them in mind, not writing them down.

The wave function contains all of the information about the state of the system. In order to extract it, operators are applied. An operator can be understood by an analogy to a function. The function ascribes a number to a number and the operator ascribes a function to another function. In other words, the operator is a recipe for how to obtain one function from another:

$$\widehat{A}f = g. \quad (8)$$

We will denote operators by hats above the symbol to distinguish them from functions and numbers. One of the particularly interesting cases is when the function g is proportional to the function f ,

$$g = af, \quad (9)$$

where a is a number. Then Eq. 8 takes the form

$$\widehat{A}f = af. \quad (10)$$

Equation 10 is called an eigenvalue equation of the operator \widehat{A} . The function f fulfilling this equation is called an eigenfunction and a is an eigenvalue of the operator \widehat{A} . The Schrödinger equation

$$\widehat{H}\Psi = E\Psi \quad (11)$$

is a typical eigenvalue equation in which the Hamilton operator \widehat{H} extracts the information about the energy E of the system from the wave function Ψ .

Like in the case of the wave function, we will not consider operators in general. Let us concentrate on the Hamilton operator and its properties to simplify our discussion. We need our operators ascribed to observables (Hamiltonian among others) to satisfy the following requirements:

- Linearity – the operators must fulfill the condition

$$\widehat{A}(\alpha f + \beta g) = \alpha\widehat{A}f + \beta\widehat{A}g, \quad (12)$$

where now α and β are numbers. This seems simple and obvious; however, it is not the property of all operators. For instance, the square root is not a linear operator, since the square root of the sum is not equal to the sum of the square roots.

- Real eigenvalues – only real values can be measured in a laboratory.

For these reasons, we will be interested in so-called Hermitian operators that can be defined by the relation

$$\iint \dots \int f_1^* (\widehat{A}f_2) d\tau = \iint \dots \int f_2 (\widehat{A}f_1)^* d\tau. \quad (13)$$

All the functions, variables, and integration intervals remain the same as in Eq. 3. Writing of all these things in the expressions was already troublesome enough, and things become even more complicated when operators appear. In order to make life easier, Dirac notation can be applied. In this notation Eq. 13 has the form

$$\langle f_1 | \widehat{A}f_2 \rangle = \langle \widehat{A}f_1 | f_2 \rangle, \quad (14)$$

where the left-hand side can be equivalently written as $\langle f_1 | \widehat{A} | f_2 \rangle$, and the integral of Eq. 3 becomes simply

$$\langle f | f \rangle = q. \quad (15)$$

In this very convenient notation, it is also assumed that the integration intervals and variables flow from the context.

Let us look at Hermitian operators more carefully, considering them in the example of Hamiltonian. It has already been mentioned that such operators have real eigenvalues. Furthermore, the eigenfunctions of the Hermitian operator that correspond to different eigenvalues are orthogonal. In other words, for

$$\widehat{H} f_1 = E_1 f_1 \quad \text{and} \quad \widehat{H} f_2 = E_2 f_2, \quad (16)$$

where $(E_1 \neq E_2)$, one has

$$\langle f_1 | f_2 \rangle = \langle f_2 | f_1 \rangle = 0. \quad (17)$$

This will be a very useful property, since it will cause various terms in complicated expressions to vanish. In the case of degeneration, or, in other words, when one of the eigenvalues corresponds to two or more eigenfunctions, the eigenfunctions f_1 and f_2 can be orthogonalized.

It is worth considering the integral

$$\langle f_1 | \widehat{H} | f_1 \rangle. \quad (18)$$

Since f_1 is the eigenfunction of \widehat{H} with the eigenvalue E_1 , it is obvious that

$$\langle f_1 | \widehat{H} | f_1 \rangle = \langle f_1 | E_1 f_1 \rangle = E_1 \langle f_1 | f_1 \rangle. \quad (19)$$

It would be certainly more convenient if the result was a single number – the eigenvalue E_1 . This would be the case if $\langle f_1 | f_1 \rangle = 1$ or if we say the function f_1 was normalized to unity. It would be consistent with the interpretation of the integral $\langle f_1 | f_1 \rangle$ as the probability of finding the system in the whole space – it should be surely equal 1. This is a very handy requirement. Any function that does not possess this property can be normalized by multiplying by the normalization factor $\mathcal{N} = 1/\sqrt{\langle f_1 | f_1 \rangle}$. Then, the new function \tilde{f}_1 will be given as

$$\tilde{f}_1 = \mathcal{N} f_1. \quad (20)$$

This new function \tilde{f}_1 is also an eigenfunction of the Hamiltonian, since f_1 was only divided by the number $\sqrt{\langle f_1 | f_1 \rangle}$ and Hamiltonian is linear:

$$\widehat{H} \tilde{f}_1 = \widehat{H} \mathcal{N} f_1 = \mathcal{N} \widehat{H} f_1 = \mathcal{N} E f_1 = E \mathcal{N} f_1 = E \tilde{f}_1. \quad (21)$$

In the case of unnormalized functions, the expression for the eigenvalue E_1 can be obtained from Eq. 19:

$$E_1 = \frac{\langle f_1 | \widehat{H} | f_1 \rangle}{\langle f_1 | f_1 \rangle}. \quad (22)$$

However, many of the applied functions are not the eigenfunctions of the Hamiltonian. Therefore, let us investigate another interesting integral,

$$\langle g | \widehat{H} | g \rangle, \quad (23)$$

where g is not an eigenfunction of \widehat{H} . In order to calculate this integral, an alternate important property of the Hermitian operators needs to be exploited: the fact that their eigenfunctions constitute a complete basis set. Each function depending on the same variables as the eigenfunctions can be expressed as the linear combination of the basis functions. Now, this concept seems to be hard-core mathematics; however, anybody using computational techniques knows well that the two things one must input to the ab initio program are the method and the basis set. Hence, let us make a break from the general considerations of operators and abstract space functions and concentrate for a while on the basis set concept in the example of simple trigonometric functions.

In a calculus course, one learns how to express a function using a set of other functions. For example, consider $\sin x$ function and expand it in the Taylor series around 0:

$$\sin x = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{(2i-1)!} x^{2i-1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \quad (24)$$

In Eq. 24, $\sin x$ function is expressed in the basis set of monomials:

$$\sin x = \sum_{k=1}^{\infty} c_k x^k, \quad (25)$$

where c_k are the expansion coefficients that need to be determined. In our case it is simple, since c_k result directly from the Taylor expansion and are equal:

Table 1 Taylor expansion of x function and standard deviation for various expansion lengths

n	F_n	σ_n
1	x	3.93×10^{-1}
3	$x - \frac{1}{6}x^3$	4.05×10^{-2}
5	$x - \frac{1}{6}x^3 + \frac{1}{120}x^5$	2.08×10^{-3}
7	$x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5,040}x^7$	6.39×10^{-5}

$$c_k = \begin{cases} \frac{(-1)^{(k-1)/2}}{k!} & \text{for odd } k, \\ 0 & \text{for even } k. \end{cases} \quad (26)$$

The summation in Eqs. 24 and 25 goes from 1 to ∞ . In practice, finite and possibly short expansions are applied:

$$F_n(x) = \sum_{i=1}^n c_i x^i. \quad (27)$$

This truncation of the series introduces an approximation to our function.

Let us analyze the x function in the range $x \in \langle -\frac{\pi}{2}; \frac{\pi}{2} \rangle$. The standard deviation works well as the accuracy measure:

$$\sigma_n \sqrt{\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\sin x - F_n(x))^2 dx}. \quad (28)$$

Table 1 summarizes data for small n values. Increasing the number of expansion terms causes a decrease of the standard deviation values and more accurate representation of the original x function. Given the required accuracy of the calculation, the necessary value length of the expansion, n , can be found.

The following question arises: Why use Taylor expansion instead of x function itself, if one needs to worry about the expansion accuracy? The answer is straightforward: simplifications and savings. It is much easier to operate on the polynomials than on the trigonometric functions (for instance, the integral $\int (x^i)^2 dx$ is much easier to handle than $\int \sin^2 x dx$). Moreover, the required accuracy can often be obtained with a relatively short expansion.

Let us now make the considerations more general. As was stated before, the set of eigenfunctions of the Hermitian operator \hat{H} is complete and orthonormal – functions are orthogonal and normalized:

$$\forall_{i,j} \left\langle f_i \left| f_j \right. \right\rangle = \delta_{ij}, \quad (29)$$

where δ_{ij} is a Kronecker symbol that takes value 1 for $i = j$ and 0 otherwise. The basis set completeness means that each function depending on the same set of variables can be expressed by the basis functions:

$$g = \sum_{i=1}^{\infty} c_i f_i, \quad (30)$$

where coefficients c_i need to be found. Knowing the normalized g function makes this task simple, because of the orthonormality of the $\{f_i\}$ set, the coefficients will be equal:

$$c_i = \langle f_i | g \rangle, \quad (31)$$

since

$$\langle f_i | g \rangle = \sum_{j=1}^{\infty} c_j \langle f_i | f_j \rangle = \sum_{j=1}^{\infty} c_j \delta_{ij} = c_i. \quad (32)$$

(Only one term for $i = j$, c_i , remains; all other vanish for the Kronecker delta equals zero if $i \neq j$.) Similarly,

$$\langle g | g \rangle = \sum_{j=1}^{\infty} c_j^* c_j. \quad (33)$$

However, things are not that easy, since we usually apply the expansion (Eq. 30) when we do not know the g function. Thus, the integrals (Eqs. 31 and 33) should be perceived rather as the interpretation of the c_i coefficients than the direct recipe for the calculations. From Eq. 30, the g function can be treated as the linear combination of the f_i functions. Moreover (see Eq. 33), the probability that a system described by the function g is in the state f_j is given by $c_j^* c_j$.

Now let us consider the expression

$$\langle g | \hat{H} | g \rangle \equiv \langle \hat{H} \rangle_g, \quad (34)$$

when g is not the Hamiltonian eigenfunction. Using the expansion Eq. 30 and the fact that f_i are the Hamiltonian eigenfunctions (Eq. 16), one obtains

$$\langle g | \hat{H} | g \rangle = \sum_{j=1}^{\infty} c_j^* c_j E_j. \quad (35)$$

The above integral is called an average (expectation) value, and Eq. 35 for Hamiltonian carries the information about the average energy of the system in the state described by the g function. Looking closer at Eq. 35 shows that this average energy is simply a weighted average of all possible E_j energies of the system. The weights are determined by the $c_j^* c_j$ products – the probability of finding the system in the f_j state. It should be noticed that we used the linearity of the Hamiltonian operator to achieve this result.

Conclusion? Very optimistic: We can say something about the sought energy value not knowing the eigenfunctions of the operator of interest, since for calculations of Eq. 35, we do not need f_j functions. Strange? Not at all, if we recall some linear algebra: Using three basis vectors, we can describe each and every point in the 3D space. Likewise, the wave function can be perceived as a vector, the Hermitian operator as the symmetric transformation matrix, an integral $\langle f | g \rangle$ as the dot product, orthogonality of functions as orthogonality of vectors, and normalization as dividing the vector components by its length.

Now, when the “linear algebra” term has already appeared, let us see how it is applied for solving the eigenequation. Almost all calculations are performed by applying the basis functions. This means that the unknown function Ψ describing the investigated system is expressed in the basis of known functions χ_i (see Eq. 30):

$$\Psi \approx \sum_{i=1}^n c_i \chi_i = \Phi, \quad (36)$$

where Ψ is the eigenfunction of the Hamiltonian corresponding to a given eigenvalue E (Eq. 11). The task is to find such c_i coefficients that the function Φ would be the best approximation to Ψ . Since Φ is an approximation to the wave function, the corresponding energy will also be only approximated. Let us call the approximation E_Φ . Basis functions χ_i are not the Hamiltonian eigenfunctions; therefore, to estimate the energy, an average value must be calculated. Therefore, substituting Eq. 36 for Eq. 11 and multiplying by $\Psi^* = \langle \sum_i c_i \chi_i |$ on both sides gives¹

$$\text{LHS} = \left\langle \sum_i c_i \chi_i \left| \hat{H} \left| \sum_j c_j \chi_j \right. \right. \right\rangle = \sum_i \sum_j c_i^* c_j \langle \chi_i | \hat{H} | \chi_j \rangle, \quad (37)$$

$$\text{RHS} = E_\Phi \sum_i \sum_j c_j^* c_j \langle \chi_i | \chi_j \rangle. \quad (38)$$

¹Here, different subscripts appear on the both sides of the integral. The sum does not depend on the name of the summation index; thus, any subscript can be applied. However, one should not apply the same index on both sides of the integral, since it can cause the erroneous omission of the off-diagonal terms. Compare the overlap integral

$$\begin{aligned} \langle \Psi | \Psi \rangle &= \left\langle \sum_i c_i \chi_i \left| \sum_j c_j \chi_j \right. \right\rangle = \langle c_1 \chi_1 + c_2 \chi_2 + c_3 \chi_3 + \dots | c_1 \chi_1 + c_2 \chi_2 + c_3 \chi_3 + \dots \rangle \\ &= c_1^* c_1 \langle \chi_1 | \chi_1 \rangle + c_1^* c_2 \langle \chi_1 | \chi_2 \rangle + c_1^* c_3 \langle \chi_1 | \chi_3 \rangle + \dots \end{aligned}$$

The explicit writing of all terms shows that not only the integrals $\langle \chi_i | \chi_i \rangle$ with the same function on both sides are present, but also the contributions $\langle \chi_i | \chi_j \rangle$ with $i \neq j$. Therefore, the diversification of the subscripts prevents mistakes.

In order to further simplify the notation, let us denote $\langle \chi_i | \widehat{H} | \chi_j \rangle$ by H_{ij} and $\langle \chi_i | \chi_j \rangle$ by S_{ij} . Then,

$$\sum_i \sum_j c_j^* c_j H_{ij} = E_\Phi \sum_i \sum_j c_i^* c_j S_{ij}. \quad (39)$$

Equivalently, in the matrix form

$$\mathbf{H}\mathbf{c} = \mathbf{S}\mathbf{c}E_\Phi, \quad (40)$$

where \mathbf{H} is the Hamiltonian matrix with the elements H_{ij} , \mathbf{S} is called the overlap matrix and is built of the overlap integrals S_{ij} , and \mathbf{c} denotes the vector of the c_i coefficients. The basis sets applied in practice are usually non-orthogonal, which causes the off-diagonal terms in the \mathbf{S} matrix to not vanish.

Such a method of finding approximate eigenvalues and eigenvectors of the Hamiltonian is known as the Ritz method and is frequently applied in quantum chemistry.

This simple introduction of basic terms of quantum mechanics is obviously far from complete. One can notice the lack of further discussion of the degeneration, the continuum spectrum, and many other topics. For these we encourage the reader to dive into the following excellent books on quantum mechanics and chemistry: Atkins and Friedman (2005), Griffiths (2004), Levine (2008), Lowe and Peterson (2005), McQuarrie and Simon (1997), Piela (2007), Ratner and Schatz (2000), and Szabo and Ostlund (1996).

On the Way to Quantum Chemistry

For the sake of simplification, we assume that the energy of the ground state of our system differs from other energy values. This allows one to avoid embroilment in technical details that are unnecessary at this point. Our system is described by the wave function Ψ fulfilling the Schrödinger Eq. 11. It is important to notice that this eigenvalue equation can be solved exactly only for hydrogen atoms. Any more complicated system requires approximate techniques. In order to explain this complication, let us look into the Hamilton operator. For the system of N electrons and M nuclei, the full Hamiltonian is a sum of the following terms:

- Kinetic energy of electrons, \widehat{T}_e
- Kinetic energy of nuclei, \widehat{T}_n
- Energy of interactions between electrons, \widehat{V}_{ee}
- Energy of interactions between nuclei, \widehat{V}_{nn}
- Energy of interactions between a nucleus and an electron, \widehat{V}_{ne}

In the atomic units, these terms have the following form:

$$\hat{T}_e = -\frac{1}{2} \sum_{i=1}^N \nabla_{\mathbf{r}_i}^2 \quad (41)$$

$$\hat{T}_n = -\sum_{i=1}^M \frac{1}{2m_i} \nabla_{\mathbf{R}_i}^2 \quad (42)$$

$$\hat{V}_{ee} = \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \quad (43)$$

$$\hat{V}_{nn} = \sum_{i=1}^M \sum_{j>i}^M \frac{Z_i Z_j}{R_{ij}} \quad (44)$$

$$\hat{V}_{ne} = -\sum_{i=1}^N \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} \quad (45)$$

where m_i is the mass of the nucleus i , Z_i stands for the nuclear charge, and r_{ij} denotes the distance between the electrons i and j , $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. Likewise, R_{ij} refers to the internuclear distance, $R_{ij} = |\mathbf{R}_i - \mathbf{R}_j|$. The presence of the mutual distances between the particles causes a serious problem when solving the Schrodinger equation; it does not allow one to decouple the equations.

Fortunately, from the chemist's point of view, such a Hamiltonian is not very useful. The chemist is not interested in each and every bit of information one can get about any N -electron M -nuclei system; however, she or he is focused on the given molecule, its conformations, interactions with the environment, and properties (spectroscopic, magnetic, electric, and so on). What makes quantum mechanics a valuable tool for chemists is the Born–Oppenheimer approximation, discussed in detail in the previous chapter of this volume. Let us briefly summarize it to maintain consistent notation throughout the chapter.

The chemist is concerned with the relative positions of the nuclei in the molecule and with the internal energy, but not with the motions of the molecule as a whole. This motion can be excluded from our consideration, for example, by elimination of the center-of-mass translation. Moreover, the intermolecular (electrostatic) forces acting on electrons and nuclei would be similar. This would cause much slower internal motion of the heavy nuclei in comparison to light electrons. For this reason, the approximate description of electron motion with parametric dependence on the static positions of nuclei is justified. Such reasoning leads to adiabatic approximation and finally to Born–Oppenheimer approximation.

According to this approximation, the Hamiltonian can be written as

$$\hat{H} = \hat{T}_n + \hat{H}_e + \hat{V}_{nn}, \quad (46)$$

where \widehat{T}_n now has a meaning of nuclear kinetic energy of the molecule for which the center of mass is stopped (however, there are still oscillations and rotations), and

$$\widehat{H}_e = \widehat{T}_e + \widehat{V}_{ee} + \widehat{V}_{en} \quad (47)$$

is called an electronic Hamiltonian, and it represents the energy of the system after omitting the nuclear kinetic energy term and nuclear repulsion. One can now focus on the solution of the equation of the form

$$\left[\widehat{T}_n(\mathbf{R}) + \widehat{H}_e(\mathbf{r}; \mathbf{R}) + \widehat{V}_{nn}(\mathbf{R}) \right] \Psi(\mathbf{r}, \mathbf{R}) = E \Psi(\mathbf{r}, \mathbf{R}). \quad (48)$$

Here, the dependence on the electronic spatial variables $\mathbf{r} = (x_1, y_1, z_1, \dots, x_N, y_N, z_N)$ and the nuclear spatial variables $\mathbf{R} = (X_1, Y_1, Z_1, \dots, X_M, Y_M, Z_M)$ is written explicitly. The semicolon sign in the \widehat{H}_e term denotes the parametric dependence – for various \mathbf{R} the various electronic equations are obtained.

With such a Hamiltonian, it seems reliable to distinguish also the nuclear $f(\mathbf{R})$ and electronic $\Psi_e(\mathbf{r}; \mathbf{R})$ part in the wave function

$$\Psi(\mathbf{r}, \mathbf{R}) \approx \Psi_e(\mathbf{r}; \mathbf{R}) f(\mathbf{R}), \quad (49)$$

which leads to a significant reduction of the problem.

Now Eq. 48 can be separated into three equations:

$$\widehat{H}_e(\mathbf{r}; \mathbf{R}) \Psi_e(\mathbf{r}; \mathbf{R}) = E_e(\mathbf{R}) \Psi_e(\mathbf{r}; \mathbf{R}) \quad (50)$$

$$\left(\widehat{H}_e(\mathbf{r}; \mathbf{R}) + \widehat{V}_{nn}(\mathbf{R}) \right) \Psi_e(\mathbf{r}; \mathbf{R}) = U(\mathbf{R}) \Psi_e(\mathbf{r}; \mathbf{R}) \quad (51)$$

$$\left(\widehat{T}_n(\mathbf{R}) + \widehat{U}(\mathbf{R}) \right) f(\mathbf{R}) = E f(\mathbf{R}) \quad (52)$$

The first two describe electronic motion for a given position of nuclei. The difference between $E_e(\mathbf{R})$ and $U(\mathbf{R})$ is that in $U(\mathbf{R})$ nuclear repulsion energy is taken into account. These equations are milestones in our considerations for two reasons. First, since we are now talking about “fixed positions of the nuclei,” finally we have got molecules instead of an unspecified system containing some electrons and some nuclei. The second is hidden in Eq. 52: Electronic energy and nuclear repulsion energy constitute the potential, in which nuclei are moving. That is why the proper description of electronic movement in a molecule is so important: The electrons glue the whole molecule together.

Our attention in the rest of the chapter will be focused only on Eq. 50; hence, to simplify notation, all the subscripts denoting the electronic case will be omitted:

$$\widehat{H}_e \rightarrow \widehat{H} \quad (53)$$

$$\Psi_e \rightarrow \Psi \quad (54)$$

$$\widehat{H}_e \Psi_e (r; R) = E_e \Psi_e (r; R) \rightarrow \widehat{H} \Psi = E \Psi \quad (55)$$

The electronic wave function Ψ satisfies all the requirements discussed in the previous sections, depends on the coordinates of N electrons, and additionally must be antisymmetric with respect to the exchange of coordinates of two electrons.²

It should be noted that the analytic solution of Eq. 50 is not known even for the smallest molecule, i.e., H_2 . Therefore, the approximate techniques must be applied to extract the necessary information about molecules of interest. Quantum mechanics provides two tools:

- Variational principle
- Perturbation theory

Variational Principle: An Indicator

The variational principle allows one to judge the quality of the obtained solutions. It can be formulated as follows: For the arbitrary trial function χ that is square-integrable, differentiable, and antisymmetric and depends on the same set of variables as a sought ground-state Ψ_0 function, we have

$$E_0 \leq \frac{\langle \chi | \widehat{H} | \chi \rangle}{\langle \chi | \chi \rangle}, \quad (56)$$

where E_0 is the ground-state energy corresponding to Ψ_0 (Eq. 50). The important consequence of the variational principle is that to estimate the energy of the system, one does not need to solve the eigenequation (this we already know; see Eq. 35), and moreover – what is crucial – the estimated energy value will always not be lower than the exact eigenvalue E_0 .

The proof of the inequality (Eq. 56) is straightforward and can be derived from Eqs. 30 and 35. The function χ that satisfies the above requirements can be expanded on the basis of the Hamiltonian eigenfunctions:

$$\chi = \sum_{i=0}^{\infty} c_i f_i, \quad (57)$$

²In order to explain the antisymmetry requirement, we have to refer again to theory that is beyond the scope of the present chapter. Let us simply state here that wave functions must be antisymmetric without belaboring the point. This will mean that the exchange of the coordinates of the two electrons causes the wave function to change the sign: $\Psi(\tau_1, \tau_2) = -\Psi(\tau_2, \tau_1)$.

where f_i fulfill the eigenproblem $\widehat{H} f_i = E_i f_i$. Thus,

$$\frac{\langle \chi | \widehat{H} | \chi \rangle}{\langle \chi | \chi \rangle} = \frac{\sum_{i=0}^{\infty} c_i^* c_i E_i}{\sum_{i=0}^{\infty} c_i^* c_i} \geq \frac{\sum_{i=0}^{\infty} c_i^* c_i E_0}{\sum_{i=0}^{\infty} c_i^* c_i} = E_0, \quad (58)$$

with the assumption that E_0 is the lowest of all Hamiltonian eigenvalues (Atkins and Friedman 2005; Levine 2008; Lowe and Peterson 2005; McQuarrie and Simon 1997; Piela 2007; Szabo and Ostlund 1996).

Perturbation Calculus: The Art of Estimation

Due to the variational principle that is satisfied for the electronic Hamiltonian, the group of methods of searching for parameters optimizing the energy value can be constructed. The way of verification of the given wave function is the corresponding energy value: The lower, the better. Besides this “quality control,” the variational principle does not give the prescription for the choice of the trial wave functions. Here comes the perturbation calculus – the method frequently applied in physics for the estimation of the functions or values on the basis of partial knowledge about the solutions of the investigated problem. We will consider here the Rayleigh–Schrödinger variant of the perturbation calculus (Atkins and Friedman 2005; Levine 2008; Lowe and Peterson 2005; McQuarrie and Simon 1997; Piela 2007; Ratner and Schatz 2000).

Let us assume that the total electronic Hamiltonian of the investigated system can be divided into

$$\widehat{H} = \widehat{H}^0 + \widehat{H}^1, \quad (59)$$

in such a fashion that we know the exact solutions of

$$\widehat{H}^0 \Psi_k^{(0)} = E_k^{(0)} \Psi_k^{(0)}, \quad (60)$$

where the subscript k enumerates the eigenvalues of the \widehat{H}^0 operator in such a way that $E_0^{(0)}$ is the lowest energy. Now, one can say that the operator \widehat{H} describes the system for which \widehat{H}^0 is an unperturbed operator and \widehat{H}^1 denotes a perturbation. We can assume that if the change in the system represented by \widehat{H}^1 is minor, then the functions $\Psi_k^{(0)}$ would be a good approximation to Ψ_k . Considering \widehat{H}^0 , one postulates its Hermiticity and that its eigenvalues are not degenerate (in our case, for the ground state at least $E_0^{(0)}$ must not be equal to any other eigenvalue). This condition will become clear in a moment.

Knowing only the unperturbed solutions (Eq. 60), we would like to say something more about the ground-state energy of the investigated system. Nothing is easier – we can calculate the average value of the full electronic Hamiltonian with

the $\Psi_0^{(0)}$ function. The variational principle states that the resulting energy will be not lower than the exact energy:

$$E_0 \leq \left\langle \Psi_0^{(0)} \left| \widehat{H}^0 + \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle = E_0^{(0)} + E_0^{(1)}. \quad (61)$$

The term modifying $E_0^{(0)}$ is simply

$$E_0^{(1)} = \left\langle \Psi_0^{(0)} \left| \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle. \quad (62)$$

So far, the only new thing is the manner of partitioning the total energy into the energy of the unperturbed system and the corrections (where $E_0^{(1)}$ is not the only term):

$$E_0 = E_0^{(0)} + E_0^{(1)} + E_0^{(2)} + \dots \quad (63)$$

Likewise, the wave function can be written as

$$\Psi_0 = \Psi_0^{(0)} + \Psi_0^{(1)} + \Psi_0^{(2)} + \dots, \quad (64)$$

where $\Psi_0^{(1)}, \Psi_0^{(2)}$ and so forth are the corrections to the wave function of the unperturbed system $\Psi_0^{(0)}$. Now, the electronic Schrödinger Eq. 50 becomes

$$\begin{aligned} & \left(\widehat{H}^0 + \widehat{H}^1 \right) \left(\Psi_0^{(0)} + \Psi_0^{(1)} + \Psi_0^{(2)} + \dots \right) \\ & = \left(E_0^{(0)} + E_0^{(1)} + E_0^{(2)} + \dots \right) \left(\Psi_0^{(0)} + \Psi_0^{(1)} + \Psi_0^{(2)} + \dots \right). \end{aligned} \quad (65)$$

Introducing the expansions (Eqs. 63 and 64) does not increase our knowledge about the energy or the wave function; it is only a different way of expressing the unknowns by other unknowns. However, now we have a starting point for further investigations.

The comparison of the terms on the left- and right-hand side of the above expression is instructive. Let us regard as similar the terms with the same sum of the superscripts (so-called perturbation order, by analogy to the multiplication and ordering of polynomials). Simple multiplication in Eq. 65 and directing the terms of the same order to separate equations gives

$$\widehat{H}^0 \Psi_0^{(0)} = E_0^{(0)} \Psi_0^{(0)}, \quad (66)$$

$$\widehat{H}^0 \Psi_0^{(1)} + \widehat{H}^1 \Psi_0^{(0)} = E_0^{(0)} \Psi_0^{(1)} + E_0^{(1)} \Psi_0^{(0)}, \quad (67)$$

$$\begin{aligned} \widehat{H}^0 \Psi_0^{(2)} + \widehat{H}^1 \Psi_0^{(1)} &= E_0^{(0)} \Psi_0^{(2)} + E_0^{(1)} \Psi_0^{(1)} + E_0^{(2)} \Psi_0^{(0)}. \\ &\vdots \end{aligned} \quad (68)$$

These equations link the corrections to the wave function and to the energy. Before detailed investigation of the subsequent corrections, one more thing should be underlined. Up to now, the function Ψ_0 is not normalized; only $\Psi_k^{(0)}$ are normalized. Until the corrections to Ψ_0 were found, we would not be able to normalize it. We can only write the normalization constant as $\mathcal{N} = \frac{1}{\sqrt{\langle \Psi_0 | \Psi_0 \rangle}}$. However, it is not necessary at this moment. Now the intermediate normalization condition is more useful:

$$\langle \Psi_0^{(0)} | \Psi_0 \rangle = 1. \quad (69)$$

Such a concept is based on the information that the eigenfunctions of \widehat{H}^0 form an orthonormal complete set of functions (that is one of the reasons why the Hermiticity of \widehat{H}^0 was required) and they can be applied to express any other function, for instance, Ψ_0 as

$$\Psi_0 = \sum_{k=0}^{\infty} c_k \Psi_k^{(0)} + \sum_{k \neq 0}^{\infty} c_k \Psi_k^{(0)}. \quad (70)$$

In this linear combination, the function $\Psi_0^{(0)}$ has a distinguished meaning ($c_0 = 1$); this is the approximation of the wave function of the considered system. Therefore, one can require that $\Psi_0^{(0)}$ does not have a contribution to the higher corrections: $\Psi_0^{(1)}$, $\Psi_0^{(2)}$, and so on:

$$\sum_{k \neq 0}^{\infty} c_k \Psi_k^{(0)} = \Psi_0^{(1)} + \Psi_0^{(2)} + \dots \quad (71)$$

Here, the benefits from the intermediate normalization are obvious: The function $\Psi_0^{(0)}$ is orthogonal to each of the corrections (or, in other words, the corrections are defined in such a way that they are orthogonal to $\Psi_0^{(0)}$).

Therefore, there is an additional set of equations to be satisfied:

$$\langle \Psi_0^{(0)} | \Psi_0^{(n)} \rangle = \delta_{0n}, \quad (72)$$

where the superscript n denotes the n th-order correction to the ground-state wave function Ψ_0 . Now we can go back to Eqs. 66, 67, and 68 and extract the corrections to energy. For this purpose, each of the equations must be multiplied from the left-hand side by $\Psi_0^{(0)}$ and integrated:

$$E_0^{(0)} = \langle \Psi_0^{(0)} | \widehat{H}^0 | \Psi_0^{(0)} \rangle, \quad (73)$$

$$E_0^{(1)} = \langle \Psi_0^{(0)} | \widehat{H}^1 | \Psi_0^{(0)} \rangle, \quad (74)$$

$$E_0^{(2)} = \langle \Psi_0^{(0)} | \widehat{H} | \Psi_0^{(1)} \rangle, \quad \vdots \quad (75)$$

The integrals $\langle \Psi_0^{(0)} | \widehat{H}^0 | \Psi_0^{(n)} \rangle$ vanish, since

$$\langle \Psi_0^{(0)} | \widehat{H}^0 | \Psi_0^{(n)} \rangle = \langle \widehat{H}^0 \Psi_0^{(0)} | \Psi_0^{(n)} \rangle = E_0^{(0)} \langle \Psi_0^{(0)} | \Psi_0^{(n)} \rangle = 0. \quad (76)$$

Thus, obtaining the energy corrections of any order is straightforward. The general expression for the n th-order correction can be written as

$$E_0^{(n)} = \langle \Psi_0^{(0)} | \widehat{H}^1 | \Psi_0^{(n-1)} \rangle \quad \text{for } n > 1. \quad (77)$$

The problem is that to obtain the corrections to the energy in the second or higher orders, the corrections to the wave function are necessary. Then, let us try to find $\Psi_0^{(1)}$. This function can be expressed as the linear combination of the functions from the orthonormal set $\{\Psi_k^{(0)}\}$ for $k \neq 0$:

$$\Psi_0^{(1)} = \sum_{k \neq 0}^{\infty} c_k^{(1)} \Psi_k^{(0)}, \quad (78)$$

where $c_k^{(1)}$ are the expansion coefficients in the first-order correction. Again, the whole thing reduces to finding the coefficients c_k . Substituting Eq. 78 to Eq. 67 gives

$$\left(\widehat{H}^0 - E_0^{(0)} \right) \sum_{k \neq 0}^{\infty} c_k^{(1)} \Psi_k^{(0)} = \left(E_0^{(1)} - \widehat{H}^1 \right) \Psi_0^{(0)}. \quad (79)$$

Integrating this equation with the $\Psi_l^{(0)}$ function leads to

$$\begin{aligned} \text{LHS} &= \left\langle \Psi_l^{(0)} \left| \widehat{H}^0 - E_0^{(0)} \right| \sum_{k \neq 0}^{\infty} c_k \Psi_k^{(0)} \right\rangle = \sum_{k \neq 0}^{\infty} c_k \left\langle \Psi_l^{(0)} \left| \widehat{H}^0 - E_0^{(0)} \right| \Psi_k^{(0)} \right\rangle \\ &= \sum_{k \neq 0}^{\infty} c_k \left(E_l^{(0)} - E_0^{(0)} \right) \left\langle \Psi_l^{(0)} \left| \Psi_k^{(0)} \right\rangle = \sum_{k \neq 0}^{\infty} c_k \left(E_l^{(0)} - E_0^{(0)} \right) \delta_{lk} \\ &= c_l \left(E_l^{(0)} - E_0^{(0)} \right) \end{aligned} \quad (80)$$

and

$$\begin{aligned} \text{RHS} &= \left\langle \Psi_l^{(0)} \left| E_0^{(1)} - \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle = E_0^{(1)} \left\langle \Psi_l^{(0)} \left| \Psi_0^{(0)} \right\rangle - \left\langle \Psi_l^{(0)} \left| \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle \\ &= - \left\langle \Psi_l^{(0)} \left| \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle. \end{aligned} \quad (81)$$

Altogether, these allow one to write the coefficients of the expansion (Eq. 78) as

$$c_l = \frac{\langle \Psi_l^{(0)} | \widehat{H}^1 | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_l^{(0)}}. \quad (82)$$

Hence, the first correction to the wave function is already known:

$$\Psi_0^{(1)} = \sum_{k \neq 0}^{\infty} \frac{\langle \Psi_k^{(0)} | \widehat{H}^1 | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_k^{(0)}} \Psi_k^{(0)}, \quad (83)$$

and, thereby, the second-order correction to the energy can be calculated as

$$E_0^{(2)} = \sum_{k \neq 0}^{\infty} \frac{\langle \Psi_k^{(0)} | \widehat{H}^1 | \Psi_0^{(0)} \rangle}{E_0^{(0)} - E_k^{(0)}} \langle \Psi_0^{(0)} | \widehat{H}^1 | \Psi_k^{(0)} \rangle. \quad (84)$$

This is also equivalently written as

$$E_0^{(2)} = \sum_{k \neq 0}^{\infty} \frac{\left| \langle \Psi_k^{(0)} | \widehat{H}^1 | \Psi_0^{(0)} \rangle \right|^2}{E_0^{(0)} - E_k^{(0)}}. \quad (85)$$

The energy difference in the denominator of the above expression cannot be equal to zero, and for this reason the non-degenerated ground state was assumed. The higher-order corrections are sought in a similar manner, which just requires more operations.

One of the interesting issues is the problem of variationality of the perturbation calculus built upon the variational Hamiltonian. This is, however, a sophisticated problem for advanced readers and will not be discussed here. It should be added, in summary, that the manner of partitioning of the Hamilton operator was arbitrary. The only prerequisites were the Hermitian character of the operators (for the eigenfunctions to form the orthonormal set) and the non-degenerated ground-state eigenenergy and nothing more. One should also remember that, in practice, even solving the unperturbed problem cannot be performed exactly and the approximations must be applied. The consequence could be the loss of accuracy for higher-order corrections. Moreover, good convergence of the perturbation expansion can be expected when the consecutive corrections are small in comparison to the total estimated value. However, in such a case, the low orders of the series would reproduce the sought value with relatively good accuracy. Thus, application of the low orders of perturbation calculus is highly recommended.

One-Electron Approximation: Describe One and Say Something About All

Equipped with general knowledge about the tools for the Schrödinger equation solution, one can move to many-electron systems.

The electronic Hamiltonian for any many-electron system in atomic units has the following form (compare Eqs. 41, 42, 43, 44, and 45):

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^N \Delta_{r_i} + \sum_{i=1}^N \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} + \sum_{i=1}^N \sum_{j=i+1}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (86)$$

Let us look more closely. In the first term, we sum up over the number of electrons N ; in the second term, the summations run over the number of electrons N and number of nuclei M ; and the third term contains the double sum over the number of electrons N . Thus, one can simplify the notation of the first two terms:

$$-\frac{1}{2} \sum_{i=1}^N \Delta_{r_i} + \sum_{i=1}^N \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} = \sum_{i=1}^N \left(-\frac{1}{2} \Delta_{r_i} + \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|} \right). \quad (87)$$

Now, denoting the term in parenthesis by $\hat{h}(i)$,

$$\hat{h}(i) = -\frac{1}{2} \Delta_{r_i} + \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|}, \quad (88)$$

we get the part of the Hamiltonian depending only on the coordinates of one electron i (and nuclear coordinates, but it does not bother us). Hence, the total electronic Hamiltonian (Eq. 86) can be rewritten as the sum of one-electron and two-electron contributions:

$$\hat{H} = \sum_{i=1}^N \hat{h}(i) + \sum_{i=1}^N \sum_{j=i+1}^N \hat{g}(i, j), \quad (89)$$

where we introduced a symbol:

$$\hat{g}(i, j) = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (90)$$

It should be noticed that each of the one-electron Hamiltonians $\hat{h}(i)$ describes a single electron in the field of some potentials. Therefore, the exact solutions of the eigenvalue problem for these one-electron operators are available. The problem lies

in the $\hat{g}(i, j)$ operator that couples two electrons together: It is not possible to separate their coordinates exactly.

All electronic Hamiltonians have the same general form; they differ only in the number of electrons N and the nuclear potential hidden in $\hat{h}(i)$. One can choose any possible chemical compounds and try to write the corresponding equations; however, one would soon note the similarity of all of them. Therefore, we will not invest time in describing the procedure for a polypeptide or a nanotube, but for simplicity we will start from the two-electron helium atom (the simplest many-electron system) and later try to generalize the considerations. For the helium atom:

- $N = 2$ – two electrons
- $M = 1$ – one nucleus

The generalization of the helium discussion into the larger (N -electron) systems should be straightforward:

$$\hat{h}(1) + \hat{h}(2) = \sum_{i=1}^2 \hat{h}(i) \rightarrow \sum_{i=1}^N \hat{h}(i), \quad (91)$$

$$\hat{g}(1, 2) = \sum_{i=1}^2 \sum_{j=i+1}^2 \hat{g}(i, j) \rightarrow \sum_{i=1}^N \sum_{j=i+1}^N \hat{g}(i, j), \quad (92)$$

and finally

$$\hat{h}(i) = -\frac{1}{2}\Delta_{r_i} + \frac{Z_1}{|\mathbf{r}_i - \mathbf{R}_1|} \rightarrow -\frac{1}{2}\Delta_{r_i} + \sum_{j=1}^M \frac{Z_j}{|\mathbf{r}_i - \mathbf{R}_j|}. \quad (93)$$

Recall that the exact solutions of the one-electron problem are known, and the task is to solve the full problem. The ideas of the perturbation theory were explained in the previous section. Now it is the time to apply the knowledge. The one-electron part can be treated as the unperturbed Hamiltonian and the rest as the perturbation:

$$\hat{H}^0 = \hat{h}(1) + \hat{h}(2), \quad \hat{H}^1 = \hat{g}(1, 2), \quad (94)$$

where the normalized solutions for the one-electron part are known:

$$\hat{h}(1)\phi_i(1) = \epsilon_i\phi_i(1), \quad (95)$$

$$\hat{h}(2)\phi_j(2) = \epsilon_j\phi_j(2). \quad (96)$$

These functions require more attention. Although the electronic Hamiltonian – and thereby the one-electron operators – do not act on the spin variables, the wave

functions ϕ_k must carry on the spin dependence. Therefore, the function $\phi_k(l)$ is a product of the spatial part depending on the three spatial coordinates of the electron l and on the spin part and is called a spin-orbital. For simplicity, the set of coordinates τ_l is written as a label of the electron, i.e., l . Such convention will be applied from now on, with an exception where the τ_l labeling is really needed.

If an operator can be written as a sum of contributions acting on different variables, its eigenfunction takes a form of the product of the eigenfunctions of the subsequent operators in the summation. In the case of the helium atom, where \widehat{H}^0 is a sum of $\widehat{h}(1)$ and $\widehat{h}(2)$, the wave function $\Psi(1, 2)$ can be denoted as the product of one-electron functions:

$$\Psi(1, 2) = \phi_i(1)\phi_j(2). \quad (97)$$

The eigenproblem for such a function gives the eigenvalue that is simply the sum of the one-electron eigenvalues:

$$\begin{aligned} [\widehat{h}(1) + \widehat{h}(2)] \phi_i(1)\phi_j(2) &= \widehat{h}(1)\phi_i(1)\phi_j(2) + \widehat{h}(2)\phi_i(1)\phi_j(2) \\ &= [\widehat{h}(1)\phi_i(1)] \phi_j(2) + [\widehat{h}(2)\phi_j(2)] \phi_i(1) \\ &= [\epsilon_i \phi_i(1)] \phi_j(2) + [\epsilon_j \phi_j(2)] \phi_i(1) \\ &= [\epsilon_i + \epsilon_j] \phi_i(1)\phi_j(2). \end{aligned} \quad (98)$$

However, here the antisymmetry requirement should also be taken into account. The many-electron function must change the sign with respect to the exchange of labels of any two electrons:

$$\Psi(1, 2) = -\Psi(2, 1). \quad (99)$$

The product (Eq. 97) is not antisymmetric; the interchange of the electron labels leads to

$$\Psi(2, 1) = \phi_i(2)\phi_j(1) \neq -\Psi(1, 2), \quad (100)$$

and the result is a function different from the original $\Psi(1, 2)$. But another function,

$$\Psi(1, 2) \sim [\phi_i(1)\phi_j(2) - \phi_i(2)\phi_j(1)], \quad (101)$$

satisfies the antisymmetry condition.

Moreover, the wave function $\Psi(1, 2)$ has to be normalized. This can be achieved by calculating the following (overlap) integral:

$$\begin{aligned}
\langle \Psi(1, 2) | \Psi(1, 2) \rangle &= \langle \mathcal{N}(\phi_i(1)\phi_j(2) - \phi_j(1)\phi_i(2)) | \mathcal{N}(\phi_i(1)\phi_j(2) - \phi_j(1)\phi_i(2)) \rangle \\
&= \mathcal{N}^2 \left(\langle \phi_i(1)\phi_j(2) | \phi_i(1)\phi_j(2) \rangle - \langle \phi_i(1)\phi_j(2) | \phi_j(1)\phi_i(2) \rangle \right. \\
&\quad \left. - \langle \phi_j(1)\phi_i(2) | \phi_i(1)\phi_j(2) \rangle + \langle \phi_j(1)\phi_i(2) | \phi_j(1)\phi_i(2) \rangle \right). \tag{102}
\end{aligned}$$

The spin-orbitals of electrons 1 and 2 are mutually independent; thus,

$$\langle \phi_i(1)\phi_j(2) | \phi_k(1)\phi_l(2) \rangle = \langle \phi_i(1) | \phi_k(1) \rangle \langle \phi_j(2) | \phi_l(2) \rangle = \delta_{ik}\delta_{jl}, \tag{103}$$

where $\langle \phi_i(1) | \phi_k(1) \rangle = \delta_{ik}$ arises from the fact that the eigenfunctions of the Hermitian operator $\widehat{h}(1)$ form the orthonormal set. Hence, in Eq. 102 only the first and last integral will be nonvanishing, and, finally,

$$\langle \Psi(1, 2) | \Psi(1, 2) \rangle = 2\mathcal{N}^2 = 1, \tag{104}$$

and, therefore, the normalization constant \mathcal{N} must equal $1/\sqrt{2}$. In order to fulfill the normalization and the antisymmetry request, the trial wave function can be written as

$$\Psi(1, 2) = \frac{1}{\sqrt{2}} (\phi_i(1)\phi_j(2) - \phi_j(1)\phi_i(2)) = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_i(1) & \phi_i(2) \\ \phi_j(1) & \phi_j(2) \end{vmatrix}. \tag{105}$$

Thereafter, the expectation value of the Hamiltonian can be calculated using the same tricks as in Eq. 102³:

$$\begin{aligned}
\langle \widehat{H}(1, 2) \rangle_{\Psi(1,2)} &= \langle \Psi(1, 2) | \widehat{H}(1, 2) | \Psi(1, 2) \rangle \\
&= \langle \phi_i(1) | \widehat{h}(1) | \phi_i(1) \rangle + \langle \phi_j(2) | \widehat{h}(2) | \phi_j(2) \rangle \\
&\quad + \langle \phi_i(1)\phi_j(2) | \frac{1}{r_{12}} | \phi_i(1)\phi_j(2) \rangle \\
&\quad - \langle \phi_i(1)\phi_j(2) | \frac{1}{r_{12}} | \phi_j(1)\phi_i(2) \rangle. \tag{106}
\end{aligned}$$

³Since the value of the integral does not depend on the name of the variable, the obvious equality $\int_a^b f(x)dx = \int_a^b f(y)dy$ in the above case takes the form

$$\langle \phi_i(1)\phi_j(2) | \frac{1}{r_{12}} | \phi_j(1)\phi_i(2) \rangle = \langle \phi_i(2)\phi_j(1) | \frac{1}{r_{12}} | \phi_j(2)\phi_i(1) \rangle.$$

The only difference with respect to the overlap integral is that the integrals containing $1/r_{12}$ cannot be separated.

In the spirit of perturbation calculus, the expectation value (Eq. 106) can be treated as the energy corrected up to the first order of perturbation expansion. According to the variational principle, this integral is an upper bound of the exact energy of the two-electron system under consideration. The recipe for correcting the trial function is given by the perturbation theory: Apply functions corresponding to the remaining states of the unperturbed system in the expansion. In other words, in order to improve the wave function, the expansion built from products of the remaining states of the systems should be used.

Let us summarize. The many-electron function of the system can be approximately written as the antisymmetrized product of the one-electron functions being the solutions for the one-electron eigenvalue problem. This is the idea of the popular one-electron approximation (Atkins and Friedman 2005; Lowe and Peterson 2005; McQuarrie and Simon 1997; Ratner and Schatz 2000). In the N -electron case, one obtains the trial function as the antisymmetrized product of N one-electron functions that can be written in the form of the Slater determinant:

$$\Psi(1, 2, \dots, N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \dots & \phi_1(N) \\ \phi_2(1) & \phi_2(2) & \dots & \phi_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(1) & \phi_N(2) & \dots & \phi_N(N) \end{vmatrix}. \quad (107)$$

$1/\sqrt{N!}$ factor ensures the normalization of the wave function. Often, instead of writing the whole determinant, only the diagonal is written down:

$$\begin{vmatrix} \phi_1(1) & \phi_1(2) & \dots & \phi_1(N) \\ \phi_1(2) & \phi_2(2) & \dots & \phi_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(1) & \phi_N(2) & \dots & \phi_N(N) \end{vmatrix} = |\phi_1(1)\phi_2(2) \dots \phi_N(N)|. \quad (108)$$

It is important to realize that the many-electron function in the form of the Slater determinant is only an approximation. Intuition tells us that describing the many-electron system using only one-electron functions cannot be exact. One needs to be aware of the fact that such an approach causes the loss of some information included in the sought wave function. In particular, the one-electron function cannot “see” another electron; therefore, the terms coupling the mutual electron positions are missing in the one-electron approximation.

For example, consider the two-electron function (Piela 2007)

$$F(1, 2) \sim (e^{-ar_1-br_2-cr_{12}} - e^{-ar_2-br_1-cr_{12}}), \quad (109)$$

where r_1, r_2 are the electron–nucleus distances, r_{12} denotes the distance between two electrons, and a, b , and c stand for the coefficients. $F(1, 2)$ contains the factor c correlating the electron motions. In the one-electron approach, such a function would be approximated by the antisymmetrized function

$$f(1, 2) \sim (e^{-ar_1-br_2} - e^{-ar_2-br_1}), \quad (110)$$

with total neglect of this correlation. From the practical point of view, this means that the trial function written as the single determinant does not allow reproduction of the exact electronic energy, even when using the best possible one-electron functions for its construction.

The question arises: Why use such an approach, knowing from the very beginning that it is bad? The answer is simple. First, including the electron correlation in the wave function is very expensive. For two electrons, one additional term appears; for three, three terms; for four, six. In general, for N electrons there are $N(N - 1)/2$ terms (the triangle of the $N \times N$ matrix without the diagonal elements). Therefore, the number of coefficients describing the electron correlation is much bigger than the number of one-electron terms. Calculating even the one-electron coefficients is very time-consuming, and moreover, calculating the overlap integrals and Hamiltonian matrix elements becomes prohibitively complicated with the correlated functions. Second, perturbation theory provides ways of improving the results. Expansion built on a higher number of determinants will lead to better energy. Third, a chemist does not usually need exact data but only an appropriate accuracy (furnishing a house does not require caliper measurements, just a quick glance to estimate the size of the door and the furniture).

Therefore, let us stick to the one-electron approximation. The next section will explain how to find the best possible spin-orbitals.

Hartree–Fock Method: It Is Not That Sophisticated

Now we are prepared to concentrate on the methods for solving the electron equation. As you will see, they are only an extension of the already-discussed techniques. We start with the fundamental Hartree–Fock method.

The main goal of this approximation is to find the spin-orbitals applied for construction of the Slater determinant that will best reproduce the exact wave function. We again begin our considerations with the two-electron system. The problem is that the operator $\hat{h}(i)$ does not contain the part arising from the potential of the second electron, and an operator responsible for this missing part must be found. What we do know is that such an interaction is included in the two-electron part of Eq. 106:

$$\left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)} = \left\langle \phi_1(1)\phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1)\phi_j(2) \right\rangle - \left\langle \phi_i(1)\phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_j(1)\phi_i(2) \right\rangle. \quad (111)$$

Unfortunately, Eq. 111 contains the integrals that cannot be exactly separated into a product of the simpler integrals depending only on the coordinates of one electron. What we can propose is rewriting this equation in the following form:

$$\left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)} = \langle \phi_i(1) | \widehat{v}(1) | \phi_i(1) \rangle + \langle \phi_j(2) | \widehat{v}(2) | \phi_j(2) \rangle + \text{the rest}, \quad (112)$$

where $\widehat{v}(1)$ and $\widehat{v}(2)$ are introduced to extract only the terms depending on the coordinates of a single electron from Eq. 111 and “the rest” is what remains from $\left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)}$ after this extraction.

If such extraction were possible, the operators $\widehat{v}(1)$ and $\widehat{v}(2)$ could be applied to improve our one-electron operators, which leads to the following equations:

$$\left[\widehat{h}(1) + \widehat{v}(1) \right] \phi_i(1) = \epsilon_i \phi_i(1), \quad (113)$$

$$\left[\widehat{h}(2) + \widehat{v}(2) \right] \phi_j(2) = \epsilon_j \phi_j(2). \quad (114)$$

The solutions of such equations (spin-orbitals ϕ_i, ϕ_j) can be used to build up the Slater determinant. They ensure better approximation than Eqs. 95 and 96, since they somehow provide for the influence of the second electron.

Therefore, our goal now is to utilize Eq. 106 (treating ϕ_i and ϕ_j) as known functions) to find the best form of the operators $\widehat{v}(1)$ and $\widehat{v}(2)$. Adding zero, written as

$$0 = \left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)} - \left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)}, \quad (115)$$

to Eq. 111 allows one to ascribe, for instance,

$$\begin{aligned} \langle \phi_i(1) | \widehat{v}(1) | \phi_i(1) \rangle &= \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_j(2) - \phi_j(1) \phi_i(2) \right\rangle, \\ \langle \phi_j(2) | \widehat{v}(2) | \phi_j(2) \rangle &= \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_j(2) - \phi_j(1) \phi_i(2) \right\rangle, \\ \text{the rest} &= - \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_j(2) - \phi_j(1) \phi_i(2) \right\rangle. \end{aligned}$$

Consider in more detail the integral containing $\widehat{v}(1)$. We want it to be expressed in such a way that only the coordinates of the electron labeled by 1 are explicitly written under the integral:

$$\begin{aligned} \langle \phi_i(1) | \widehat{v}(1) | \phi_i(1) \rangle &= \left\langle \phi_i(1) \phi_j(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_j(2) - \phi_j(1) \phi_i(2) \right\rangle \\ &= \int \int \phi_i^*(\tau_1) \phi_j^*(\tau_2) \frac{1}{r_{12}} \phi_i(\tau_1) \phi_j(\tau_2) d\tau_1 d\tau_2 \\ &\quad - \int \int \phi_i^*(\tau_1) \phi_j^*(\tau_2) \frac{1}{r_{12}} \phi_j(\tau_1) \phi_i(\tau_2) d\tau_1 d\tau_2 \\ &= \int \phi_i^*(\tau_1) \left(\int \phi_j^*(\tau_2) \frac{1}{r_{12}} \phi_j(\tau_2) d\tau_2 \right) \phi_i(\tau_1) d\tau_1 \end{aligned}$$

$$\begin{aligned}
& - \int \phi_i^* (\tau_1) \left(\int \phi_j^* (\tau_2) \frac{1}{r_{12}} \phi_i (\tau_2) d\tau_2 \right) \phi_j (\tau_1) d\tau_1 \\
& = \int \phi_i^* (\tau_1) \left(\widehat{J}_j(1) \phi_i (\tau_1) \right) d\tau_1 \\
& \quad - \int \phi_j^* (\tau_1) \left(\widehat{K}_j(1) \phi_i (\tau_1) \right) d\tau_1 \tag{116} \\
& = \left\langle \phi_i(1) \left| \widehat{J}_j(1) \right| \phi_i(1) \right\rangle - \left\langle \phi_i(1) \left| \widehat{K}_j(1) \right| \phi_i(1) \right\rangle \\
& = \left\langle \phi_i(1) \left| \widehat{J}_j(1) - \widehat{K}_j(1) \right| \phi_i(1) \right\rangle,
\end{aligned}$$

where two operators were defined:

$$\widehat{J}_j(1) \phi_i(1) = \left(\int \phi_j^* (\tau_2) \frac{1}{r_{12}} \phi_j (\tau_2) d\tau_2 \right) \phi_i(1), \tag{117}$$

$$\widehat{K}_j(1) \phi_i(1) = \left(\int \phi_j^* (\tau_2) \frac{1}{r_{12}} \phi_i (\tau_2) d\tau_2 \right) \phi_j(1). \tag{118}$$

Despite a slightly different way of defining, they are still ordinary operators. An operator is a function of a function. The operators $\widehat{J}_j(1)$ and $\widehat{K}_j(1)$ act on the function $\phi_i(1)$, producing another function, as was written in Eq. 8. The operator $\widehat{J}_j(1)$ acting on $\phi_i(1)$ transforms it into the same function:

$$\phi_i(1) \xrightarrow{\widehat{J}_j(1)} \left(\int \phi_j^* (\tau_2) \frac{1}{r_{12}} \phi_j (\tau_2) d\tau_2 \right) \phi_i(1), \tag{119}$$

and the operator $\widehat{K}_j(1)$ produces $\phi_j(1)$:

$$\phi_i(1) \xrightarrow{\widehat{K}_j(1)} \left(\int \phi_j^* (\tau_2) \frac{1}{r_{12}} \phi_i (\tau_2) d\tau_2 \right) \phi_j(1). \tag{120}$$

The expressions on the right-hand side of the arrows are some functions of a dependent variable τ_1 (the integration eliminates the dependence on τ_2 ; however, its result is not a number but a function depending on τ_1). The interpretation of the $\widehat{J}_j(1)$ and $\widehat{K}_j(1)$ operators by ascribing them to observables is not straightforward. These operators appear in the equations when we try to write the interaction between two electrons as the average value of the one-electron operator calculated with the one-electron function. However, in fact, they do appear as the difference:

$$\widehat{v}_i^{\text{HF}}(1) = \widehat{J}_j(1) - \widehat{K}_j(1). \tag{121}$$

And the physical sense should be sought in this difference. Here, $\widehat{v}_i^{\text{HF}}(1)$ is an operator of the average interaction of the electron labeled as 1, described by a spin-orbital ϕ_i with the second electron characterized by $\phi_j(2)$. It should be noticed that the potential $\widehat{v}_i^{\text{HF}}(1)$ depends on $\phi_j(2)$, since this spin-orbital is necessary to define $\widehat{J}_j(1)$ and $\widehat{K}_j(1)$ operators.

Similarly,

$$\widehat{v}_j^{\text{HF}}(2) = \widehat{J}_i(2) - \widehat{K}_i(2), \quad (122)$$

where the action of operators $\widehat{J}_j(1)$ and $\widehat{K}_j(1)$ on the function $\phi_j(2)$ is defined by

$$\widehat{J}_i(2)\phi_j(2) = \left(\int \phi_i^*(\tau_1) \frac{1}{r_{12}} \phi_i(\tau_1) d\tau_1 \right) \phi_j(2) \quad (123)$$

$$\widehat{K}_i(2)\phi_j(2) = \left(\int \phi_i^*(\tau_1) \frac{1}{r_{12}} \phi_j(\tau_1) d\tau_1 \right) \phi_i(2). \quad (124)$$

Here, one more important circumstance should be mentioned. The electronic Hamiltonian does not depend on spin. However, the electronic wave function is spin dependent. During the construction of the HF equations, this dependence is introduced to the operators \widehat{K} since they depend on two different spin-orbitals.

Let us briefly review. We want to have a one-electron operator that includes the interaction between the electrons in some averaged way. Such an operator must depend on the function describing the motion of the second, adjacent electron. Therefore, for the two-electron case, two coupled equations must be solved (Eqs. 113 and 114). The word “coupled” that distinguishes this set of equations from Eqs. 95 and 96 is crucial. Denoting

$$\widehat{f}_i(1) = \widehat{h}(1) + \widehat{v}_i^{\text{HF}}(1), \quad (125)$$

one can rewrite the above equations as

$$\widehat{f}_k(1)\phi_k(1) = \epsilon_k\phi_k(1), \quad \text{for } k = i, j. \quad (126)$$

$\widehat{f}_i(1)$ is called the Fock operator and Eq. 114 gives the Hartree–Fock equations.

It should be noted that the label of electrons determines only the name of the integration variables, and the result of the integration does not depend on the name of the variable. Therefore, what is really important is the label of the spin-orbital. It will be even more pronounced in the N -electron case, when the electron labels are applied only to show that the operator acts on one or two electron coordinates. The form of the N -electron wave function depends only on the spin-orbitals and not on the electron labels; they are just the integration variables.

The most popular way of solving Eq. 126 is the iterative procedure. It starts from the guessed or chosen spin-orbitals $\phi_i(1)$ and $\phi_j(2)$ applied to construct the potentials $\widehat{v}_i^{\text{HF}}(1)$ and $\widehat{v}_j^{\text{HF}}(2)$. Next, the obtained potentials are substituted to Eq. 125 and solutions of Eq. 126 give the improved form of the $\phi_i(1)$ and $\phi_j(2)$ spin-orbitals. These, on the other hand, are treated as the starting point again, and the whole procedure is repeated until the starting and final orbitals of the given iteration

do not differ much. This technique is called self-consistent field (SCF). Often the abbreviations HF for Hartree–Fock method and SCF are used interchangeably. They can also be joined together as SCF–HF, denoting the self-consistent way of solving Hartree–Fock equations.

Let us assume that the spin-orbitals are already known. Concentrate on the calculations of the average value of the Hamiltonian with the determinant build of these spin-orbitals using the operators defined previously. Writing down the two-electron part:

$$\begin{aligned} \left\langle \frac{1}{r_{12}} \right\rangle_{\Psi(1,2)} &= \left\langle \phi_i(1) \left| \widehat{J}_j(1) - \widehat{K}_j(1) \right| \phi_i(1) \right\rangle \\ &= \left\langle \phi_j(2) \left| \widehat{J}_i(2) - \widehat{K}_i(2) \right| \phi_j(2) \right\rangle, \end{aligned}$$

one obtains the integrals that can be denoted by

$$\begin{aligned} J_{ij} &= \left\langle \phi_i(1) \left| \widehat{J}_j(1) \right| \phi_i(1) \right\rangle, \\ K_{ij} &= \left\langle \phi_i(1) \left| \widehat{K}_j(1) \right| \phi_i(1) \right\rangle. \end{aligned}$$

With this notation, the expression (Eq. 106) takes the following form:

$$\left\langle \widehat{H}(1, 2) \right\rangle_{\Psi(1,2)} = \left\langle \phi_i(1) \left| \widehat{f}(1) \right| \phi_i(1) \right\rangle + \left\langle \phi_j(2) \left| \widehat{f}(2) \right| \phi_j(2) \right\rangle - (J_{ij} - K_{ij}), \quad (127)$$

or if one wants to apply the spin-orbital energies ϵ calculated earlier:

$$\left\langle \widehat{H}(1, 2) \right\rangle_{\Psi(1,2)} = \epsilon_i + \epsilon_j - (J_{ij} - K_{ij}). \quad (128)$$

Now it is time to generalize these considerations into the N -electron case. The recipe for this transformation was given previously (Eqs. 91, 92, and 93). For the system of N electrons, the set of N -coupled equations of the form

$$\widehat{f}_i(1)\phi_i(1) = \epsilon_i\phi_i(1), \quad \text{for } i = 1, \dots, N \quad (129)$$

must be solved. Here,

$$f_i(1) = h(1) + \sum_{j \neq i}^N \left(\widehat{J}_j(1) - \widehat{K}_j(1) \right). \quad (130)$$

The summation in the expression for the one-electron Fock operator arises from the fact that now the given electron described by the spin-orbital i interacts with $(N - 1)$ electrons in the states determined by the remaining spin-orbitals.

In searching for the ground state energy, one is interested in the lowest possible energy. Therefore, the functions of choice are the spin-orbitals corresponding to the lowest values of ϵ . Such a set of spin-orbitals, called occupied, is opposite to any other solutions of the Fock equations corresponding to higher energies. These are known as virtual (unoccupied) orbitals.

Finally, the average value of the Hamiltonian can be written as

$$\left\langle \widehat{H}(1, 2, \dots, N) \right\rangle_{\Psi^{(0)}(N)} = \sum_{i=1}^N \epsilon_i - \sum_{i=1}^N \sum_{j>i}^N (J_{ij} - K_{ij}) = E_0^{\text{HF}}$$

It should be emphasized that the energy estimated in this manner is not the simple sum of the orbital energies. If the terms $\sum_{i=1}^N \sum_{j>i}^N (J_{ij} - K_{ij})$ are neglected, the double summation of the electron–electron interaction would take place (Atkins and Friedman 2005; Cramer 2004; Jensen 2006; Levine 2008; Lowe and Peterson 2005; McQuarrie and Simon 1997; Piela 2007; Ratner and Schatz 2000; Roos and Widmark 2002; Szabo and Ostlund 1996).

Møller–Plesset Perturbation Theory: HF Is Just the Beginning

From here forward, we will treat the Hartree–Fock function as the basis for the further investigations and denote it as $\Psi_0^{(0)}$, where the subscript 0 indicates the ground state and the superscript (0) is the reference function. We will also omit the explicit writing of the dependence of the Hamiltonian and the wave function on the coordinates of N electrons. As a consequence,

$$\left\langle \Psi_0^{(0)} \left| \widehat{H} \right| \Psi_0^{(0)} \right\rangle = E_0^{\text{HF}}. \quad (131)$$

One of the possible ways of improving the HF results is the application of the perturbation theory (Atkins and Friedman 2005; Cramer 2004; Jensen 2006; Levine 2008; McQuarrie and Simon 1997; Piela 2007; Ratner and Schatz 2000; Szabo and Ostlund 1996). The Hamiltonian (Eq. 89) can be partitioned into the unperturbed part and the perturbation in the following manner:

$$\widehat{H}^0 = \sum_{i=1}^N \widehat{f}(i), \quad \widehat{H}^1 = \sum_{i=1}^N \left(\sum_{j>i}^N \frac{1}{r_{ij}} - \widehat{v}^{\text{HF}}(i) \right) \quad (132)$$

(compare Eq. 125). Up to the first order in the perturbation theory, the energy is

$$E_0^{\text{HF}} = E_0^{(0)} + E_0^{(1)} = \left\langle \Psi_0^{(0)} \left| \widehat{H}^0 + \widehat{H}^1 \right| \Psi_0^{(0)} \right\rangle = \left\langle \widehat{H} \right\rangle_{\Psi_0^{(0)}}. \quad (133)$$

The correction to the HF energy appears in the second order:

$$E_0^{(2)} = \sum_{k \neq 0}^{\infty} \frac{\left| \langle \Psi_0^{(0)} | \widehat{H}^1 | \Psi_k^{(0)} \rangle \right|^2}{E_0^{(0)} - E_k^{(0)}}. \quad (134)$$

During calculations, the expansion of the spin-orbitals in the finite basis set is applied. This allows identification of only the finite number of spin-orbitals. But the number of all possible $\Psi_k^{(0)}$ functions can still be horrifyingly large. In practice, this is equivalent to the finite but long expansion in the above summation. In quantum chemistry we are interested in the best quality results with moderate expenses and are continuously searching for more economical methods.

Careful examination of the second-order energy correction $E_0^{(2)}$ shows that a significant number of its terms do not contribute to the final result. Some work invested in the manipulation of expressions allows one not only to learn the basic computational apparatus but also to save a lot of effort.

In order to proceed comfortably, the notation should again be simplified. Let us drop the superscript (0) denoting the zeroth-order functions (other functions will not appear in our considerations). Additionally, we omit the subscript k from the determinants containing the virtual spin-orbitals. In return, we explicitly specify the pattern of spin-orbital exchange. Using Eq. 108, the Slater determinant can be written as

$$\left| \Psi_0 \right\rangle = |\phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4) \dots \phi_N(N)\rangle. \quad (135)$$

The numbers of the occupied spin-orbitals vary from 1 to N . Thus, the virtual spin-orbitals will be labeled starting from $N + 1$ onward. Now consider the example of the determinant in which the occupied spin-orbital ϕ_3 is exchanged for the virtual one, ϕ_{N+8} . The new determinant can be written as

$$\left| \Psi_3^{N+8} \right\rangle = |\phi_1(1)\phi_2(2)\phi_{N+8}(3)\phi_4(4) \dots \phi_N(N)\rangle, \quad (136)$$

where the subscript in Ψ_3^{N+8} denotes the occupied orbital that is exchanged and the superscript denotes the virtual one that takes its place. To generalize, one could denote the occupied orbitals building the Ψ_0 function by first alphabet letters a, b, c, d, \dots and the virtuals by p, q, r, s, \dots . Therefore, Ψ_3^{N+8} can be written as Ψ_a^p , where $a = 3$ and $p = N + 8$. Likewise, any determinant can be represented.

Thereafter, the influence of the function choice on the values of the integrals appearing in the electronic energy calculations can be investigated. Similarly, as in the case of Hamiltonian, the integrals can also be divided into two groups with one-electron operator

$$\widehat{o}_1 = \sum_{i=1}^N \widehat{o}(i), \quad (137)$$

where in place of $\hat{o}(i)$ operators $\hat{h}, (i)$ or $\hat{f}(i)$ will be used, and with two-electron operator,

$$\hat{O}_2 = \sum_i \sum_{j>i} \left(\frac{1}{r_{ij}} \right). \quad (138)$$

To get a full picture, we should analyze the following types of integrals:

$$\langle \Psi_0 | \hat{o}_1 | \Psi_a^p \rangle \quad \text{and} \quad \langle \Psi_0 | \hat{O}_2 | \Psi_a^p \rangle, \quad (139)$$

$$\langle \Psi_0 | \hat{o}_1 | \Psi_{ab}^{pq} \rangle \quad \text{and} \quad \langle \Psi_0 | \hat{O}_2 | \Psi_{ab}^{pq} \rangle, \quad (140)$$

$$\langle \Psi_0 | \hat{o}_1 | \Psi_{abc}^{pqr} \rangle \quad \text{and} \quad \langle \Psi_0 | \hat{O}_2 | \Psi_{abc}^{pqr} \rangle. \quad (141)$$

Any other will be equal to zero (Levine 2008; Szabo and Ostlund 1996).

Recall that the determinant of the $N \times N$ matrix can be represented as the sum of the $N!$ products of the matrix elements. According to Laplace's formula, a determinant can be expanded along a row or a column. Thus, calculation of the $\langle \Psi_0 | \hat{o}_1 | \Psi_a^p \rangle$ integrals, when the occupied orbital a in Ψ_0 has been exchanged with the virtual orbital p in Ψ_a^p , can be performed by expanding the determinant Ψ_0 along the a -th row and the determinant Ψ_a^p along the p -th row:

$$\Psi_0 = \frac{1}{\sqrt{N!}} \sum_{i=1}^N \phi_a(i) C_{ai}, \quad (142)$$

$$\Psi_a^p = \frac{1}{\sqrt{N!}} \sum_{i=1}^N \phi_p(i) C_{pi}. \quad (143)$$

The cofactors C can be perceived as the $(N - 1)$ -electron determinants that have been obtained from Ψ_0 and Ψ_a^p via the elimination of the a -th and p -th spin-orbitals, respectively. Thus, after elimination of what is different in the two determinants, we get both cofactors equal to each other. Hence,

$$\begin{aligned} \langle \Psi_0 | \hat{o}_1 | \Psi_a^p \rangle &= \frac{1}{N!} \left\langle \sum_{i=1}^N \phi_a(i) C_{ai} \left| \sum_{j=1}^N \hat{o}(j) \right| \sum_{k=1}^N \phi_p(k) C_{pi} \right\rangle \\ &= \frac{1}{N!} \sum_{i=1}^N \langle \phi_a(i) | \hat{o}(i) | \phi_p(i) \rangle \langle C_{ai} | C_{pi} \rangle \\ &= \frac{1}{N} \sum_{i=1}^N \langle \phi_p(i) | \hat{o}(i) | \phi_p(i) \rangle \\ &= \langle \phi_a(1) | \hat{o}(1) | \phi_p(1) \rangle, \end{aligned} \quad (144)$$

where

$$\langle C_{ai} | C_{pi} \rangle = (N - 1)!, \quad (145)$$

$$\sum_{i=1}^N \langle \phi_a(i) | \widehat{\partial}(i) | \phi_a(i) \rangle = N \langle \phi_a(1) | \widehat{\partial}(1) | \phi_a(1) \rangle \quad (146)$$

was applied. If the cofactors are not the same – this would be the case when the determinants differ from two or more spin-orbitals – the corresponding overlap matrix is equal to zero according to the orthogonality condition. Thus,

$$\begin{aligned} \langle \Psi_0 | \widehat{\partial}_1 | \Psi_a^p \rangle &= \langle \phi_a(1) | \widehat{\partial}(1) | \phi_p(1) \rangle, \\ \langle \Psi_0 | \widehat{\partial}_1 | \Psi_{ab}^{pq} \rangle &= 0, \\ \langle \Psi_0 | \widehat{\partial}_1 | \Psi_{abc}^{pqr} \rangle &= 0. \end{aligned} \quad (147)$$

Similar (but a little more time-consuming) considerations for the two-electron operators lead to the following expressions:

$$\begin{aligned} \langle \Psi_0 | \widehat{\mathcal{O}}_2 | \Psi_a^p \rangle &= \sum_{i=1}^N \left(\left\langle \phi_a(1) \phi_i(2) \left| \frac{1}{r_{12}} \right| \phi_p(1) \phi_i(2) \right\rangle \right. \\ &\quad \left. - \left\langle \phi_a(1) \phi_i(2) \left| \frac{1}{r_{12}} \right| \phi_i(1) \phi_p(2) \right\rangle \right) \\ \langle \Psi_0 | \widehat{\mathcal{O}}_2 | \Psi_{ab}^{pq} \rangle &= \left\langle \phi_a(1) \phi_b(2) \left| \frac{1}{r_{12}} \right| \phi_p(1) \phi_q(2) \right\rangle \\ &\quad - \left\langle \phi_a(1) \phi_b(2) \left| \frac{1}{r_{12}} \right| \phi_q(1) \phi_p(2) \right\rangle \\ \langle \Psi_0 | \widehat{\mathcal{O}}_2 | \Psi_{abc}^{pqr} \rangle &= 0. \end{aligned} \quad (148)$$

In the first contact with these equations, one can have the feeling that something is lost here. We start from the integrals with the N -electron functions and we finish with the integral of only the electrons labeled by 1 and 2. What happened to the rest? Again, it should be emphasized here that the electron labels symbolize only the integration variables. What matters is the functions of these variables, namely, spin-orbitals. Thus, in the integral with the one-electron operator, the electron label 1 means that the integration is performed only over the variables of one electron. Likewise, the two-electron operator integral depends on the variables of two electrons, which is symbolized by two labels: 1 and 2.

In the integration of one-electron expressions, the electron label can be omitted without any harm: $\langle \phi_a(1) | \widehat{\partial}_1(1) | \phi_a(1) \rangle$. Likewise, in the two-electron case, we can declare that the spin-orbitals are written in the given order. So,

$$\left\langle \phi_x(1) \phi_y(2) \left| \frac{1}{r_{12}} \right| \phi_v(1) \phi_z(2) \right\rangle = \left\langle \phi_x \phi_y \left| \frac{1}{r_{12}} \right| \phi_v \phi_z \right\rangle. \quad (149)$$

Moreover, now there is no reason to explicitly write the ϕ symbol. Thereby, the next simplification of the notation is obvious:

$$\begin{aligned}\langle \phi_a | \widehat{o}_1 | \phi_a \rangle &= \langle a | \widehat{o}_1 | a \rangle, \\ \langle \phi_x \phi_y | \frac{1}{r_{12}} | \phi_v \phi_z \rangle &= \langle xy | \frac{1}{r_{12}} | vz \rangle,\end{aligned}\tag{150}$$

and finally, since the combination of the integrals $\langle xy | 1/r_{12} | vz \rangle - \langle xy | 1/r_{12} | zv \rangle$ appears frequently, the following symbol is introduced:

$$\langle xy || zv \rangle = \langle xy | \frac{1}{r_{12}} | vz \rangle - \langle xy | \frac{1}{r_{12}} | zv \rangle.\tag{151}$$

Now Eqs. 147 and 148 can be rewritten as

$$\begin{aligned}\langle \Psi_0 | \widehat{o}_1 | \Psi_a^p \rangle &= \langle a | \widehat{o}_1 | p \rangle, \\ \langle \Psi_0 | \widehat{o}_1 | \Psi_{ab}^{pq} \rangle &= 0, \\ \langle \Psi_0 | \widehat{o}_1 | \Psi_{abc}^{pqr} \rangle &= 0, \\ \langle \Psi_0 | \widehat{O}_2 | \Psi_a^p \rangle &= \sum_{i=1}^N \langle ai || pi \rangle, \\ \langle \Psi_0 | \widehat{O}_2 | \Psi_{ab}^{pq} \rangle &= \langle ab || pq \rangle, \\ \langle \Psi_0 | \widehat{O}_2 | \Psi_{abc}^{pqr} \rangle &= 0.\end{aligned}\tag{152}$$

These simple equations, known as the Slater rules, allow for the following general remark: If the one-electron operator is integrated with functions that differ by more than one spin-orbital, the corresponding integral vanishes. Similarly, the result is zero for the integration of two-electron operators with functions differing by more than two spin-orbitals. Hitherto, only the integrals with Ψ_0 were considered. However, it is easy to notice that functions Ψ_{ab}^{pq} and Ψ_{abc}^{pqr} differ by only one exchange (spin-orbital $\phi_c \rightarrow \phi_r$), etc. Therefore, the above considerations can be also applied in any other cases.

In this abundance of equations, our main goal cannot be lost: All these derivations were necessary to limit the types of the Ψ_k functions present in the MP2 energy expression. Now, with a recognition of the above Slater rules, one can safely neglect the integrals with the pairs of the Ψ_k functions differing with more than two spin-orbital exchanges. But this is not everything.

Let us consider the integral

$$\langle \Psi_0 | \widehat{H} | \Psi_a^p \rangle = \langle a | \widehat{h} | p \rangle + \sum_j \left(\langle aj | \frac{1}{r_{ij}} | pj \rangle - \langle aj | \frac{1}{r_{ij}} | jp \rangle \right).\tag{153}$$

Since

$$\left\langle aj \left| \frac{1}{r_{ij}} \right| pj \right\rangle = \left\langle a \left| \widehat{J}_j \right| p \right\rangle, \quad (154)$$

and

$$\left\langle aj \left| \frac{1}{r_{ij}} \right| jp \right\rangle = \left\langle a \left| \widehat{K}_j \right| p \right\rangle, \quad (155)$$

one gets

$$\left\langle \Psi_0 \left| \widehat{H} \right| \Psi_a^p \right\rangle = \left\langle a \left| \widehat{H}^0 \right| p \right\rangle, \quad (156)$$

which has to vanish, because

$$\left\langle a \left| \widehat{H}^0 \right| p \right\rangle = \left\langle a \left| \sum_{i=1}^N f(i) \right| p \right\rangle = \langle a | f | p \rangle = \epsilon_p \delta_{ap} = 0, \quad (157)$$

for it was assumed that $a \neq p$.

The obtained result,

$$\left\langle \Psi_0 \left| \widehat{H} \right| \Psi_a^p \right\rangle = 0, \quad (158)$$

is known as the Brillouin theorem. Applying it allows one also to show that

$$\left\langle \Psi_0 \left| \widehat{H}^1 \right| \Psi_a^p \right\rangle = 0. \quad (159)$$

Indeed, using

$$\widehat{H}^1 = \widehat{H} - \widehat{H}^0, \quad (160)$$

one obtains

$$\begin{aligned} \left\langle \Psi_0 \left| \widehat{H}^1 \right| \Psi_a^p \right\rangle &= \left\langle \Psi_0 \left| \widehat{H} \right| \Psi_a^p \right\rangle - \left\langle \Psi_0 \left| \widehat{H}^0 \right| \Psi_a^p \right\rangle \\ &= 0 - \left(\sum_{i=1}^N \epsilon_i \right) \left\langle \Psi_0 \left| \Psi_a^p \right\rangle \right. \\ &= 0. \end{aligned} \quad (161)$$

Now is the time for an important conclusion: In order to calculate the correction to the energy in MP2, the functions arising from Ψ_0 only by the exchange of precisely two spin-orbitals need to be applied, not more, not less. Life becomes easier when not calculating zero contributions in a complicated way.

Let us exploit the above knowledge to transform Eq. 134. The integral on the right-hand side will be calculated with the functions Ψ_{ab}^{pq} . Assuming $b > a$ and $q > p$ allows avoidance of the integration with the same functions. The upper limit for

the summations over a and b will be equal to N . For the remaining spin-orbitals, it should be ∞ ; however, in practice, the finite basis is applied and the upper limit will be determined by the basis set size. Let us look into the numerator of Eq. 134 carefully. Denoting the total Fock operator as $\widehat{F} = \sum_i \widehat{f}(i)$, one gets

$$\begin{aligned} \langle \Psi_0 | \widehat{H}^1 | \Psi_{ab}^{pq} \rangle &= \langle \Psi_0 | \widehat{H} | \Psi_{ab}^{pq} \rangle - \langle \Psi_0 | \widehat{F} | \Psi_{ab}^{pq} \rangle \\ &= \langle ab | \frac{1}{r_{ij}} | pq \rangle - \langle ab | \frac{1}{r_{ij}} | pq \rangle, \end{aligned} \quad (162)$$

since

$$\langle \Psi_0 | \widehat{F} | \Psi_{ab}^{pq} \rangle = 0 \quad (163)$$

(integration of the one-electron operator with the functions differing by two spin-orbitals; see Eq. 152).

Now consider the denominator of Eq. 134. The energy of the ground state is simply a sum of N lowest spin-orbital energies:

$$E_0^{(0)} = \sum_{i=1}^N \epsilon_i. \quad (164)$$

The energy of the zeroth order for the function Ψ_{ab}^{pq} is a similar sum with ϵ_a and ϵ_b replaced by ϵ_p and ϵ_q :

$$E_{(pq)}^{(0)} = E_0^{(0)} - \epsilon_a - \epsilon_b + \epsilon_p + \epsilon_q. \quad (165)$$

Thus, the final form of Eq. 134 is

$$E_0^{(2)} = \sum_{b>a} \sum_{q>p} \frac{|\langle ab || pq \rangle|^2}{\epsilon_p + \epsilon_q - \epsilon_a - \epsilon_b}. \quad (166)$$

Why was it worth our hard work? Not only for satisfaction. These derivations are necessary to understand the mechanisms employed in computational methods of quantum chemistry. With the Slater rules, it is straightforward to recognize the vanishing integrals. The double exchanges appear to be most important, since these are the only exchanges that give rise to the energy corrections in the MP2 method.

According to the perturbation calculus, the wave function corrected in the first order can be written as

$$\Psi \approx \Psi_0 + \sum_{k \neq 0} c_k \Psi_k, \quad (167)$$

where c_k is given by

$$c_k = \frac{\langle \Psi_0 | \widehat{H}^1 | \Psi_k \rangle}{E_0 - E_k}. \quad (168)$$

From the Slater rules, it is easy to estimate that the only nonvanishing terms will arise from the double spin-orbital exchange in the wave function. The intuitive statement is that the low-order corrections should have the larger impact and the above considerations lead to the most important accomplishment of this section: The largest contribution to the corrections of the Hartree–Fock function arises from the functions with the doubly exchanged spin-orbitals.

Beyond the HF Wave Function

Having done all this hard work, one can now sit comfortably in an armchair and think. The main goal of the quantum chemistry is to find the best possible description of the state of the system (the best possible wave function). The exact solutions of the Hamiltonian eigenproblem are unavailable and all we have are approximations. We have become used to approximations in everyday life. The important thing is to realize that the Hartree–Fock solutions can be improved. At the beginning of this chapter, various properties of operators were discussed. Among others, it was stated that the eigenfunctions of the Hermitian operators constitute the complete sets and any other function of the same variables can be represented by applying them. The one-electron spin-orbitals that are the eigenfunctions of the Fock operator are accessible. They form the complete set, but only for the one-electron functions. However, they can be applied to build up the N -electron determinants. The set of all possible determinants is also complete and, therefore, can be applied to express any N -electron function (Cramer 2004; Jensen 2006; Levine 2008; Lowe and Peterson 2005; Piela 2007; Ratner and Schatz 2000; Roos and Widmark 2002; Szabo and Ostlund 1996):

$$\Psi = c_0 \Psi_0 + \sum_{a,p} c_a^p \Psi_a^p + \sum_{a,b,p,q} c_{ab}^{pq} \Psi_{ab}^{pq} + \sum_{a,b,c,p,q,r} c_{abc}^{pqr} \Psi_{abc}^{pqr} + \dots \quad (169)$$

For this purpose, only the coefficients $c_0, c_a^p, c_{ab}^{pq}, c_{abc}^{pqr}, \dots$ need to be found. In the ideal case, all the summations would be infinite and the problem must be reduced. Still, instead of using the infinite expansions, the finite and relatively small number of terms can be sufficient. Moreover, solving the Hartree–Fock equations in the finite basis, one possesses only the finite number of orbitals that can be exchanged.

Anyway, it is instructive to see how large the number of terms in Eq. 169 can be. Consider the methane molecule CH_4 . Calculations with the minimal basis set (each orbital described by a single one-electron function; for carbon single functions for each of the orbitals, 1 s , 2 s , $2p_x$, $2p_y$, $2p_z$, and for hydrogen a single function for 1 s orbital) require 9 orbitals/18 spin-orbitals for the 10-electron system. From the probability theory, the number of combinations (K) of k elements from the n -element set can be calculated as

$$K = \binom{n}{k} = \frac{n!}{(n-k)!k!}. \quad (170)$$

In the case of methane, 10 electrons can be placed in 18 spin-orbitals on $\binom{18}{10} = 43,758$ ways. This is equivalent to the 43,758 terms in the expansion (Eq. 169). Impressive. And one needs to remember that the minimal basis set gives relatively bad Hartree–Fock solutions and is not recommended in *ab initio* calculations. However, increasing the basis set size causes the number of expansion terms to grow dramatically. For instance, in the case of the so-called double- ζ basis set (two functions per each orbital) for methane, one has 36 spin-orbitals, which makes 254,186,856 combinations! And double- ζ is still not much . . .

Therefore, it is necessary to find some way to reduce the size of the problem. The symmetry of the molecules can be applied here, and the fact that the chosen determinants (or their linear combinations) must be the given functions of the spin operators can be beneficial. Moreover, one would like to eliminate from the expansion the determinants that are not crucial for the quality of the wave function, and their neglect does not cause the deterioration of the description of the system (or causes only slight deterioration). In other words, only the determinants that have the significant contribution to the total energy must be chosen for the wave function construction. Let us begin with the classification of the determinants, taking into account the number of the spin-orbitals exchanged with respect to Ψ_0 . For this purpose, the averaged value of Hamiltonian calculated with Ψ will be useful. We can write the wave function expansion as

$$\Psi = c_0\Psi_0 + \mathbf{S}\mathbf{c}_S + \mathbf{D}\mathbf{c}_D + \mathbf{T}\mathbf{c}_T + \mathbf{Q}\mathbf{c}_Q + \dots \quad (171)$$

The symbols' meaning can be clearly deciphered by comparison with Eq. 169. \mathbf{S} denotes a vector build of the determinants constructed from Ψ_0 by single exchanges, and \mathbf{c}_S is a vector of coefficients corresponding to the functions in \mathbf{S} :

$$\mathbf{S}\mathbf{c}_S = \sum_a \sum_q c_a^q \Psi_a^q. \quad (172)$$

In other words, \mathbf{S} contains all the functions with the single exchanged spin-orbital. Similarly, \mathbf{D} would be the combination of the functions with double exchanges, \mathbf{T} with triple exchanges, and so forth. With such a notation, the function Ψ can be treated as the scalar product of the basis vectors $\Psi_0, \mathbf{S}, \mathbf{D}, \dots$, and the coefficient vector

$$|\Psi\rangle = [|\Psi_0\rangle, |\mathbf{S}\rangle, |\mathbf{D}\rangle, |\mathbf{T}\rangle, |\mathbf{Q}\rangle, \dots] \cdot \begin{bmatrix} c_0 \\ \mathbf{c}_S \\ \mathbf{c}_D \\ \mathbf{c}_T \\ \mathbf{c}_Q \\ \vdots \end{bmatrix}. \quad (173)$$

Using this notation, the Hamiltonian \widehat{H} of the system can be linked in an elegant way to a matrix \mathbf{H} . Let us apply this form of the wave function for the calculation of the Hamiltonian average value. To simplify the expressions, let us limit ourselves to the truncated expansion:

$$\Psi_{\text{SD}} = [\Psi_0 \mathbf{S} \mathbf{D}] \begin{bmatrix} c_0 \\ \mathbf{c}_S \\ \mathbf{c}_D \end{bmatrix} = c_0 \Psi_0 + \mathbf{S} \mathbf{c}_S + \mathbf{D} \mathbf{c}_D. \quad (174)$$

The average value of the Hamiltonian can now be written as

$$\langle H \rangle_{\Psi_{\text{SD}}} = \langle \Psi_{\text{SD}} | \widehat{H} | \Psi_{\text{SD}} \rangle = [c_0 \mathbf{c}_S \mathbf{c}_D]^\dagger \begin{bmatrix} \langle \Psi_0 | \\ \langle \mathbf{S} | \\ \langle \mathbf{D} | \end{bmatrix} \widehat{H} \begin{bmatrix} | \Psi_0 \rangle, | \mathbf{S} \rangle, | \mathbf{D} \rangle \end{bmatrix} \begin{bmatrix} c_0 \\ \mathbf{c}_S \\ \mathbf{c}_D \end{bmatrix}. \quad (175)$$

The vector multiplication leads to the following expression:

$$\begin{aligned} \langle H \rangle_{\Psi_{\text{SD}}} &= c_0^* \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle c_0 + c_0^* \langle \Psi_0 | \widehat{H} | \mathbf{S} \rangle \mathbf{c}_S + c_0^* \langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle \mathbf{c}_D \\ &\quad + \mathbf{c}_S^\dagger \langle \mathbf{S} | \widehat{H} | \Psi_0 \rangle c_0 + \mathbf{c}_S^\dagger \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle \mathbf{c}_S + \mathbf{c}_S^\dagger \langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle \mathbf{c}_D \\ &\quad + \mathbf{c}_D^\dagger \langle \mathbf{D} | \widehat{H} | \Psi_0 \rangle c_0 + \mathbf{c}_D^\dagger \langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle \mathbf{c}_S + \mathbf{c}_D^\dagger \langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle \mathbf{c}_D. \end{aligned} \quad (176)$$

Such an equation is not very useful, since we still do not know the c_0 , \mathbf{c}_S and \mathbf{c}_D coefficients determining the Ψ_{SD} function. The only thing that can be said about them so far comes from the normalization requirement for Ψ_{SD} :

$$1 = c_0^* c_0 + \mathbf{c}_S^\dagger \mathbf{c}_S + \mathbf{c}_D^\dagger \mathbf{c}_D. \quad (177)$$

This is not enough to uniquely determine the wave function. However, going back to Eq. 175 and multiplying only the inside vectors, one obtains

$$\mathbf{H}_{\text{SD}} = \begin{bmatrix} \langle \Psi_0 | \\ \langle \mathbf{S} | \\ \langle \mathbf{D} | \end{bmatrix} \widehat{H} \begin{bmatrix} | \Psi_0 \rangle, | \mathbf{S} \rangle, | \mathbf{D} \rangle \end{bmatrix} = \begin{bmatrix} \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle & \langle \Psi_0 | \widehat{H} | \mathbf{S} \rangle & \langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle \\ \langle \mathbf{S} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle \\ \langle \mathbf{D} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle \end{bmatrix}. \quad (178)$$

We can then associate finding the approximate Hamiltonian eigenvalues with its matrix in the Ψ_{SD} basis:

$$\langle \mathbf{H} \rangle_{\Psi_{\text{SD}}} = [c_0^* \mathbf{c}_S^\dagger \mathbf{c}_D^\dagger] = \begin{bmatrix} \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle & \langle \Psi_0 | \widehat{H} | \mathbf{S} \rangle & \langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle \\ \langle \mathbf{S} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle \\ \langle \mathbf{D} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle \end{bmatrix} \begin{bmatrix} c_0 \\ \mathbf{c}_S \\ \mathbf{c}_D \end{bmatrix}. \quad (179)$$

Because of the Hermitian character of the Hamilton operator, the \mathbf{H}_{SD} matrix is symmetric and real. Its diagonalization provides the set of the eigenvalues corresponding to its eigenvectors. We are interested in the ground-state energy and thus, we need only the lowest eigenvalue of the \mathbf{H}_{SD} matrix and the respective normalized eigenvector Ψ_{SD} .

Knowing the procedure for the finite basis (only single- and double-orbital exchanges), we can see how it looks for the full (Eq. 171) expansion. The matrix notation leads to the average value of the Hamiltonian, written as

$$\langle H \rangle_\Psi = [c_0, \mathbf{c}_S, \mathbf{c}_D, \mathbf{c}_T, \mathbf{c}_Q] \begin{bmatrix} \langle \Psi_0 | \\ \langle \mathbf{S} | \\ \langle \mathbf{D} | \\ \langle \mathbf{T} | \\ \langle \mathbf{Q} | \\ \vdots \end{bmatrix} \widehat{H} [\Psi_0 | \mathbf{S} \rangle, | \mathbf{D} \rangle, | \mathbf{T} \rangle, | \mathbf{Q} \rangle, \dots] \begin{bmatrix} c_0 \\ \mathbf{c}_S \\ \mathbf{c}_D \\ \mathbf{c}_T \\ \mathbf{c}_Q \\ \vdots \end{bmatrix}. \quad (180)$$

The vector multiplication permits one to perceive the average value as the eigenproblem of the Hamiltonian matrix:

$$\mathbf{H} = \begin{bmatrix} \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle & 0 & \langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle & 0 & 0 & \dots \\ 0 & \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{T} \rangle & 0 & \dots \\ \langle \mathbf{D} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{T} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{Q} \rangle & \dots \\ 0 & \langle \mathbf{T} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{T} | \widehat{H} | \mathbf{D} \rangle & \langle \mathbf{T} | \widehat{H} | \mathbf{T} \rangle & \langle \mathbf{T} | \widehat{H} | \mathbf{Q} \rangle & \dots \\ 0 & 0 & \langle \mathbf{Q} | \widehat{H} | \mathbf{D} \rangle & \langle \mathbf{Q} | \widehat{H} | \mathbf{T} \rangle & \langle \mathbf{Q} | \widehat{H} | \mathbf{Q} \rangle & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \quad (181)$$

It can be clearly seen that some blocks in this matrix are equal to zero. This happens in two cases:

- The integrals between Ψ_0 and functions of the \mathbf{S} type (single exchange of spin-orbitals) vanish due to the Brillouin theorem, as was shown in the previous section.

- The integrals between the functions that differ by more than two exchanges, for instance, **S** and **Q** type, vanish due to the Slater rules.

Even not knowing combinatorics, one can expect that the number of functions in a block will grow drastically with and increase in the number of exchanges (block **S** will contain less functions than **D** etc.). A bit of thinking in the beginning would help to save a lot of time by not calculating zero integrals. Let's see: Only in the case of $\langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle$ and $\langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle$ blocks should all the elements be calculated. The remaining matrices are sparse. For example, the $\langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle$ block contains the integrals of the following types: $\langle \Psi_{ab}^{pq} | \widehat{H} | \Psi_{ab}^{pq} \rangle$ (the same function on both sides), $\langle \Psi_{ab}^{pq} | \widehat{H} | \Psi_{ab}^{pr} \rangle$ (differing by one exchange), $\langle \Psi_{ab}^{pq} | \widehat{H} | \Psi_{ab}^{rs} \rangle$ (differing by two exchanges), $\langle \Psi_{ab}^{pq} | \widehat{H} | \Psi_{ac}^{pq} \rangle$ (differing by three exchanges), and $\langle \Psi_{ab}^{pq} | \widehat{H} | \Psi_{dc}^{pq} \rangle$ (differing by four exchanges). Obviously, the two latter cases produce zeros.

With the large number of exchanges, the size of the blocks grows abruptly, but most of the elements would be equal to zero. The simplification in this case would be the limitation of the **H** matrix size by the elimination of the functions including more than a given number of exchanges from the expansion. Let us leave only the single exchange block. Thus, the Hamiltonian matrix has the form

$$\mathbf{H}_s = \begin{bmatrix} \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle & 0 \\ 0 & \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle \end{bmatrix}. \quad (182)$$

This is the block diagonal matrix. One of the properties of such matrices is that their eigenvalue set is a sum of the eigenvalues of the diagonal blocks. This means that the lowest possible eigenvalue is $E_0 = \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle$, and in consequence, there is no improvement in the ground-state energy with respect to the Hartree–Fock theory when taking only the single orbital exchanges.

Hence, let us also include the functions of the **D** type. Now the Hamilton matrix can be written as

$$\mathbf{H}_{SD} = \begin{bmatrix} \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle & 0 & \langle \Psi_0 | \widehat{H} | \mathbf{D} \rangle \\ 0 & \langle \mathbf{S} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle \\ \langle \mathbf{D} | \widehat{H} | \Psi_0 \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle & \langle \mathbf{D} | \widehat{H} | \mathbf{D} \rangle \end{bmatrix}. \quad (183)$$

It is no longer a block diagonal matrix – all blocks contribute to its eigenvalues and one can count on some improvement. An interesting observation, however, is that here the functions with the single spin-orbital exchange also have influence on the energy via the $\langle \mathbf{S} | \widehat{H} | \mathbf{D} \rangle$ and $\langle \mathbf{D} | \widehat{H} | \mathbf{S} \rangle$ blocks.

Next, subsequent groups of functions can be applied containing more than two spin-orbital exchanges. However, the calculations become prohibitively expensive, even for moderate size of the systems, and the consecutive corrections are smaller and smaller. The distinguished character of the double spin-orbital exchange was already discussed within the MP2 method. Now one can also expect that including double exchanges produces reasonable results with the moderate computational cost. Then, why not save more and diagonalize only

$$\mathbf{H}_D = \begin{bmatrix} \langle \Psi_0 | \hat{H} | \Psi_0 \rangle & \langle \Psi_0 | \hat{H} | \mathbf{D} \rangle \\ \langle \mathbf{D} | \hat{H} | \Psi_0 \rangle & \langle \mathbf{D} | \hat{H} | \mathbf{D} \rangle \end{bmatrix} \quad (184)$$

instead of \mathbf{H}_{SD} ? This can be done; however, savings are not that great, since the number of \mathbf{S} functions is significantly smaller than the number of \mathbf{D} functions. Thus, if one can afford \mathbf{H}_D diagonalization, \mathbf{H}_{SD} diagonalization is probably also within easy reach.

The above reasoning has led to the sequence of quantum chemistry methods. The best results can be obtained within full CI (FCI) by applying the full expansion (Eq. 171) within the given basis set. This is certainly the most expensive variant. Cheaper – but also worse – are, respectively, CISD based on the \mathbf{H}_{SD} matrix and CID neglecting single exchanges (Cramer 2004; Jensen 2006; Levine 2008; Lowe and Peterson 2005; Piela 2007; Ratner and Schatz 2000; Szabo and Ostlund 1996).

So far, the reference function has been a single determinant. Such an approach is very limited. For instance, it does not allow one to describe a dissociation process. Correct description of dissociation requires at least one determinant for each subsystem. And, even in the cases when multi-determinant reference state description is not obligatory, such an elastic wave function will provide an improved description of the system of interest (Cramer 2004; Jensen 2006; Levine 2008; Piela 2007; Roos and Widmark 2002).

The multi-determinant wave function Ψ depends both on the expansion coefficients and on the spin-orbitals building up the determinants. Both these sets of variables can be optimized simultaneously. The particular case of this procedure, when taking only the first expansion term, is the Hartree–Fock approximation (SCF–HF). Therefore, the optimization of the multi-determinant wave function is called the multiconfiguration (MC) SCF method. Even without a detailed study of the MC–SCF equations, an improvement in the results with respect to the HF energy can be expected. However, this approach is much more expensive, since the spin-orbitals are optimized several times. Again, a time saving is desired. Therefore, let us search for the spin-orbitals with the highest influence on the total energy value. It has been observed that not all doubly exchanged functions provide the same contribution to the energy. Some improve the result more and others less. This is due to the spin-orbital energy differences. The exchange of the spin-orbitals of significantly different energies does not contribute much to the total energy improvement. Therefore, it can be requested that the exchange is included in calculations only if the energy difference between the involved spin-orbitals is

smaller than some given value. Hence, only some groups of spin-orbitals can be exchanged.

Up to now, the spin-orbital notion was used. However, let us switch to the orbital language that is frequently used for MC–SCF considerations.

For the N -electron system, the orbitals can be divided into three groups:

- Core orbitals, which are not varied, since they have too low orbital energies, but are applied in the wave function expansion (doubly occupied orbitals)
- Active orbitals, which are exchanged in the expansion (partially occupied orbitals)
- Virtual orbitals, which are not varied and not applied in the expansion (unoccupied orbitals)

Instead of optimization of all the orbitals, only the active orbitals will be varied within the complete active space self-consistent field (CASSCF) approximation. In the acronym of this method, the number of active orbitals and active electrons is also provided for the given system. For instance, CASSCF (6,4) denotes the calculations with the expansion including all the possible exchanges of the four electrons within the six active orbitals. The CASSCF approach leads to all possible exchanges in the given active space, and for a moderately sized system, the size of the active spaces can quickly exceed the computational resources. In such a case, the solution can be the restricted active space self-consistent field (RASSCF) method, which supplies a way of limiting the size of the active space.

Additionally, one needs to remember that for a powerful tool such as perturbation theory, there is no obstacle to applying the multi-determinant reference function as the unperturbed function in perturbation calculus. Thus, similar to the SCF–HF and MP2 approaches, CASPT2 would be the second-order perturbation theory complete active space method – the perturbationally corrected CASSCF.

Coupled Cluster Approximation: The Operator Strikes Back

It would seem that all the straightforward ways to improve wave function in the one-electron approximation have been exploited. However, we now next discuss one of the most accurate (and simultaneously most expensive) methods applied in quantum chemistry.

The idea is simple. Consider again the expansion (Eq. 171). Introducing an operator

$$\widehat{C} = \widehat{C}_0 + \widehat{C}_1 + \widehat{C}_2 + \widehat{C}_3 + \widehat{C}_4 + \dots, \quad (185)$$

defined as

$$\widehat{C}_0 |\Psi_0\rangle = c_0 |\Psi_0\rangle, \quad (186)$$

$$\widehat{C}_1 |\Psi_0\rangle = \mathbf{S}c_s, \quad (187)$$

$$\widehat{C}_2 |\Psi_0\rangle = \mathbf{Dc}_D, \quad (188)$$

$$\widehat{C}_3 |\Psi_0\rangle = \mathbf{Tc}_T, \quad (189)$$

$$\widehat{C}_4 |\Psi_0\rangle = \mathbf{Qc}_Q, \quad (190)$$

⋮

allows one to write Eq. 171 in a very compact form:

$$\Psi = \widehat{C} \Psi_0. \quad (191)$$

Now the problem of finding the appropriate expansion can be replaced by the problem of finding the adequate operator. This is the essence of the coupled cluster (CC) method. Here the assumption is made that the wave function can be expressed by

$$\Psi = e^{\widehat{T}} \Psi_0, \quad (192)$$

where Ψ_0 is a reference function (depending on the approach, this can be the one-determinant HF function or the multi-determinant function arising from MC-SCF) and \widehat{T} is a sought operator. Applying the expansion of the exponential function, it can be written that

$$e^{\widehat{T}} = \widehat{1} + \widehat{T} + \frac{1}{2!} \widehat{T}^2 + \frac{1}{3!} \widehat{T}^3 + \dots \quad (193)$$

Such an expanded form makes the interpretation of the \widehat{T} operators easier. Putting

$$\widehat{T} = \widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4 + \dots, \quad (194)$$

one can identify the subsequent \widehat{T}_i operators as corresponding to i -tuple exchanges of the spin-orbitals in the reference function:

$$\widehat{T}_1 \Psi_0 = \sum_{a,p} t_a^p \Psi_a^p, \quad (195)$$

$$\widehat{T}_2 \Psi_0 = \sum_{a,b,p,q} t_{ab}^{pq} \Psi_{ab}^{pq}, \quad (196)$$

and so forth. The coefficients t (called “amplitudes”) are in general not equivalent to the c coefficients in the CI expansion (see Eq. 171). In order to find their mutual relation, let us consider the approximate operator:

$$\widehat{T} \approx \widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4. \quad (197)$$

The operator (Eq. 193) takes the form

$$\begin{aligned}
& e^{\widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4} = \widehat{1} \\
& + \widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4 \\
& + \frac{1}{2!} \left(\widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4 \right)^2 \\
& + \frac{1}{3!} \left(\widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4 \right)^3 \\
& + \frac{1}{4!} \left(\widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4 \right)^4.
\end{aligned} \tag{198}$$

Limiting ourselves to the terms corresponding to not more than four spin-orbital exchanges and writing it in the ordered way according to the number of exchanges, one gets

$$\begin{aligned}
& e^{\widehat{T}_1 + \widehat{T}_2 + \widehat{T}_3 + \widehat{T}_4} = \widehat{1} \\
& + \widehat{T}_1 \\
& + \widehat{T}_2 + \frac{1}{2} \widehat{T}_1^2 \\
& + \widehat{T}_3 + \widehat{T}_1 \widehat{T}_2 + \frac{1}{3} \widehat{T}_1^3 \\
& + \widehat{T}_4 + \widehat{T}_1 \widehat{T}_3 + \frac{1}{2} \widehat{T}_2^2 + \frac{1}{2} \widehat{T}_1^2 \widehat{T}_2 + \widehat{T}_1^4.
\end{aligned} \tag{199}$$

Now the direct comparison can be made:

$$\widehat{C}_1 = \widehat{T}_1, \tag{200}$$

$$\widehat{C}_2 = \widehat{T}_2 + \frac{1}{2} \widehat{T}_1^2, \tag{201}$$

$$\widehat{C}_3 = \widehat{T}_3 + \widehat{T}_1 \widehat{T}_2 + \frac{1}{3} \widehat{T}_1^3, \tag{202}$$

$$\widehat{C}_4 = \widehat{T}_4 + \widehat{T}_1 \widehat{T}_3 + \frac{1}{2} \widehat{T}_2^2 + \frac{1}{2} \widehat{T}_1^2 \widehat{T}_2 + \widehat{T}_1^4. \tag{203}$$

We have the relation between the \widehat{C}_i and \widehat{T}_i operators, but still neither \widehat{C}_i nor \widehat{T}_i are known. Recall from the earlier sections that the double exchanges have a significant influence on the energy improvement with respect to the Hartree–Fock results. Taking double exchanges into account within the coupled cluster formalism means that the operators \widehat{T}_1 and \widehat{T}_2 need to be determined. However, as a side effect, they also allow inclusion of some not negligible contributions arising from the triple and higher exchanges. In the above comparison, the \widehat{T}_1 and \widehat{T}_2 operators recover two out of three terms in \widehat{C}_3 and three out of five terms in \widehat{C}_4 . This is the power of the CC method.

Unfortunately, the strength of this method does not go together with ease of calculations. Obtaining the expressions for the operators \widehat{T} is ransomed with

compromises. Not only is the operator expansion (Eq. 194) truncated, but the basis set is finite. Moreover, the variational character of the method is sacrificed.

In order to realize the complications, let us consider step-by-step the energy calculation within the CC formalism. We begin, as usual, with the electron Schrödinger Eq. 50. Substituting Eq. 192 gives

$$\widehat{H}e^{\widehat{T}}\Psi_0 = Ee^{\widehat{T}}\Psi_0. \quad (204)$$

Taking into account that, due to (Eq. 193),

$$\langle \Psi_0 | \Psi \rangle = \langle \Psi_0 | e^{\widehat{T}} | \Psi_0 \rangle = \langle \Psi_0 | \Psi_0 \rangle = 1, \quad (205)$$

the energy can be calculated as

$$E = \langle \Psi_0 | \widehat{H}e^{\widehat{T}} | \Psi_0 \rangle = \langle \Psi_0 | \widehat{H} | \Psi \rangle. \quad (206)$$

This is not the Hamiltonian average value expression. Additionally, the operator inside the bracket is not Hermitian. But, until we assume that (Eq. 192) is true, such an approach works. We can also construct an integral:

$$\langle \Psi_{ab}^{pq} | \widehat{H}e^{\widehat{T}} | \Psi_0 \rangle = E \langle \Psi_{ab}^{pq} | e^{\widehat{T}} | \Psi_0 \rangle, \quad (207)$$

which is the consequence of Eq. 204 and will be applied in the near future.

We should now concentrate on the way of determining the form of amplitudes. To simplify the considerations, we can assume

$$\widehat{T} \approx \widehat{T}_2, \quad (208)$$

which is equivalent to the CCD variant. We are interested in finding the amplitudes t_{ab}^{pq} . The final result of the calculations will be the approximate energy:

$$E_{\text{CCD}} = \langle \Psi_0 | \widehat{H}e^{\widehat{T}_2} | \Psi_0 \rangle. \quad (209)$$

The information about the amplitude t_{ab}^{pq} can be extracted from the integral

$$t_{ab}^{pq} = \langle \Psi_{ab}^{pq} | \widehat{T}_2 | \Psi_0 \rangle \quad (210)$$

(see Eq. 196). However, the amplitudes are still not known, since we do not know the \widehat{T}_2 operator. Therefore, one more equation is necessary to elicit the sought information. Let us begin with the approximated expression (Eq. 207):

$$\begin{aligned} \langle \Psi_{ab}^{pq} | \widehat{H} e^{\widehat{T}_2} | \Psi_0 \rangle &= E_{\text{CCD}} \langle \Psi_{ab}^{pq} | e^{\widehat{T}_2} | \Psi_0 \rangle \\ &= \langle \Psi_0 | \widehat{H} e^{\widehat{T}_2} | \Psi_0 \rangle \langle \Psi_{ab}^{pq} | e^{\widehat{T}_2} | \Psi_0 \rangle. \end{aligned} \quad (211)$$

The expansion (Eq. 193) tailored to the present case,

$$e^{\widehat{T}_2} = \widehat{1} + \widehat{T}_2 + \frac{1}{2} \widehat{T}_2^2 + \dots, \quad (212)$$

and substituted to the left-hand side of Eq. 211 gives

$$\langle \Psi_{ab}^{pq} | \widehat{H} e^{\widehat{T}} | \Psi_0 \rangle = \langle \Psi_{ab}^{pq} | \widehat{H} \left(\widehat{1} + \widehat{T}_2 + \frac{1}{2} \widehat{T}_2^2 \right) | \Psi_0 \rangle. \quad (213)$$

Further terms are not necessary; in such a case, the functions on both sides of the integral would differ with four and more spin-orbital exchanges (and Hamiltonian is still a sum of one- and two-electron operators). Similarly, the expansion in the integral $\langle \Psi_0 | \widehat{H} e^{\widehat{T}_2} | \Psi_0 \rangle$ will also be truncated on the second term:

$$\langle \Psi_0 | \widehat{H} e^{\widehat{T}_2} | \Psi_0 \rangle = \langle \Psi_0 | \widehat{H} \left(\widehat{1} + \widehat{T}_2 \right) | \Psi_0 \rangle. \quad (214)$$

Remembering that

$$E_0^{(0)} = \langle \Psi_0 | \widehat{H} | \Psi_0 \rangle, \quad (215)$$

one gets

$$\langle \Psi_0 | \widehat{H} e^{\widehat{T}_2} | \Psi_0 \rangle = E_0^{(0)} + \langle \Psi_0 | \widehat{H} \widehat{T}_2 | \Psi_0 \rangle. \quad (216)$$

The last integral on the right-hand side of Eq. 211, $\langle \Psi_{ab}^{pq} | e^{\widehat{T}} | \Psi_0 \rangle$, can be nonvanishing only if the functions on the right and left side are the same. This is possible for

$$\langle \Psi_{ab}^{pq} | e^{\widehat{T}_2} | \Psi_0 \rangle = \langle \Psi_{ab}^{pq} | \widehat{T}_2 | \Psi_0 \rangle. \quad (217)$$

This is the integral that can provide information about the desired amplitudes of Eq. 210. Putting all these together, one gets

$$\langle \Psi_{ab}^{pq} | \widehat{H} \left(\widehat{1} + \widehat{T}_2 + \frac{1}{2} \widehat{T}_2^2 \right) | \Psi_0 \rangle = \left(E_0 + \langle \Psi_0 | \widehat{H} \widehat{T}_2 | \Psi_0 \rangle \right) \langle \Psi_{ab}^{pq} | \widehat{T}_2 | \Psi_0 \rangle. \quad (218)$$

Therefore, the amplitude $t_{ab}^{pq} = \langle \Psi_{ab}^{pq} | \hat{T}_2 | \Psi_0 \rangle$ can be expressed as

$$t_{ab}^{pq} = \frac{\langle \Psi_{ab}^{pq} | \hat{H} (\hat{1} + \hat{T}_2 + \frac{1}{2} \hat{T}_2^2) | \Psi_0 \rangle}{E_0 + \langle \Psi_0 | \hat{H} \hat{T}_2 | \Psi_0 \rangle}. \quad (219)$$

Unluckily, this does not mean that the amplitudes are known. Still, the above expression also contains the t_{ab}^{pq} amplitudes on the right-hand side in the \hat{T}_2 operators. Moreover, all other amplitudes are also present on the right-hand side. The consequence of this aggravation is that the CC equations cannot be solved separately, one by one. All together, the complicated set of nonlinear equations must be handled. The number of equations is equal to the number of sought amplitudes. This is the main reason for the huge computational cost of the CC calculations, even though the variationality of the method was abandoned (Atkins and Friedman 2005; Cramer 2004; Jensen 2006; Levine 2008; Piela 2007; Roos and Widmark 2002).

It can be seen that solving the CC equations is quite complicated, even in the simplified case of the CCD approach. If one wanted to use the variational Hamiltonian and apply its average value, the following integrals would appear:

$$\langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi_0 | e^{\hat{T}} \hat{H} e^{\hat{T}} | \Psi_0 \rangle = \langle e^{\hat{T}} \Psi_0 | \hat{H} | e^{\hat{T}} \Psi_0 \rangle. \quad (220)$$

In order to calculate them, one needs to know the form of all the \hat{T}_i operators, since not only the function on the right-hand side of the above integral will contain the exchanged spin-orbitals but also the function on the left-hand side. Therefore, one needs to calculate terms like $\langle \hat{T}_3 \Psi_0 | \hat{H} | \hat{T}_2 \Psi_0 \rangle$ and many others. This causes the significant increase of the computational costs of the CC method.

Like in the MP n case, the CC method is worth using for the short expansion of the \hat{T} operator. Thus, relatively good accuracy is obtained with a moderate price.

Conclusions

We have finally reached the end of the zeroth iteration in the process of learning quantum chemistry methods. The beginner may feel saturated or even overwhelmed; however, we hope that this chapter arouses interest. Our aim was to show that simple ideas underlie quantum chemistry methods. The purpose is to put complicated things in a simpler and more convenient form. One of the most popular rules in computational chemistry is as follows: “If you cannot calculate something, divide it into parts in such a way that you can calculate some contribution while the other is too difficult.” For instance, nonrelativistic energy can be divided into HF energy and correlation energy. Correlation energy accounts for the contribution that we cannot calculate in practice, but methods such as MP n , CC, and CI allow one to find some

part of it. It may happen (and it often does!) that what we can calculate will be enough.

This chapter should be treated as the introduction to more advanced handbooks or as a guide through the symbols and concepts applied in the later parts of this book. Thus, some of the concepts are just touched upon, and many are omitted. If the reader noticed this and wants to know more, it means that this chapter has met its goal.

Acknowledgments Authors are grateful to Dr Krzysztof Strasburger, Dr Andrej Antušek, Agnieszka Zawada, and Łukasz Mentel for helpful comments.

Bibliography

- Atkins, P., & Friedman, R. (2005). *Molecular quantum mechanics*. Oxford: Oxford University Press.
- Cramer, C. J. (2004). *Essentials of computational chemistry: Theories and models*. Chichester: Wiley.
- Griffiths, D. J. (2004). *Introduction to quantum mechanics* (2nd ed.). San Francisco: Benjamin Cummings.
- Jensen, F. (2006). *Introduction to computational chemistry*. Chichester: Wiley.
- Levine, I. N. (2008). *Quantum chemistry* (6th ed.). Englewood Cliffs: Prentice Hall.
- Lowe, J. P., & Peterson, K. (2005). *Quantum chemistry* (3rd ed.). Boston: Academic.
- McQuarrie, D. A., & Simon, J. D. (1997). *Physical chemistry: A molecular approach*. Sausalito: University Science.
- Piela, L. (2007). *Ideas of quantum chemistry*. Amsterdam: Elsevier.
- Ratner, M. A., & Schatz, G. C. (2000). *Introduction to quantum mechanics in chemistry*. Upper Saddle River: Prentice Hall.
- Roos, B. O., & Widmark, P.-O. (Eds.). (2002). *European summerschool in quantum chemistry*. Lund: Lund University.
- Szabo, A., & Ostlund, N. S. (1996). *Modern quantum chemistry: Introduction to advanced electronic structure theory*. Mineola: Dover.

Adiabatic, Born-Oppenheimer, and Non-adiabatic Approaches

5

Monika Stanke

Contents

Multi-electron Molecule	175
The Equation for the Motion of Nuclei	178
Adiabatic Approximation	182
Qualitative Analysis of the Adiabatic Approximations	184
Born-Oppenheimer Approximation	188
Rotations and Vibrations of Diatomic Molecules	191
Separation of Rotation and Vibration	191
Nuclear Vibrations: Harmonic Approximation	194
A More Detailed Look at the Molecule	197
Nuclear Vibrations: Anharmonic Approximation	197
Rotations: Non-rigid Rotor	198
The Influence of the Electronic Motion	200
The Collapse of the Adiabatic Approximation	202
Apart from the BO Approximation	208
Charge Asymmetry in Pure Vibrational States of the HD Molecule	213
Molecular Relativistic Corrections Determined in the Framework Where the Born-Oppenheimer Approximation Is Not Assumed	215
Appendix A	217
Appendix B	218
Bibliography	221

Abstract

A detailed derivation of the adiabatic approximation and the Born-Oppenheimer approximation is presented, the difference between these two approximations is discussed and the circumstances under which the adiabatic approximation

M. Stanke (✉)

Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicholas Copernicus
University, Toruń, Poland

e-mail: monika@fizyka.umk.pl

collapses are discussed. It is shown that the solution of the Schrödinger equation in the adiabatic approximation can be divided into one representing the motion of electrons in the field of fixed nuclei and another one representing the motion of nuclei in the potential generated by the presence of the electrons. The shapes of the potential energy curves generated by the electrons and the motion of the nuclei in these potentials are also analyzed. Finally, the state-of-the-art highly accurate calculations for diatomic molecules performed without the use of the Born-Oppenheimer approximation is presented.

Solving the Schrödinger equation is very complex, even in the case of atoms. For molecules the complexity of the calculations becomes so great that often we are unable to perform them at all. Then we have to resort to the Born-Oppenheimer (BO) approximation (Born and Oppenheimer 1927) which is one of the most fundamental approximations in quantum chemistry. The BO approximation allows to separate the electronic and the nuclear motion in the molecule. The high accuracy of this approximation¹ is a consequence of a very small ratio of the mass of electron and the masses of the nuclei. Due to this difference, the nuclei are sluggish in their motions relative to the fast-moving electrons. For clarity and simplicity, the discussion is restricted to the case of diatomic molecules.

In this chapter, the reader will find a detailed derivation of the adiabatic approximation and the BO approximation. The significant difference between the two approximations will be shown. It will be also shown that the solution of the Schrödinger equation in the adiabatic approximation can be divided into the solution representing the motion of electrons in the field of fixed nuclei and the solution representing the motion of nuclei in the potential generated by the presence of the electrons in areas around the nuclei and between the nuclei in the molecule. In the next part, we will discuss the possible shapes of the potential energy curve generated by the electrons and how the motion of nuclei in this potential can be described. The circumstances in which the adiabatic approximation collapses will also be discussed. At the end the section, the reader will be introduced to the area of state-of-the-art highly accurate calculations for diatomic molecules performed without the use of the BO approximation.

The atomic units are used in this work; however, the electron mass is always explicitly shown in the formulas.

¹Precisely, the BO and adiabatic approximations are based on the difference of the time scales of movements of the nuclei and the electrons. For example, in some weakly bound molecular anions, the excess electron is very weakly bound (e.g., in the water anion). Such an electron moves at a speed comparable to the oscillating motion of atomic nuclei. A complete description of the dynamics of such an anion in the framework of the adiabatic approximation is doomed to failure. The failure can be attributed to the fact that the electron may be unbound for some large areas of the configuration space of the nuclei. In these areas the electronic wave function is not square integrable and the adiabatic corrections are divergent.

Multi-electron Molecule

Let us consider a molecule composed of N_n nuclei with masses M_a (for $a = 1, \dots, N_n$) and N_e electrons of masses m in a Cartesian reference frame referred to as the laboratory reference frame. For simplicity let us take a three-electron diatomic molecule with the nuclear masses M_1, M_2 located, respectively, at $\mathbf{R}'_1 = [X'_1, Y'_1, Z'_1]$, $\mathbf{R}'_2 = [X'_2, Y'_2, Z'_2]$. The electron coordinates are denoted as $\mathbf{r}'_i = [x'_i, y'_i, z'_i]$, $i = 1, 2, 3$ Fig. 1.

The Hamiltonian for this molecule is a sum of the kinetic energy operator $\hat{T}(\mathbf{r}') + \hat{T}(\mathbf{R}')$, and the Coulomb potential energy operator $\hat{V}(|\mathbf{R}'|, |\mathbf{r}'|)$, which depends on the inter-particle distances only:

$$\hat{H} = \hat{T}(\mathbf{R}') + \hat{T}(\mathbf{r}') + \hat{V}(|\mathbf{R}'|, |\mathbf{r}'|), \quad (1)$$

where

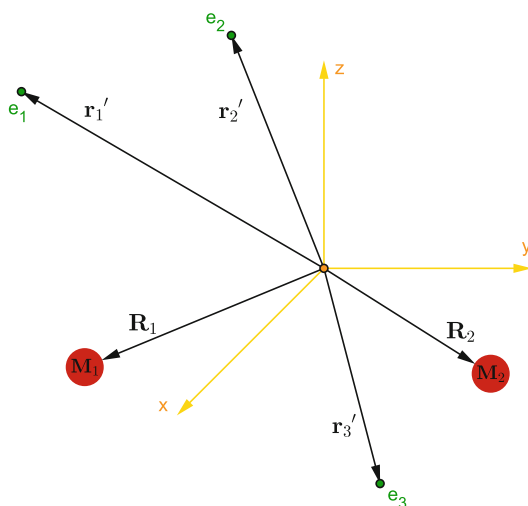
$$\hat{T}(\mathbf{R}') = -\frac{1}{2} \left(\frac{1}{M_1} \Delta_{\mathbf{R}'_1} + \frac{1}{M_2} \Delta_{\mathbf{R}'_2} \right) \quad (2)$$

$$\hat{T}(\mathbf{r}') = -\frac{1}{2m} \sum_{i=1}^3 \Delta_{\mathbf{r}'_i} \quad (3)$$

$$\hat{V}(|\mathbf{R}'|, |\mathbf{r}'|) = \sum_{a < b}^2 \frac{Z_a Z_b}{|\mathbf{R}'_{ab}|} - Z_a \sum_{i=1}^3 \frac{1}{|\mathbf{r}'_{ia}|} - Z_b \sum_{i=1}^3 \frac{1}{|\mathbf{r}'_{ib}|} + \sum_{i < j}^3 \frac{1}{|\mathbf{r}'_{ij}|}. \quad (4)$$

The first term describes the Coulomb repulsion between nuclei, the second and third the attraction between electrons and nuclei, and the fourth the mutual repulsion between the electrons. In formulas we have used the abbreviations

Fig. 1 Diatomic molecule in a Cartesian coordinate system. By $\mathbf{R}'_a = [X'_a, Y'_a, Z'_a]$ and $\mathbf{r}'_i = [x'_i, y'_i, z'_i]$, where $a = 1, 2$ and $i = 1, 2, 3$, are denoted as the position vectors of nuclei and electrons, respectively



$$|\mathbf{R}'_{ab}| = |\mathbf{R}'_a - \mathbf{R}'_b|, \quad |\mathbf{r}'_{ia}| = |\mathbf{r}'_i - \mathbf{R}'_a|, \quad |\mathbf{r}'_{ib}| = |\mathbf{r}'_i - \mathbf{R}'_b|, \quad |\mathbf{r}'_{ij}| = |\mathbf{r}'_i - \mathbf{r}'_j|.$$

In a general case, the Hamiltonian reads

$$\hat{T}(\mathbf{R}') + \hat{T}(\mathbf{r}') = -\frac{1}{2m} \sum_{i=1}^{N_e} \Delta_{\mathbf{r}'_i} - \frac{1}{2} \sum_{a=1}^{N_n} \frac{1}{M_a} \Delta_{\mathbf{R}'_a} \quad (5)$$

$$\hat{V}(|\mathbf{R}'|, |\mathbf{r}'|) = \underbrace{\sum_{a < b}^{N_n} \frac{Z_a Z_b}{|\mathbf{R}'_{ab}|}}_{\text{nuclears repulsion}} - \underbrace{\sum_{a=1}^{N_n} Z_a \sum_{i=1}^{N_e} \frac{1}{|\mathbf{r}'_{ia}|}}_{\text{electron-nucleus-attraction}} + \underbrace{\sum_{i < j}^{N_e} \frac{1}{|\mathbf{r}'_{ij}|}}_{\text{electrons repulsion}}. \quad (6)$$

Since we are not interested in molecular collisions, we assume that the center of mass of the molecule does not move relative to the laboratory reference frame. Therefore, it is convenient to introduce *internal coordinates*. The new coordinates, in the case of the diatomic molecule, are defined as follows²

$$\mathbf{R} = [R_x, R_y, R_z] = \mathbf{R}'_1 - \mathbf{R}'_2 \quad (7)$$

\mathbf{R} vector of the relative position of the nuclei

$$\mathbf{R}_{\text{CM}} = [X_{\text{CM}}, Y_{\text{CM}}, Z_{\text{CM}}] = \frac{1}{M} \left(\sum_{a=1}^2 M_a \mathbf{R}'_a + \sum_{i=1}^3 m \mathbf{r}'_i \right) \quad (8)$$

\mathbf{R}_{CM} coordinates of the center of mass (CM) of the molecule, where

$$M = \sum_{a=1}^2 M_a + \sum_{i=1}^3 m_i \quad (9)$$

M the total mass of the molecule

$$X_{\text{CM}} = \frac{1}{M} \left(\sum_{a=1}^2 M_a X'_a + \sum_{i=1}^{N_e} m x'_i \right), \quad Y_{\text{CM}} = \frac{1}{M} \left(\sum_{a=1}^2 M_a Y'_a + \sum_{i=1}^{N_e} m y'_i \right),$$

$$Z_{\text{CM}} = \frac{1}{M} \left(\sum_{a=1}^2 M_a Z'_a + \sum_{i=1}^{N_e} m z'_i \right)$$

$$\mathbf{r}_i = [x_i, y_i, z_i] = \mathbf{r}'_i - \frac{1}{2}(\mathbf{R}'_1 + \mathbf{R}'_2) \quad (10)$$

\mathbf{r}_i is the vector of the electron position anchored at the geometric center of the molecule, and

²We use the most common definition of the new coordinates. Internal coordinates may be defined in many different ways, see, e.g., Piela (2007).

$$x_i = x'_i - \frac{1}{2} (X'_1 + X'_2), \quad y_i = y'_i - \frac{1}{2} (Y'_1 + Y'_2), \quad z_i = z'_i - \frac{1}{2} (Z'_1 + Z'_2). \quad (11)$$

Similarly, in the extended formalism, for any molecule with N_n nuclei of mass M_a and n_n electrons, we have

$$M = \sum_{a=1}^{N_n} M_a + N_e \cdot m, \quad (12)$$

$$\mathbf{R}_{ab} = \mathbf{R}'_a - \mathbf{R}'_b \quad (a \neq b) \quad a, b = 1, \dots, N_n, \quad (13)$$

$$\mathbf{R}_{\text{CM}} = [X_{\text{CM}}, Y_{\text{CM}}, Z_{\text{CM}}] = \frac{1}{M} \left(\sum_{a=1}^{N_n} M_a \mathbf{R}'_a + \sum_{i=1}^{N_e} m \mathbf{r}'_i \right), \quad (14)$$

$$\mathbf{r}_i = \mathbf{r}_i - \frac{1}{2} (\mathbf{R}^{(n)} + \mathbf{R}'_{N_n}) \quad \text{where} \quad \mathbf{R}^{(n)} = \frac{1}{2} (\mathbf{R}^{(n-1)} + \mathbf{R}'_{N_{n-1}}) \quad \text{for} \quad n = N_n - 1. \quad (15)$$

In order to transform the molecular Hamiltonian to the new reference system, one has to perform some rather tedious algebra (for details, see Appendix A). The molecular Hamiltonian in the new reference frame may be written as

$$\hat{H} = \hat{H}_{\text{CM}} + \hat{H}_0 + \hat{H}', \quad (16)$$

$$\hat{H}_{\text{CM}} = -\frac{1}{2M} \Delta_{\mathbf{R}_{\text{CM}}}, \quad (17)$$

$$\hat{H}_0 = -\frac{1}{2m} \sum_{i=1}^{N_e} \Delta_{\mathbf{r}_i} + \hat{V}(|\mathbf{R}|, |\mathbf{r}|), \quad (18)$$

$$\hat{H}' = \underbrace{-\frac{1}{2} \left(\frac{1}{M_1} + \frac{1}{M_2} \right) \Delta_{\mathbf{R}}}_{-\frac{1}{2\mu} \Delta_{\mathbf{R}}} - \underbrace{\frac{1}{8} \left(\frac{1}{M_1} + \frac{1}{M_2} \right) \left(\sum_{i=1}^{N_e} \nabla_{\mathbf{r}_i} \right)^2 - \frac{1}{2} \left(\frac{1}{M_1} - \frac{1}{M_2} \right) \nabla_{\mathbf{R}} \sum_{i=1}^{N_e} \nabla_{\mathbf{r}_i}}_{\hat{H}''} \quad (19)$$

where

$$\hat{V}(|\mathbf{R}|, |\mathbf{r}|) = \frac{Z_1 Z_2}{|\mathbf{R}|} - \sum_{i=1}^{N_e} \frac{Z_1}{|\mathbf{r}_i - \frac{1}{2}\mathbf{R}|} - \sum_{i=1}^{N_e} \frac{Z_2}{|\mathbf{r}_i + \frac{1}{2}\mathbf{R}|} + \sum_{i>j=1}^{N_e} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (20)$$

The kinetic energy of CM is described by \hat{H}_{CM} and is responsible for the translational motion of the entire molecule. Therefore, it does not carry any information essential for us. Hereafter, we will consider the molecular Hamiltonian of the following form

$$\hat{H} = \hat{H}_0 + \hat{H}'. \quad (21)$$

The motion of electrons in the Coulomb field generated by fixed nuclei is described by \hat{H}_0 with \hat{V} describing Coulomb interactions between all particles. The nuclear motion and the coupling between the nuclei and electrons is described by \hat{H}' . Here the nuclear part $-1/2\mu \Delta_{\mathbf{R}}$ (in (19)) is equal to the kinetic energy of a "pseudonucleus" with mass $\mu = M_1 M_2 / (M_1 + M_2)$. The second term, \hat{H}'' , describes the coupling of the nuclear and the electronic motion. The last term in \hat{H}'' , known as the mass-polarization term, vanishes in the case of homonuclear diatomic molecules.³

The coordinate transformation (7), (8), (9), and (10) is not unique. It is possible, for example, to use another transformation, proposed by Author in Dalgarno and McCarroll (1956), where instead of the geometric center of the molecule, one takes the center of the nuclear masses:

$$\mathbf{R} = \mathbf{R}'_a - \mathbf{R}'_b \quad \mathbf{r}_i = \mathbf{r}'_i - \frac{1}{M_a + M_b} (M_a \mathbf{R}'_a + M_b \mathbf{R}'_b).$$

In this case,

$$\hat{H}_0 = -\frac{1}{2m} \sum_{i=1}^{N_e} \Delta_{\mathbf{r}_i} + \hat{V}(|\mathbf{R}|, |\mathbf{r}|) \quad (22)$$

$$\hat{H}' = -\frac{1}{2} \left(\frac{1}{M_1} + \frac{1}{M_2} \right) \Delta_{\mathbf{R}} - \frac{1}{2} \left(\frac{1}{M_1} + \frac{1}{M_2} \right) \left(\sum_{i=1}^{N_e} \nabla_{\mathbf{r}_i} \right)^2. \quad (23)$$

There is no mass-polarization term, but the coordinates \mathbf{r}_i and, consequently, also $\hat{V}(|\mathbf{R}|, |\mathbf{r}|)$ depend on the nuclear masses. Hereafter transformation (16), (17), (18), and (19) is used.

The Equation for the Motion of Nuclei

Our aim is to solve the stationary Schrödinger equation with Hamiltonian (21),

$$(\hat{H}_0 + \hat{H}')\Psi(\mathbf{r}, \mathbf{R}) = E\Psi(\mathbf{r}, \mathbf{R}).$$

Let us allow our imagination to help – this will facilitate the understanding our further considerations. The nuclei, due to their heaviness, move very slowly.

³“Frolov’s calculations showed that when one mass increased without limit (the atomic case), any discrete spectrum persisted, but when two masses were allowed to increase without limit (the molecular cause), the Hamiltonian ceased to be well defined, and this failure led to what he called adiabatic divergence in attempts to compute discrete eigenstates of (21). This divergence is discussed in some mathematical detail in the Appendix to Frolov (1999). It does not arise from the choice of a translationally invariant form for the electronic Hamiltonian; rather it is due to the lack of any kinetic energy term to dominate the Coulomb potential. Thus it really is essential to characterize the spectrum of H_{elec} to see whether the traditional approach can be validated.” (Sutcliffe 2003)

Electrons, on the contrary, quickly react to changes in the positions of the nuclei. This means that the electronic charge distribution depends on the instant positions of the nuclei rather than on their motion. As a consequence, the motions of the nuclei and electrons in a molecule can be separated to a very good *approximation*. In terms of formalism, this implies that the wave function $\Psi(\mathbf{r}, \mathbf{R})$ can be expressed as a product of its electronic and nuclear parts. In fact, exactly, the wave function can be expressed as an infinite expansion

$$\Psi(\mathbf{r}, \mathbf{R}) = \sum_k \psi_k^{\text{el}}(\mathbf{r}; R) \chi_k(\mathbf{R}), \quad (24)$$

where

- $\psi_k^{\text{el}}(\mathbf{r}; R)$ is the wave function that describes k -th state of electrons in the molecule at a given position of the nuclei ($|\mathbf{R}| = R$). This function is a solution of the electronic Schrödinger equation:

$$\hat{H}_0 \psi_k^{\text{el}}(\mathbf{r}; R) = E_k^0(R) \psi_k^{\text{el}}(\mathbf{r}; R). \quad (25)$$

- $\chi_k(\mathbf{R})$ is an expansion coefficient which depends on \mathbf{R} ; it is also the wave function describing the nuclear motion.

Let us return to the eigenvalue problem of Hamiltonian (21) and the wave function (24). It is convenient to rewrite the molecular Hamiltonian (21) in a slightly modified form by splitting (19) to the pseudonuclei kinetic energy operator $-1/2\mu\Delta_{\mathbf{R}}$ and the operators which couple electron-nuclear variables (\hat{H}''):

$$\left(-\frac{1}{2m} \sum_{i=1}^3 \Delta_{\mathbf{r}_i} + \hat{V}(|\mathbf{R}|, |\mathbf{r}|) + \hat{H}' \right) \Psi(\mathbf{r}; R) = E \Psi(\mathbf{r}; R), \quad (26)$$

where

$$\Psi(\mathbf{r}, \mathbf{R}) = \sum_k \psi_k^{\text{el}}(\mathbf{r}; R) \chi_k(\mathbf{R}),$$

$$\hat{H}' = -\frac{1}{2\mu} \Delta_{\mathbf{R}} + \hat{H}'', \quad (27)$$

$$\hat{H}'' = -\frac{1}{8\mu} \left(\sum_{i=1}^3 \nabla_{\mathbf{r}_i} \right)^2 - \frac{1}{2} \left(\frac{1}{M_a} - \frac{1}{M_b} \right) \nabla_{\mathbf{R}} \sum_{i=1}^3 \nabla_{\mathbf{r}_i}. \quad (28)$$

Multiplying Eq. (26) on the lhs by $\psi_l^{\text{el}}(\mathbf{r}; R)$ and integrating over the electron coordinates, we get⁴:

⁴For the sake of simplicity, we denote $\psi_k^{\text{el}}(\mathbf{r}; R) \equiv \psi_k^{\text{el}}$ and $\chi_k(\mathbf{R}) \equiv \chi_k$.

$$\begin{aligned}
& \sum_k \langle \psi_l^{\text{el}} | \hat{H}_0 + \hat{H}' | \psi_k^{\text{el}} \chi_k \rangle_{\text{el}} = E \sum_k \langle \psi_l^{\text{el}} | \psi_k^{\text{el}} \chi_k \rangle_{\text{el}} \\
& \sum_k \left(\int \mathbf{dr} \psi_l^{\text{el}} \hat{H}_0 \psi_k^{\text{el}} \right) \chi_k + \sum_k \int \mathbf{dr} \psi_l^{\text{el}} \hat{H}' \psi_k^{\text{el}} \chi_k = E \sum_k \int \mathbf{dr} \psi_l^{\text{el}} \psi_k^{\text{el}} \chi_k \\
& \sum_k E_k^0(R) \delta_{kl} \chi_k + \sum_k \int \mathbf{dr} \psi_l^{\text{el}} \hat{H}' \psi_k^{\text{el}} \chi_k = E \sum_k \delta_{kl} \chi_k \\
& (E_l^0(R) - E) \chi_l(\mathbf{R}) + \sum_k \int \mathbf{dr} \psi_l^{\text{el}}(\mathbf{r}; R) \hat{H}' \psi_k^{\text{el}}(\mathbf{r}; R) \chi_k(\mathbf{R}) = 0.
\end{aligned}$$

$\langle \rangle_{\text{el}}$ denotes integration over electronic coordinates. The last equation, in the Dirac notation, takes a simple form:

$$(E_l^0(R) - E) \chi_l(\mathbf{R}) + \sum_k \langle \psi_l^{\text{el}}(\mathbf{r}; R) | \hat{H}' | \psi_k^{\text{el}}(\mathbf{r}; R) \chi_k(\mathbf{R}) \rangle = 0. \quad (29)$$

Using the explicit form of \hat{H}' , one can easily transform the matrix element of \hat{H}' in Eq. (29) in the following way:

$$\langle \psi_l^{\text{el}} | \hat{H}' | \psi_k^{\text{el}} \chi_k \rangle = \langle \psi_l^{\text{el}} | -\frac{1}{2\mu} \Delta_{\mathbf{R}} + \hat{H}'' | \psi_k^{\text{el}} \chi_k \rangle, \quad (30)$$

where

$$\begin{aligned}
\Delta_{\mathbf{R}} (\psi_k^{\text{el}} \chi_k) &= \nabla_{\mathbf{R}}^T \cdot \nabla_{\mathbf{R}} (\psi_k^{\text{el}} \chi_k) = \nabla_{\mathbf{R}}^T \cdot [(\nabla_{\mathbf{R}} \psi_k^{\text{el}}) \chi_k + \psi_k^{\text{el}} (\nabla_{\mathbf{R}} \chi_k)] \\
&= [(\Delta_{\mathbf{R}} \psi_k^{\text{el}}) \chi_k + (\nabla_{\mathbf{R}}^T \psi_k^{\text{el}}) (\nabla_{\mathbf{R}} \chi_k) + (\nabla_{\mathbf{R}}^T \chi_k) (\nabla_{\mathbf{R}} \psi_k^{\text{el}}) + \psi_k^{\text{el}} (\Delta_{\mathbf{R}} \chi_k)] \\
&= [(\Delta_{\mathbf{R}} \psi_k^{\text{el}}) \chi_k + 2 (\nabla_{\mathbf{R}}^T \psi_k^{\text{el}}) (\nabla_{\mathbf{R}} \chi_k) + \psi_k^{\text{el}} (\Delta_{\mathbf{R}} \chi_k)]. \quad (31)
\end{aligned}$$

The first term in the last equation corresponds to the kinetic energy of the pseudonuclei. The remaining two terms are the so-called non-adiabatic contributions (to be discussed hereafter). Substituting (31) to (30), we get

$$\begin{aligned}
\langle \psi_l^{\text{el}} | \hat{H}' | \psi_k^{\text{el}} \chi_k \rangle &= -\frac{1}{2\mu} \langle \psi_l^{\text{el}} | \Delta_{\mathbf{R}} | \psi_k^{\text{el}} \rangle \chi_k - \frac{1}{\mu} \langle \psi_l^{\text{el}} | \nabla_{\mathbf{R}}^T | \psi_k^{\text{el}} \rangle \nabla_{\mathbf{R}} \chi_k \\
&\quad - \frac{1}{2\mu} \langle \psi_l^{\text{el}} | \psi_k^{\text{el}} \rangle \Delta_{\mathbf{R}} \chi_k + \langle \psi_l^{\text{el}} | \hat{H}'' | \psi_k^{\text{el}} \rangle \chi_k. \quad (32)
\end{aligned}$$

Now Eq. (29) may be rewritten as

$$\begin{aligned}
(E_l^0(R) - E) \chi_l &- \frac{1}{2\mu} \sum_k \langle \psi_l^{\text{el}} | \Delta_{\mathbf{R}} | \psi_k^{\text{el}} \rangle \chi_k - \frac{1}{\mu} \sum_k \langle \psi_l^{\text{el}} | \nabla_{\mathbf{R}}^T | \psi_k^{\text{el}} \rangle \nabla_{\mathbf{R}} \chi_k \\
&- \frac{1}{2\mu} \sum_k \langle \psi_l^{\text{el}} | \psi_k^{\text{el}} \rangle \Delta_{\mathbf{R}} \chi_k + \sum_k \langle \psi_l^{\text{el}} | \hat{H}'' | \psi_k^{\text{el}} \rangle \chi_k = 0
\end{aligned} \quad (33)$$

or, in a modified form, as

$$\begin{aligned}
 (E_l^0(\mathbf{R}) - E) \chi_l + \sum_k \langle \psi_l^{\text{el}} | \underbrace{-\frac{1}{2\mu} \Delta_{\mathbf{R}} + \hat{H}''}_{\hat{H}'(R)} | \psi_k^{\text{el}} \rangle \chi_k \\
 - \frac{1}{\mu} \sum_k \langle \psi_l^{\text{el}} | \nabla_{\mathbf{R}}^T | \psi_k^{\text{el}} \rangle \nabla_{\mathbf{R}} \chi_k - \frac{1}{2\mu} \sum_k \delta_{kl} \Delta_{\mathbf{R}} \chi_k = 0 \\
 (E_l^0(\mathbf{R}) - E) \chi_l + \sum_k c_{kl}(\mathbf{R}) \chi_k - \frac{1}{2\mu} \Delta_{\mathbf{R}} \chi_l = 0, \quad (34)
 \end{aligned}$$

where

$$c_{kl}(\mathbf{R}) = \langle \psi_l^{\text{el}} | \hat{H}' | \psi_k^{\text{el}} \rangle - \frac{1}{\mu} \langle \psi_l^{\text{el}} | \nabla_{\mathbf{R}}^T | \psi_k^{\text{el}} \rangle \nabla_{\mathbf{R}}. \quad (35)$$

Note that \hat{H}' depends on the length of \mathbf{R} but not on its direction. Equation (34) may be expressed in a more convenient form as

$$\left(-\frac{1}{2\mu} \Delta_{\mathbf{R}} + E_l^0(\mathbf{R}) \right) \chi_l(\mathbf{R}) + \sum_k c_{kl}(\mathbf{R}) \chi_k(\mathbf{R}) = E \chi_l(\mathbf{R}), \quad (36)$$

where

$$c_{kl}(\mathbf{R}) = \langle \psi_l^{\text{el}} | \hat{H}' | \psi_k^{\text{el}} \rangle - \frac{1}{\mu} \langle \psi_l^{\text{el}} | \nabla_{\mathbf{R}}^T | \psi_k^{\text{el}} \rangle \nabla_{\mathbf{R}}.$$

The above equation is an equation for the nuclear wave function $\chi_l(\mathbf{R})$. This equation together with the Schrödinger equation for the electronic energy and the wave function (Eq. 25) forms a system of coupled equations. The coupling between the electron and nuclear function is in the $c_{kl}(\mathbf{R})$ term. This is a *non-adiabatic coupling operator*. The coupling coefficients $c_{kl}(\mathbf{R})$ are matrix elements of differential operators. Thus, at this point, the concept of the potential energy surface (PES) has no physical meaning. (It should be emphasized at this point that c_{kl} cannot be regarded as part of the potential energy operator.) A physically justified concept of PES appears only when we separate the electronic and the nuclear motions. This happens when we apply the adiabatic or BO approximation (see section “Born-Oppenheimer Approximation”). Then the PSE $E_l^0(\mathbf{R})$ is defined by a presumed configuration of the nuclei in a given electronic state.

Adiabatic Approximation

Equations (36) may be decoupled if we neglect all $c_{kl}(R)$ except the diagonal ones, i.e., if we set $c_{kl} = c_{ll} \delta_{kl}$. The decoupled equations read

$$\left(-\frac{1}{2\mu} \Delta_{\mathbf{R}} + E_l^0(R) + c_{ll}(R) \right) \chi_l(\mathbf{R}) = E \chi_l(\mathbf{R}) - \sum_{k \neq l} c_{kl}(R) \chi_k(\mathbf{R}). \quad (37)$$

Diagonal terms $c_{ll}(R)$ can be simplified using the normalization condition for ψ_l^{el} :

$$\langle \psi_l^{\text{el}} | \psi_l^{\text{el}} \rangle = 1 \quad \Rightarrow \quad \nabla_{\mathbf{R}}^T \langle \psi_l^{\text{el}} | \psi_l^{\text{el}} \rangle = 2 \langle \psi_l^{\text{el}} | \nabla_{\mathbf{R}}^T | \psi_l^{\text{el}} \rangle = 0.$$

Then, from Eq. (35) results that

$$c_{ll}(R) = \langle \psi_l^{\text{el}} | \hat{H}' | \psi_l^{\text{el}} \rangle. \quad (38)$$

Thus, the diagonal elements in the expression (35) take the form⁵

$$c_{ll}(R) = \langle \psi_l^{\text{el}} | \hat{H}' | \psi_l^{\text{el}} \rangle = -\frac{1}{2\mu} \langle \psi_l^{\text{el}} | \Delta_{\mathbf{R}} | \psi_l^{\text{el}} \rangle = +\frac{1}{2\mu} \int d\mathbf{r} |\nabla_{\mathbf{R}} \psi_l^{\text{el}}|^2. \quad (39)$$

This diagonal term is usually very small compared to $E_l^0(R)$. Using the definition (39), Eq. (36) fulfilled by $\chi_l(\mathbf{R})$ can now be written as

$$\left(-\frac{1}{2\mu} \Delta_{\mathbf{R}} + E_l^0(R) + \frac{1}{2\mu} \int d\mathbf{r} |\nabla_{\mathbf{R}} \psi_l^{\text{el}}|^2 \right) \chi_l(\mathbf{R}) = E \chi_l(\mathbf{R}) - \sum_{k \neq l} c_{kl}(R) \chi_k(\mathbf{R}). \quad (40)$$

The off-diagonal terms $c_{kl}(R)$, $k \neq l$, in Eq. (40) usually cause difficulties in solving the relevant Schrödinger equation. For this reason and due to the fact that this term is very small, they are usually neglected in calculation. Such an approximation is called *the adiabatic approximation*. Thus, in the zero-order approximation, the functions $\chi_l(\mathbf{R})$ are solutions of the following equation:

$$\left(-\frac{1}{2\mu} \Delta_{\mathbf{R}} + U_l(R) \right) \chi_l(\mathbf{R}) = E \chi_l(\mathbf{R}) \quad (41)$$

$$\text{for } U_l(R) = E_l^0(R) + \frac{1}{2\mu} \int d\mathbf{r} |\nabla_{\mathbf{R}} \psi_l^{\text{el}}(\mathbf{r}; \mathbf{R})|^2,$$

⁵As a result of simple transformations:

$$\Delta \langle \psi | \psi \rangle = 0$$

but

$$\begin{aligned} \nabla^T (\nabla \langle \psi | \psi \rangle) &= 2 \nabla^T \langle \psi | \nabla | \psi \rangle = 2 \langle \nabla \psi | \nabla | \psi \rangle + 2 \langle \psi | \Delta | \psi \rangle \\ &\Rightarrow \langle \nabla \psi | \nabla | \psi \rangle = - \langle \psi | \Delta | \psi \rangle = - |\nabla \psi|^2 \end{aligned}$$

where $U_i(R)$ is called *effective potential*, where the diagonal correction $c_{ii}(R)$ (39) is usually very small compared to $E_i^0(R)$. The effective potential $U_i(R)$ depends on the nuclear masses ($1/\mu = 1/M_a + 1/M_b$). Therefore, it is different for different isotopes. The term $c_{ii}(R)$ does not appear when in the description of the molecules we use the BO approximation (section “Born-Oppenheimer Approximation”). Therefore, in the BO approximation, the effective potential is the same for all isotopes, e.g., for H₂, HD, D₂.

In the adiabatic approximation, one assumes that the expansion of the wave function (24) may be reduced to a single term (Kołos and Wolniewicz 1963; Kołos 1970; Davydov 1965), i.e.,

$$\Psi_k(\mathbf{r}, \mathbf{R}) \approx \psi_k^{\text{el}}(\mathbf{r}; R) \chi_k(\mathbf{R}). \quad (42)$$

This is the essence of the adiabatic approximation.

Adiabatic approximation: to solve the eigenvalue problem of operator (21), we use function (42). Because we cannot provide the eigenfunctions (21) as a combination of products of functions solely dependent on the nuclear coordinates ($\{\phi_a(\mathbf{R})\}$) and functions solely dependent on the electrons coordinates ($\{\phi_i(\mathbf{r})\}$)

$$\Psi(\mathbf{r}, \mathbf{R}) = \sum_{i,a} A_{i,a} \phi_i(\mathbf{r}) \phi_a(\mathbf{R}),$$

we use the adiabatic approximation. First, we solve the electronic Schrödinger equation:

$$\hat{H}_0(R) \psi_k^{\text{el}}(\mathbf{r}; R) = E_k^0(R) \psi_k^{\text{el}}(\mathbf{r}; R),$$

where H_0 does not depend on the motion of the nuclei, although $\hat{V}(|\mathbf{R}|, |\mathbf{r}|)$ depends on the positions of the nuclei. Solutions $E_k^0(R)$ and $\psi_k^{\text{el}}(\mathbf{r}; R)$ depend on \mathbf{R} parametrically. The parametric dependence is indicated by the semicolon. By solving equation (25) for fixed internuclear distances \mathbf{R} , we get the energy $E_k^0(R)$ for a rigid molecule. Hence, $|\psi_k^{\text{el}}(\mathbf{r}; R)|^2$ determines the charge distribution in the k -th electronic state (for more details, the reader is referred to Piela 2007). In the next step, for the new positions of the nuclei, we again solve the problem (25), to get a new value $E_k^0(R)$. Eigenfunctions of the molecular Hamiltonian (21) may then be expanded in terms of the eigenfunctions of the electronic Hamiltonian \hat{H}_0 :

$$\Psi(\mathbf{r}, \mathbf{R}) \approx \psi_k^{\text{el}}(\mathbf{r}; R) \chi_k(\mathbf{R}),$$

where the coefficients $\chi_k(\mathbf{R})$ are functions of \mathbf{R} . The following orthonormal conditions are fulfilled:

$$\langle \psi_k^{\text{el}}(\mathbf{r}; R) | \psi_l^{\text{el}}(\mathbf{r}; R) \rangle_{\text{el}} = \delta_{kl} \quad \text{integration over electronic coordinates} \quad (43)$$

$$\langle \chi_k(\mathbf{R}) | \chi_l(\mathbf{R}) \rangle_n = \delta_{kl} \quad \text{integration over nuclear coordinates} \quad (44)$$

Qualitative Analysis of the Adiabatic Approximations

In this section a sketch of a rather simple exercise (Kręglewski 1979) that allows us to understand the criteria for determining the range of applicability of the adiabatic approximation is presented. The task is to solve the Schrödinger equation

$$\left(-\frac{1}{2\mu_1} \frac{\partial^2}{\partial x_1^2} + \frac{1}{2} k_1 x_1^2 - \frac{1}{2\mu_2} \frac{\partial^2}{\partial x_2^2} + \frac{1}{2} k_2 x_2^2 + k_{12} x_1 x_2 \right) \Psi(x_1, x_2) = E(1, 2) \Psi(x_1, x_2), \quad (45)$$

exactly using the adiabatic approximation. The equation we have to solve describes two coupled harmonic oscillators. The choice of the harmonic oscillator model is not accidental. First, each of the readers, after a basic course of quantum mechanics, can solve the corresponding eigenvalue problem. Second, harmonic oscillator model always describes bound states. Hence, the proposed exercise avoids difficulties and traps that may arise if continuous spectrum is present, though many aspects related to processes where the continuous spectrum is involved (e.g., dissociation) are completely missing. Another advantage of this problem is its exact solubility. By comparing the exact solutions with the approximate ones, we can see the limits of the applicability of the adiabatic approximation.

Step 1: Adiabatic approximation. Let us introduce new variables:

$$\sqrt{\mu_i} x_i = y_i \quad \text{for } i = 1, 2 \quad (46)$$

and new designations:

$$\lambda_i \equiv \frac{k_i}{\mu_i}, \quad \kappa \equiv \frac{k_{12}}{\sqrt{\mu_1 \mu_2}} \quad \text{for } i = 1, 2. \quad (47)$$

Now Eq. (45) reads

$$\left(\hat{H}(y_1) + \hat{H}(y_2) + \kappa y_1 y_2 \right) \Psi(y_1, y_2) = E(1, 2) \Psi(y_1, y_2), \quad (48)$$

where, in order to simplify the equations, we have introduced the following markings on one particle Hamiltonians

$$\hat{H}(y_1) \equiv -\frac{1}{2} \frac{\partial^2}{\partial y_1^2} + \frac{1}{2} \lambda_1 y_1^2 \quad \text{and} \quad \hat{H}(y_2) \equiv -\frac{1}{2} \frac{\partial^2}{\partial y_2^2} + \frac{1}{2} \lambda_2 y_2^2. \quad (49)$$

According to the adiabatic approximation, we assume that the eigenfunctions in Eq. (48) can be approximated as

$$\Psi(y_1, y_2) = \phi_1(y_1, y_2) \phi_2(y_2). \quad (50)$$

Substituting functions (50) to Eq. (48), we obtain

$$\begin{aligned} & \left(\hat{H}(y_1) + \hat{H}(y_2) + \kappa y_1 y_2 \right) \phi_1(y_1, y_2) \phi_2(y_2) = \varepsilon \phi_1(y_1, y_2) \phi_2(y_2) \\ & \left[\left(\hat{H}(y_1) + \kappa y_1 y_2 \right) \phi_1(y_1, y_2) \right] \phi_2(y_2) + \\ & \left[\hat{H}(y_2) \phi_1(y_1, y_2) \right] \phi_2(y_2) + \left[\hat{H}(y_2) \phi_2(y_2) \right] \phi_1(y_1, y_2) = \varepsilon \phi_1(y_1, y_2) \phi_2(y_2). \end{aligned}$$

Assuming that

$$\left(-\frac{1}{2} \frac{\partial^2}{\partial y_1^2} + \frac{1}{2} \lambda_1 y_1^2 + \kappa y_1 y_2 \right) \phi_1(y_1, y_2) = E_1(y_2) \phi_1(y_1, y_2), \quad (51)$$

we get

$$E_1(y_2) \phi_1(y_1, y_2) \phi_2(y_2) + \left[\hat{H}(y_2) \phi_2(y_2) \right] \phi_1(y_1, y_2) = \varepsilon \phi_1(y_1, y_2) \phi_2(y_2).$$

Multiplying the last equation on the lhs by $\phi_1(y_1, y_2)$ and integrating with respect to y_1 , we obtain

$$\left(\hat{H}(y_2) + E_1(y_2) \right) \phi_2(y_2) + \langle \phi_1(y_1, y_2) | -\frac{1}{2} \frac{\partial^2}{\partial y_1^2} \phi_1(y_1, y_2) \rangle \phi_2(y_2) = \varepsilon \phi_2(y_2). \quad (52)$$

Equation (51) can be rewritten as

$$\left(-\frac{1}{2} \frac{\partial^2}{\partial y_1^2} + \frac{1}{2} \lambda_1 \left(y_1 + \frac{\kappa}{\lambda_1} y_2 \right)^2 \right) \phi_1(y_1, y_2) = \left(E_1(y_2) + \frac{1}{2} \frac{\kappa^2}{\lambda_1} y_2^2 \right) \phi_1(y_1, y_2). \quad (53)$$

It can be recognized as the harmonic oscillator Schrödinger equation. The eigenfunctions are the well-known Hermite polynomials and energy is equal to

$$\phi_1(\tilde{y}) = \phi_{n_1}(\tilde{y}) = N_{n_1} H_{n_1} \left(\sqrt[4]{\lambda_1} \tilde{y} \right) \exp \left[-\frac{1}{2} \sqrt{\lambda_1} \tilde{y}^2 \right] \quad \text{where} \quad \tilde{y} \equiv y_1 + \frac{\kappa}{\lambda_1} y_2$$

$$E_1 = E_{n_1} = \sqrt{\lambda_1} \left(n_1 + \frac{1}{2} \right) - \frac{1}{2} \frac{\kappa^2}{\lambda_1} y_2^2 \quad \text{where} \quad n_1 = 0, 1, 2, \dots$$

In order to solve Eq. (52), let us first calculate the value of

$$\langle \phi_1(y_1, y_2) | -\frac{1}{2} \frac{\partial^2}{\partial y_2^2} \phi_1(y_1, y_2) \rangle = -\frac{1}{2} \int_0^\infty dy_1 \phi_1(y_1, y_2) \frac{\partial^2}{\partial y_2^2} \phi_1(y_1, y_2).$$

By introducing a new variable $y_1 + \frac{\kappa}{\lambda_1} y_2 \equiv \tilde{y}$, we get

$$-\frac{1}{2} \left(\frac{\kappa^2}{\lambda_1} \right)^2 \int_0^\infty d\tilde{y}_1 \phi_1(y_1, y_2) \frac{\partial^2}{\partial \tilde{y}^2} \phi_1(y_1, y_2) = \frac{1}{2} \sqrt{\lambda_1} \left(n_1 + \frac{1}{2} \right) \frac{\kappa^2}{\lambda_1}.$$

In this way Eq. (52) may be rewritten as

$$\left(\hat{H}(y_2) + \sqrt{\lambda_1} \left(n_1 + \frac{1}{2} \right) - \frac{1}{2} \frac{\kappa^2}{\lambda_1} y_2^2 \right) \phi_2(y_2) + \frac{1}{2} \sqrt{\lambda_1} \left(n_1 + \frac{1}{2} \right) \frac{\kappa^2}{\lambda_1} \phi_2(y_2) = \varepsilon \phi_2(y_2)$$

$$\left(-\frac{1}{2} \frac{\partial^2}{\partial y_2^2} + \frac{1}{2} \left(\lambda_2 - \frac{\kappa^2}{\lambda_1} \right) y_2^2 \right) \phi_2(y_2) = \left(\varepsilon - \frac{1}{2} \sqrt{\lambda_1} \left(n_1 + \frac{1}{2} \right) \left(\frac{\kappa^2}{\lambda_1} + 2 \right) \right) \phi_2(y_2). \quad (54)$$

As we see, Eq. (54), similarly to Eq. (53), is of the form of one-dimensional harmonic oscillator, and its solutions are expressed in terms of one-dimensional harmonic oscillator wave functions and energies:

$$\phi_2(y_2) = \phi_{n_2}(y_2) = N_{n_2} H_{n_2} \left(\sqrt[4]{\lambda_2 - \frac{\kappa^2}{\lambda_1}} y_2 \right) \exp \left[-\frac{1}{2} \sqrt{\lambda_2 - \frac{\kappa^2}{\lambda_1}} y_2^2 \right]$$

$$\varepsilon_{n_1, n_2} = \sqrt{\lambda_1} \left(1 + \frac{1}{2} \frac{\kappa^2}{\lambda_1^2} \right) \left(n_1 + \frac{1}{2} \right) + \sqrt{\lambda_2 - \frac{\kappa^2}{\lambda_1}} \left(n_2 + \frac{1}{2} \right) \quad \text{where} \quad n_1, n_2 = 0, 1, 2, \dots \quad (55)$$

In this way we get the solution of the Schrödinger equation (48) in the adiabatic approximation.

Step 2: Exact solution. The second step of the task is the exact solution of the eigenvalue problem of (48). By introducing new coordinates, which diagonalize the quadratic form $\lambda_1 y_1^2 + \lambda_2 y_2^2 + 2\kappa y_1 y_2$, we find

$$\text{exact} E_{n_1, n_2} = \sqrt{\text{exact} \lambda_1} \left(n_1 + \frac{1}{2} \right) + \sqrt{\text{exact} \lambda_2} \left(n_2 + \frac{1}{2} \right) \quad (56)$$

$$\text{exact} \Psi_{n_1, n_2}(y_1, y_2) = \Psi_{n_1}(q_1) \Psi_{n_2}(q_2) \quad (57)$$

where

$$\text{exact } \lambda_{1,2} = \frac{1}{2} \left[(\lambda_1 + \lambda_2) \pm \sqrt{(\lambda_1 - \lambda_2)^2 + 4\kappa^2} \right].$$

A comparison of the results

1. Let $(\lambda_1 - \lambda_2)^2 \gg 4\kappa^2$ and $\lambda_1 \gg \lambda_2$

The Condition $(\lambda_1 - \lambda_2)^2 \gg 4\kappa^2$ allows us to expand $\text{exact } \lambda_{1,2}$ in the exact energy expression (56) to a power series:

$$\begin{aligned} \text{exact } \lambda_{1,2} &= \frac{1}{2} \left[(\lambda_1 + \lambda_2) \pm \sqrt{(\lambda_1 - \lambda_2)^2 + 4\kappa^2} \right] \\ &= \frac{1}{2} \left[(\lambda_1 + \lambda_2) \pm (\lambda_1 - \lambda_2) \sqrt{1 + \left(\frac{2\kappa}{(\lambda_1 - \lambda_2)} \right)^2} \right] \\ &\cong \frac{1}{2} \left[\lambda_1 + \lambda_2 \pm (\lambda_1 - \lambda_2) \left(1 + \frac{2\kappa^2}{(\lambda_1 - \lambda_2)^2} + \dots \right) \right] \\ &\cong \frac{1}{2} \left[\lambda_1 + \lambda_2 \pm (\lambda_1 - \lambda_2) \left(1 + \frac{1}{2} \left(\frac{2\kappa}{(\lambda_1 - \lambda_2)} \right)^2 + \dots \right) \right] \\ \text{exact } \lambda_1 &\cong \lambda_1 + \frac{\kappa^2}{\lambda_1 - \lambda_2} \quad \text{and} \quad \text{exact } \lambda_2 \cong \lambda_2 - \frac{\kappa^2}{\lambda_1 - \lambda_2}. \end{aligned}$$

Assuming now that $\lambda_1 \gg \lambda_2$, we obtain

$$\text{exact } \lambda_1 \cong \lambda_1 + \frac{\kappa^2}{\lambda_1} \quad \text{and} \quad \text{exact } \lambda_2 \cong \lambda_2 - \frac{\kappa^2}{\lambda_1}. \quad (58)$$

Corollary:

When the frequencies of both oscillators are very different from each other and the first frequency is much larger than the second one (cf. the relative mobility of electrons and nuclei), then the energies obtained by the approximate and by the exact procedures remain in an excellent agreement (cf. (55) and (56)):

$$\begin{aligned} \varepsilon_{n_1, n_2} &= \sqrt{\lambda_1 \left(1 + \frac{1}{2} \frac{\kappa^2}{\lambda_1^2} \right)^2} \left(n_1 + \frac{1}{2} \right) + \sqrt{\lambda_2 - \frac{\kappa^2}{\lambda_1}} \left(n_2 + \frac{1}{2} \right) \\ &\cong \sqrt{\lambda_1 + \frac{\kappa^2}{\lambda_1}} \left(n_1 + \frac{1}{2} \right) + \sqrt{\lambda_2 - \frac{\kappa^2}{\lambda_1}} \left(n_2 + \frac{1}{2} \right) \\ \text{exact } E_{n_1, n_2} &= \sqrt{\text{exact } \lambda_1} \left(n_1 + \frac{1}{2} \right) + \sqrt{\text{exact } \lambda_2} \left(n_2 + \frac{1}{2} \right) \\ &\cong \sqrt{\lambda_1 + \frac{\kappa^2}{\lambda_1}} \left(n_1 + \frac{1}{2} \right) + \sqrt{\lambda_2 - \frac{\kappa^2}{\lambda_1}} \left(n_2 + \frac{1}{2} \right) \end{aligned}$$

In this case the adiabatic approximation works very well!

2. Let $k_1 = k_2 = k_{12}$

The conditions $(\lambda_1 - \lambda_2)^2 \gg 4\kappa^2$ and $\lambda_1 \gg \lambda_2$ where $k_1 = k_2 = k_{12}$ leads to

$$1 \ll \frac{\mu_2}{\mu_1}$$

Corollary:

The accuracy of the adiabatic approximation depends on the ratio of masses of the two oscillators. This means that in the case of molecules, in which the interaction between all particles have the same character, the accuracy of the adiabatic approximation depends on the ratio of masses of the particles. And because the mass ratio of the nucleus and the electron is of the order 10^3 – 10^5 , the condition under discussion is met.

3. Let $\mu_1 = \mu_2$

The condition $(\lambda_1 - \lambda_2)^2 \gg 4\kappa^2$ and $\lambda_1 \gg \lambda_2$ where $\mu_1 = \mu_2$ leads to

$$\frac{(k_1 - k_2)^2}{\mu^2} \gg \frac{k_{12}}{\mu^2} \quad \Rightarrow \quad k_1 \gg k_2$$

Corollary:

The accuracy of the adiabatic approximation depends on the values of the coupling constant k_{12} . The approximation is valid if the coupling constant is much smaller than the difference between k_1 and k_2 . This means that in a system where masses of the compound particles are the same, the accuracy of the adiabatic approximation depends mainly on the binding energy. For example, in the case of an atom with an infinite-mass nucleus, the accuracy of the adiabatic method is determined by the electron density distribution in the state under consideration.

Born-Oppenheimer Approximation

The equation of motion of nuclei (36) assuming

$$\sum_k c_{lk}(R) \chi_k(\mathbf{R}) \cong 0, \quad (59)$$

can be written as

$$\left(-\frac{1}{2\mu} \Delta_{\mathbf{R}} + E_l^0(R) \right) \chi_l(\mathbf{R}) = E \chi_l(\mathbf{R}). \quad (60)$$

It allows for a substantial simplification of the equations of motion of a molecular system and is the essence of BO approximation, named for its original creators, Max Born and Robert Oppenheimer (Born and Oppenheimer 1927).⁶

The omission of terms $\sum_k c_{lk}(R) \chi_k(\mathbf{R})$ in Eq. (36), except for the diagonal term $c_{ll}(R) \chi_l(\mathbf{R})$, corresponds to the adiabatic approximation. If also this term is neglected, we get the BO approximation. A description of molecules using the BO approximation ignores any kind of coupling between the movement of electrons and of the nuclei ($c_{ll}(R) = 0$). The physical idea is that light and fast-moving electrons adapt to the nuclear displacements instantaneously. If we solve the electronic Schrödinger equation for each possible arrangement of the nuclei, then we obtain the electronic effective potential $U_l(R) = E_l^0(R)$.

Equation (60) offers a possibility of examining the movement of electrons under the assumption that the nuclei in the molecule are immobile (i.e., their masses are infinite) and are located at a fixed distance $|\mathbf{R}_i|$ from each other. The function $E_l^0(R)$ determines the potential energy surface (PES). The nuclear wave function $\chi_l(\mathbf{R})$ determines the probability amplitude of the nuclear positions in the l -th quantum state. PES $E_l^0(R)$ is usually given numerically as a set of numbers evaluated for various nuclear positions. The shape of $E_l^0(R)$ for a diatomic molecule is shown in Fig. 2.

Electronic Terms of Diatomic Molecules: Symmetry

In atomic spectroscopy, the atomic electronic states are labeled by spectral term $^{2S+1}L_J$ indices. Usually the classification of atomic terms is performed according to the values of the total orbital angular momentum L and total spin angular momentum S , because the electronic Hamiltonian commutes with L^2 and S^2 (and also with L_z and S_z). In molecules the operator representing the square of the electronic angular momentum does not commute with the Hamiltonian because the electric field of several nuclei is not centrally symmetric. However, for diatomic molecules (or linear molecules in general), the field has axial symmetry about the axis passing through the two nuclei. Hence, the operator representing the angular momentum component along this axis commutes with the Hamiltonian. Therefore, the electronic terms of diatomic molecules are classified according to the projection of the electronic orbital angular momentum on the molecule axis. The absolute value of the projection of the orbital angular momentum on the axis of the molecule is usually denoted as Λ , where $\Lambda = 0, 1, 2, \dots$. The electronic

⁶More inventive faculty trying to explain to students the BO approximation takes an example of a cow (symbolizes the nucleus) and flies flying around it (electrons). Flies almost immediately adapt to the current position of the cow, just because they are lighter and move faster. Therefore, the cow only sees a cloud of flies, while the flies only see a static cow.

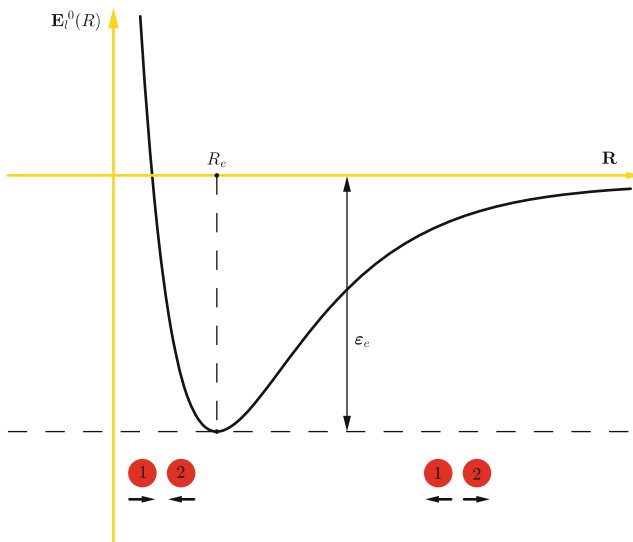


Fig. 2 Potential energy curve for a bound electronic state of a diatomic molecule. The graph shows the energy as a function of the interatomic distance. The case when R tends to infinity $E_l^{(0)}(R) = E_1 + E_2$ is equal to the energy of the isolated atoms 1 and 2. The value of R for which the potential reaches its minimum is called the equilibrium position and is denoted R_e . The binding energy of the molecule, $E_l^0(R = \infty) - E_l^0(R_e)$, is denoted as ϵ_e . For $R < R_e$, the nuclear repulsion energy dominates and $E_l^0(R)$ grows up to infinity while $R \rightarrow 0$

states with $\Lambda = 0$ are called Σ -states, those with $\Lambda = 1$ are Π -states, those with $\Lambda = 2$ are Δ -states, etc. The total spin S is denoted as an upper-left index: $^{2S+1}\Lambda$, where $2S + 1$ is equal to the spin degeneracy of the term. There is also a reflection symmetry about an arbitrary plane passing through the axis of symmetry of the molecule. In this case, the reflected and the original state of the molecule differ by sign of the angular momentum projection. This means that the electron terms with $\Lambda \neq 0$ (Π, Δ, Φ, \dots states) are double-degenerate. If $\Lambda = 0$ then the state of the molecule does not change under a reflection. Therefore, Σ terms are not degenerate. But because a double reflection in the same plane results in the identity transformation, among Σ terms we distinguish the ones for which the wave function does not change, Σ^+ , and the ones for which it changes the sign, Σ^- . Homonuclear diatomic molecules have another symmetry called *parity*. The molecule has the symmetry center at the point bisecting the line joining the two nuclei. In such a case the Hamiltonian is invariant with respect to the change of sign of all electron coordinates. The wave functions which do not change under this transformation are called *even* and labeled by an index g . The wave function changing their sign is called *odd* and labeled by u . The electronic terms for homonuclear diatomic molecules are denoted as $^{2S+1}\Lambda_u$ or $^{2S+1}\Lambda_g$.

Rotations and Vibrations of Diatomic Molecules

In the previous section, we discussed only the electronic states of the molecule. The share of the kinetic energy of the nuclei in the total energy of the molecule is usually by two orders of magnitude smaller than that of the electronic energy. Nevertheless, the effects of the nuclear motion significantly influences the molecular structure and spectra. In this section we show that the shape of the potential energy curve determines rotation and vibration of the nuclei. For the sake of simplicity of the description, we usually consider rotation and vibration of a molecule as independent of each other. In reality these motions are coupled.

Separation of Rotation and Vibration

We are going to solve the equation describing the nuclear motion (41) within adiabatic approximation (because we do not neglect the c_{ll}) term:

$$\left(-\frac{1}{2\mu} \Delta_{\mathbf{R}} + U_l(R) \right) \chi_l(\mathbf{R}) = E \chi_l(\mathbf{R}) \quad \text{where} \quad U_l(R) = E_l^0(R) + c_{ll}(R).$$

The motion of the nuclei is described by the changes of the length and the direction of $\mathbf{R} = \mathbf{R}_1 - \mathbf{R}_2$. The change of the length of \mathbf{R} corresponds to the *vibrations* of the molecule and the change of its direction – to the *rotation* of the molecule. Since the potential in Eq. (36) depends only on $R = |\mathbf{R}|$, the problem described by this equation is spherically symmetric as, e.g., in the case of the hydrogen atom. Therefore, it may be conveniently solved by introducing spherical coordinates. Using the well-known expression for the Laplace operator in the spherical coordinates (r, θ, φ)

$$\Delta_{\mathbf{R}} = \frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial}{\partial R} \right) - \frac{\hat{\mathbf{J}}^2}{R^2},$$

where

$$\hat{\mathbf{J}}^2 = -\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}$$

is the square of the angular momentum operator, we can rewrite Eq. (36) in the form

$$\frac{1}{2\mu} \left[-\frac{1}{R^2} \frac{\partial}{\partial R} \left(R^2 \frac{\partial}{\partial R} \right) + \frac{\hat{\mathbf{J}}^2}{R^2} \right] \chi_l(\mathbf{R}) + U_l(R) \chi_l(\mathbf{R}) = E \chi_l(\mathbf{R}). \quad (61)$$

The first term in the square bracket corresponds to the kinetic energy of vibrations and the second term to the rotational energy. Since the Hamiltonian operator in (61)

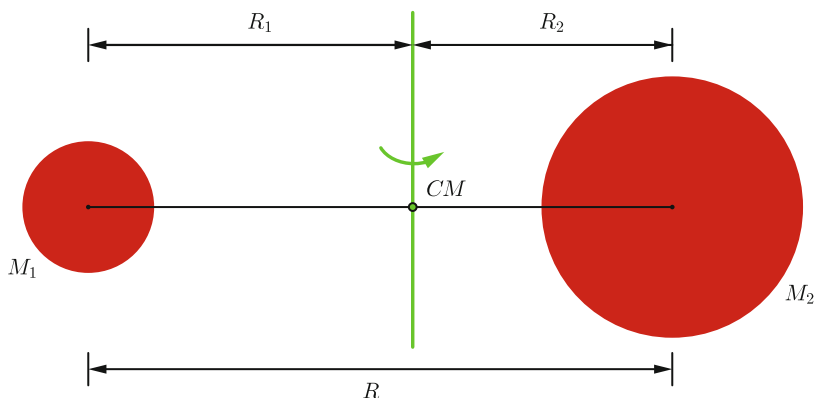


Fig. 3 Diatomic molecule as a rigid rotor. The rotation axis shown in the figure is perpendicular to the center of mass (CM) of the molecule

is a sum of operators which depend on different variables, the function $\chi_l(\mathbf{R})$ can be expressed as a product of two functions: the radial function $R^{-1}\chi_l^{\text{vib}}(R)$ and the angular function $Y_{JM}(\theta, \varphi)$

$$\chi_l(\mathbf{R}) = R^{-1}\chi_l^{\text{vib}}(R)Y_{JM}(\theta, \varphi). \quad (62)$$

The angular function is an eigenfunction of the square of the angular momentum operator

$$\hat{\mathbf{J}}^2 Y_{JM}(\theta, \varphi) = J(J + 1) Y_{JM}(\theta, \varphi),$$

where $J = 0, 1, \dots$ is the rotation quantum number.

A diatomic molecule with nuclear masses M_1 and M_2 can rotate around any axis passing through the center of the mass. The rotation energy of such a molecule is

$$E^{\text{rot}} = \frac{1}{2}I\omega^2, \quad (63)$$

where ω is the angular velocity and I (see Fig. 3) is the moment of inertia of the molecule. If the rotation axis is perpendicular to the axis of the molecule, then $I = M_1M_2/(M_1 + M_2)R^2$. Combining Eqs. (62) and (61) and using the eigenvalue equation for the rigid rotor (Appendix B), we get

$$\left[-\frac{1}{2\mu} \frac{\partial^2}{\partial R^2} + (U_l(R) - E) \right] \chi_l^{\text{vib}}(R) = -\frac{1}{2I} J(J + 1) \chi_l^{\text{vib}}(R). \quad (64)$$

The last equation may be rewritten in a more compact form as

$$\left[-\frac{1}{2\mu} \frac{\partial^2}{\partial R^2} + V_{lJ}(R) \right] \chi_{lvJ}^{\text{vib}}(R) = E_{lvJ} \chi_{lvJ}^{\text{vib}}(R), \quad (65)$$

where

$$V_{lJ}(R) = E_l^0(R) + c_{ll}(R) + \frac{1}{2I} J(J+1) \quad (66)$$

is the effective potential for the nuclear motion. As one can see, it is a sum of the adiabatic potential $U_l(R)$ and the centrifugal potential associated with the rotation of the molecule. In the BO approximation, $U_l(R)$ is replaced by $E^0(R)$:

$$V_{lJ}(R) = E_l^0(R) + \frac{1}{2I} J(J+1), \quad (67)$$

the potential generated by the cloud of electrons. The consecutive eigenfunctions $\chi_{lvJ}^{\text{vib}}(R)$ and eigenvalues E_{lvJ} are labeled by the vibrational quantum number $v = 0, 1, 2, \dots$.

$$\left[-\frac{1}{2\mu} \frac{\partial^2}{\partial R^2} + V_{lJ}(R) \right] \chi_{lvJ}^{\text{vib}}(R) = E_{lvJ} \chi_{lvJ}^{\text{vib}}(R)$$

where

$$V_{lJ}(R) = E_l^0(R) + c_{ll}(R) + J(J+1) \frac{1}{2I}, \quad (68)$$

in the adiabatic approximation, and

$$V_{lJ}(R) = E_l^0(R) + J(J+1) \frac{1}{2I}, \quad (69)$$

in the BO approximation.

- J – rotational quantum number.
- $v = 0, 1, 2, 3, \dots$ – vibrational quantum number.
- $V_{lJ}(R)$ represents the potential energy of the motion of the nuclei; depends only on the nuclear distance R , i.e., it is independent of the orientation of the molecule in space. It is spherically symmetric.
- The centrifugal term in Eq. (66) describes the dependence of the molecular rotation. The rotation energy is equal to $J(J+1)$. It describes the weakening of bond energy due to the centrifugal force.

Nuclear Vibrations: Harmonic Approximation

In the case of the low-energy states, more precisely the states with low energies, bound in sufficiently deep minima of the potential energy curves, the motion of the nuclei is limited to a small range of R , close to the equilibrium distance R_e .⁷ At $R = R_e$ the potential $V_{l(J=0)}(R) \equiv U_l(R)$ has minimum, so $R = R_e$ is the value for which the interaction energy of the atoms in the rotor has a minimum. Expanding $U_l(R)$ into a power series around $R = R_e$, we get

$$U_l(R) = U_l(R)|_{R=R_e} + \left. \frac{dU_l(R)}{dR} \right|_{R=R_e} (R-R_e) + \frac{1}{2!} \left. \frac{d^2U_l(R)}{dR^2} \right|_{R=R_e} (R-R_e)^2 + \dots \quad (70)$$

Since at $R = R_e$ the potential $U_l(R)$ reaches a minimum, and limiting the expansion (70) up to quadratic terms, we have

$$\left. \frac{dU_l(R)}{dR} \right|_{R=R_e} = 0 \quad \Rightarrow \quad U_l(R) \approx U_l(R_e) + \frac{1}{2!} \kappa (R - R_e)^2, \quad (71)$$

where

$$\kappa = \left. \frac{d^2U_l(R)}{dR^2} \right|_{R=R_e} \quad (72)$$

is called force constant. For $R = R_e$ and $J \neq 0$, we can assume that

$$J(J+1) \frac{1}{2\mu R^2} \approx J(J+1) \frac{1}{2\mu R_e^2} \equiv J(J+1) B_e, \quad (73)$$

where B_e is the *rotational constant* determined by the reduced mass μ and the equilibrium nuclear distance R_e . Thus, the potential in Eq. (68) can be written as

$$V_{lJ}(R) = U_l(R_e) + \frac{1}{2} \kappa (R - R_e)^2 + J(J+1) B_e. \quad (74)$$

The first and last terms in this equation are constant. We denote them as

$$E' \equiv U_l(R_e) + J(J+1) B_e. \quad (75)$$

With the above approximations, Eq. (65) becomes

⁷In fact, there exist rovibrational states, e.g., in the helium dimer, which cannot be properly described by the model of a harmonic oscillator.

$$\left[-\frac{1}{2\mu} \frac{\partial^2}{\partial R^2} + \frac{1}{2} \kappa (R - R_e)^2 \right] \chi_{l\nu J}^{\text{vib}}(R) = (E_{l\nu J} - E') \chi_{l\nu J}^{\text{vib}}(R). \quad (76)$$

It is easy to recognize this equation as the Schrödinger equation for the harmonic oscillator, where $R - R_e$ is the deviation from the equilibrium position. Hence, the energies (the eigenvalues of (76)) may be readily written as

$$E_{l\nu J} = E' + E_{\nu}^{\text{vib}}, \quad \text{where} \quad E_{\nu}^{\text{vib}} = \omega \left(\nu + \frac{1}{2} \right), \quad \nu = 0, 1, 2, \dots, \quad (77)$$

where $\nu = 0, 1, 2, \dots$ is vibrational quantum number and ω is the classical frequency of the oscillator. In the molecular spectroscopy, the term values rather than energies are frequently used:

$$G_{\nu} = \frac{E_{\nu}^{\text{vib}}}{hc} = \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right), \quad \text{where} \quad \bar{\nu}_{\text{vib}} = \frac{1}{2\pi c} \sqrt{\frac{\kappa}{\mu}} \quad (78)$$

and $G_0 = 1/2\bar{\nu}_{\text{vib}}$ is called *zero-energy vibration*.

Equations (75) and (77) yield the following symbolic expression for the energy levels in a molecule:

$$E_{l\nu J} = E_l^{\text{el}} + E_{\nu}^{\text{vib}} + E_J^{\text{rot}}. \quad (79)$$

The first term, E_l^{el} (see (41)) is the electronic energy, the second term is the energy of vibrations of the nuclei (77), and the third one is the rotational energy (75).

The molecular energy within adiabatic approximation and harmonic decoupling of vibronic and rotation motions model, may be expressed as a sum of the electronic, vibrational, and rotational components:

$$\begin{aligned} E_{l\nu J}(R) &= \underbrace{U_l(R_e)}_{E_l^{\text{el}}} + \underbrace{\bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right)}_{E_{\nu}^{\text{vib}}} + \underbrace{J(J+1)B}_{E_J^{\text{rot}}} = \\ &= E_l^0(R) + c_{ll}(R) + \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right) + J(J+1)B. \end{aligned} \quad (80)$$

This expression for the molecular energy levels is correct only for the low excited states which lie sufficiently close to the of ground level where the assumption on the harmonic character of the nuclear vibrations and the assumption about the separability of the vibrational and rotational motions are valid. We should also remember that in Eq. (80), the kinetic energy of the molecular center of mass is omitted.

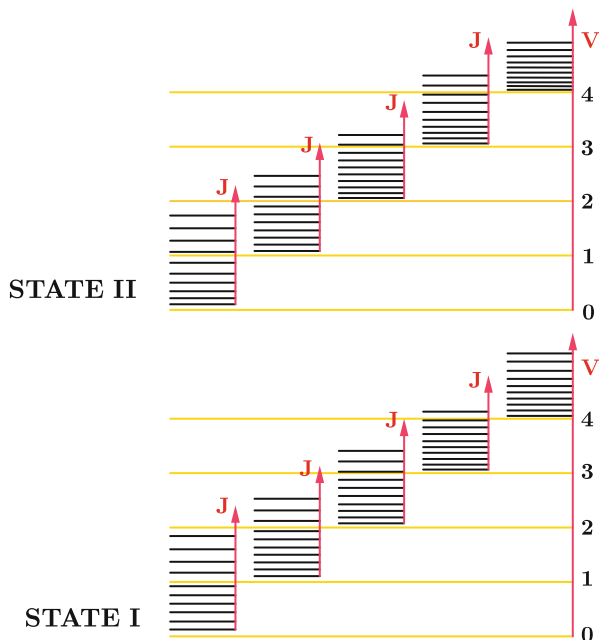


Fig. 4 Vibrational levels (labeled by the vibrational quantum numbers $\nu = 0, 1, 2, 3, \dots$) and rotational levels (quantum numbers J) of two electronic states of a molecule, denoted as “State I” and “State II.” More details on the structure of molecular spectra may be found, e.g., in Landau and Lifschitz (1981) and Haken and Wolf (2004)

According to Eq. (80), the nuclear motion causes a splitting of each electronic term to a series of energy levels characterized by quantum numbers ν and J . For each value of J , we have infinite number of levels with different values of ν , cf. Fig. 4. One can show Landau and Lifschitz (1981), Haken and Wolf (2004), and Herzberg (1951) that

$$\{E_l^{\text{el}} : E_\nu^{\text{vib}} : E_J^{\text{rot}}\} = \left\{ 1 : \sqrt{\frac{m}{M}} : \frac{m}{M} \right\},$$

where m and M are, respectively, the masses of electron and of the nucleus. Assuming that $m/M \sim 10^{-4} - 10^{-5}$, we obtain the distribution of the molecular energy levels given by the following inequality

$$\Delta E^{\text{el}} \gg \Delta E^{\text{vib}} \gg \Delta E^{\text{rot}}. \quad (81)$$

A separation of rotation and vibration in higher-order approximations is no longer possible. In higher orders of the perturbation theory, the molecular energy may be expressed as an expansion in powers of the quantum numbers ν and J (Demtröder 2010).

A More Detailed Look at the Molecule

In the following paragraphs, we present a more detailed picture of a molecule going beyond the rigid molecule approximation and beyond the harmonic description of the molecular vibrations.

Nuclear Vibrations: Anharmonic Approximation

The harmonic potential represented by a symmetric parabola, cannot correctly represent the real potential $E_i^{el}(R)$ across the variability of R . In this model the potential and the repulsion force increase indefinitely with the increasing distance from the equilibrium position. This cannot be correct for a real molecule because when $R \rightarrow \infty$, i.e., if the distance between the atoms is very large, the attractive force between them must be zero, and the potential has to be constant. This asymptotic constant value is equal to the total energy of the separated atoms. For $R < R_e$, the potential energy must be steeper than in the case of the harmonic oscillator, because reduction of the distance between atoms leads to a short-distance repulsive force, which exceeds the Coulomb interaction. These characteristics of the potential of the molecule is partly fulfilled by the Morse potential Fig. 5, which has the following functional form:

$$E(R) = D_e \{ 1 - \exp[-a(R - R_e)] \}^2, \quad (82)$$

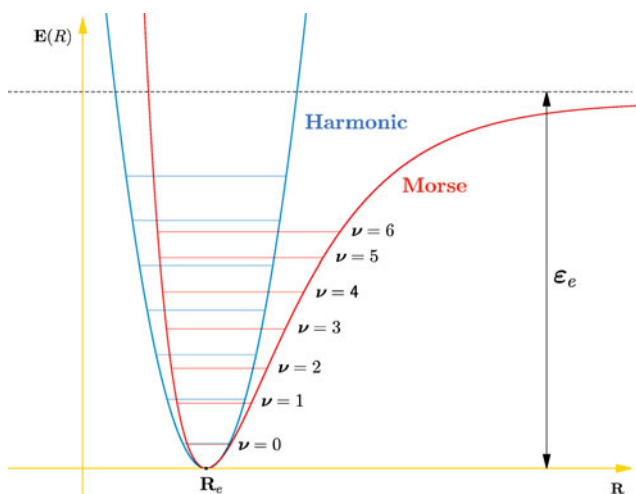


Fig. 5 The potential curve is approximately harmonic (a parabola) near the bottom. A better, and still analytic but anharmonic, potential is the Morse potential. The vibrational energy levels ν for both potentials and the dissociation energy ϵ_e for the Morse potential are also shown

where D_e is the dissociation energy and a determines the width of the potential well and is characteristic for the pair of atoms forming the molecule. One of advantages of the Morse potential is that for this potential, the Schrödinger equation is analytically solvable. As one can see, near the minimum of the potential energy curve the harmonic oscillator gives a good approximation, while at large distances from the equilibrium, $R > R_e$, one has to use an anharmonic potential. In particular, it may be relatively well described by the Morse potential. In the general case, the anharmonicity modifies the vibrational energy levels (77) to form (Herzberg 1951):

$$E_v^{\text{vib}} = \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right) - x_e \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right)^2 + y_e \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right)^3 + \dots, \quad (83)$$

where x_e , y_e are anharmonic constants, usually much smaller than the harmonic one, $\bar{\nu}_{\text{vib}}$. The anharmonic constants are always positive (Haken and Wolf 2004; Kolos and Sadlej 1998). In this way we retained the Morse potential qualitative description of the vibration spectrum. It should be remembered that in the case of high-precision calculations, one has to use more accurate potential functions. For example, good results may be obtained with a five-parameter Hulburt-Hirschfelder function (Hulburt and Hirschfelder 1941) or with the Taylor series expansion of the exact potential to high powers of $R - R_e$:

$$E(R) = D_e \{ [1 - \exp[-x]]^2 + c[x]^3 \exp[-2x] (1 + bx) \}, \quad \text{for } x = a(R - R_e), \quad (84)$$

where b, c are additional parameters that allow us to adjust the curve to the experimental data.

Rotations: Non-rigid Rotor

As we have shown (80), the energy of a molecule treated as a rigid rotor may be expressed as a sum of the electronic energy at the equilibrium positions of the nuclei, the energy of vibration, and the energy of rotation. In the rigid rotor approximation, the rotational energy of a molecule may be expressed as $E_J^{\text{rot}} = J(J + 1)B_e$. Let us note that a molecule can vibrate without rotation (the states with $J = 0$). However, it cannot rotate without vibrations, because even in the lowest energy vibrational state, it executes the $\nu = 0$ vibrations. The real molecule rotates and vibrates at the same time. Since the vibrational frequency $\bar{\nu}_{\text{vib}}$ is larger than the rotational constant B_e , the nuclear distance changes periodically during one full rotation. It follows that in describing molecular rotations, we should consider them as an average over a vibrational period. Therefore, instead of $1/R_e^2$ in Eq. (73), we should take the average value $1/R^2$, calculated for a given vibrational state:

$$B_e \rightarrow B_v = \frac{1}{2\mu R^2} \langle \chi_v^{\text{vib}}(R) \left| \frac{1}{R^2} \right| \chi_v^{\text{vib}}(R) \rangle, \quad (85)$$

where the integral is calculated with respect to the nuclear variables. Thus, the rotational constant B from Eq.(73)) is now defined as the average over the vibrational motion. In this way the rotational energy may be expressed as the average over the vibrational period

$$E_J^{\text{rot}} = \frac{1}{2\mu} J(J+1) \langle \chi_v^{\text{vib}}(R) \left| \frac{1}{R^2} \right| \chi_v^{\text{vib}}(R) \rangle. \quad (86)$$

As one can see from Eq. (85), the rotational constant depends on v . As a result, now the potential is not symmetric around the point $R = R_e$. For a general anharmonic potential, the rotational constant may be expressed as a v -dependent power series (Herzberg 1951):

$$B_v = B_e - \alpha_e \left(v + \frac{1}{2} \right) + \beta_e \left(v + \frac{1}{2} \right)^2 + \gamma_e \left(v + \frac{1}{2} \right)^3 + \dots, \quad (87)$$

where $B_e \gg \alpha_e \gg |\beta_e| \gg |\gamma_e|$ are empirically determined constants. Constants B_e and α_e are positive, but the subsequent terms can be positive or negative.

The introduction of different rotational constants for each vibrational state resulted in important amendments. However, we still remain in the rigid rotor model. The next step to correct the description of the molecular rotation is the assumption that the bond length is flexible. The degree of the flexibility is described by the force constant κ . This means that the degree to which bond length increases due to the centrifugal force is controlled by κ . The centrifugal force increases with increasing J and therefore also the bond length is larger for larger J . In the simplest model of the harmonic vibrations, the dependence of R_{min} on J is given by;

$$\frac{d}{dr} \left[\frac{1}{2} \kappa (R - R_e)^2 + \frac{J(J+1)}{2\mu R_e^2} \right]_{R=R_{\text{min}}} = \kappa (R_{\text{min}} - R_e) - \frac{J(J+1)}{\mu R_{\text{min}}^3} = 0 \quad (88)$$

$$\Rightarrow R_{\text{min}} = R_e \left(1 + \frac{J(J+1)}{\kappa \mu R_e^4} \right) \equiv R_e (1 + x), \quad \text{with} \quad x \ll 1, \quad (89)$$

where R_{min} corresponds to the minimum of potential (88) (the harmonic potential supplemented by a term describing the centrifugal force. As a result, the rotational energy of the non-rigid rotor is then

$$E_J^{\text{rot}} = \frac{1}{2} \kappa (R_{\text{min}} - R_e)^2 + \frac{J(J+1)}{2\mu R_{\text{min}}^2} = \frac{1}{2} \kappa R_e^2 x^2 + \frac{J(J+1)}{2\mu R_e^2} (1+x)^{-2}. \quad (90)$$

Inserting the power series expansion $(1 + x)^{-2} \approx 1 - 2x + 3x^2 + \dots$ to the last equation, we get

$$\begin{aligned} E_J^{\text{rot}} &\approx \frac{1}{2} \kappa R_e^2 x^2 + \frac{J(J+1)}{2\mu R_e^2} - 2 \frac{J(J+1)}{2\mu R_e^2} x + 3 \frac{J(J+1)}{2\mu R_e^2} x^2 \pm \dots \\ &= \frac{J(J+1)}{2\mu R_e^2} - \frac{J^2(J+1)^2}{2\mu^2 \kappa R_e^6} + \frac{3J^3(J+1)^3}{2\mu^3 \kappa^2 R_e^{10}} \pm \dots \end{aligned} \quad (91)$$

The resulting expression for the modified rotational energy of the molecule may be rewritten in a compact way as

$$E_J^{\text{rot}}(R) = J(J+1)B_v - J^2(J+1)^2 D_v + J^3(J+1)^3 H_v + \dots, \quad (92)$$

with the rotational constants

$$D_v = \frac{1}{2\mu^2 \kappa} \langle \chi_v^{\text{vib}}(R) | \frac{1}{R^6} | \chi_v^{\text{vib}}(R) \rangle, \quad H_v = \frac{1}{2\mu^3 \kappa^2} \langle \chi_v^{\text{vib}}(R) | \frac{1}{R^{10}} | \chi_v^{\text{vib}}(R) \rangle, \quad (93)$$

where we have introduced an average centrifugal constant D_v . For the general potential, the average centrifugal constant may be defined as

$$D_v = D_e + \tilde{\alpha}_e \left(\nu + \frac{1}{2} \right) + \tilde{\beta}_e \left(\nu + \frac{1}{2} \right)^2 + \tilde{\gamma}_e \left(\nu + \frac{1}{2} \right)^3 + \dots \quad (94)$$

The resulting expression (93) leads to the improved rotational energy (actually rotational-vibrational energy, because constants B_μ and D_μ depend on ν). We derived corrections to the energy, associated with stretching the bond in the molecule. These corrections may be neglected for small values of J . In turn, they become important for large J .

Collecting the results of this and of the previous sections, that is taking into account Eqs. (83) and (93), the energy of a rotational-vibrational level, for an anharmonic potential, can then be expressed as the power series of $(\nu + 1/2)$ and $J(J+1)$ as follows

$$E_{\nu J}^{\text{rot-vib}} = \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right) - x_e \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right)^2 + J(J+1)B_v - J^2(J+1)^2 D_v, \quad (95)$$

where the higher-order terms are neglected.

The Influence of the Electronic Motion

In order to refine the description of the rotational movement of molecules, we have to consider the contribution of the electronic motion to the energy of rotation of molecules. The reasoning presented here applies when the magnitudes of the angular

momenta of electrons and of the nuclei are comparable. Electrons process in the cylindrically symmetric potential of the nuclei. In this discussion the influence of the electronic spin is omitted. In this case, the quantity conserved is the projection of the electronic orbital angular momentum on the molecular axis. Then, the molecule can be seen as a symmetric top.⁸ In this model, two main components of the moment of inertia tensor are equal to each other (they are denoted as $I_B = I_C$), while the third one, describing the moment of inertia of electrons relative to the axis A of the molecule (I_A) is determined by the distribution of the electronic charge. Because the light electrons move much faster than heavy nuclei, the angular momentum of electrons is comparable to the angular momentum of the nuclei. Thus, the part of the Hamiltonian describing rotation may be written as

$$\hat{H}^{\text{rot}} = \frac{\hat{J}_A^2}{2I_A} + \frac{\hat{J}_B^2}{2I_B} + \frac{\hat{J}_C^2}{2I_C}. \quad (96)$$

The eigenvalues of the Hamiltonian (96) are equal to the total rotational energy of the molecule. In order to find these eigenvalues, we introduce the total angular momentum

$$\hat{J}^2 = \hat{J}_A^2 + \hat{J}_B^2 + \hat{J}_C^2, \quad \text{where} \quad |\mathbf{J}| = \sqrt{J(J+1)} \quad (97)$$

and its projection on the molecular axis A

$$\hat{J}_A^2, \quad \text{where} \quad |\mathbf{J}|_A = \Lambda \quad (98)$$

describes the contribution from the electronic motion. The quantum number Λ refers to the projection of the electronic angular momentum on the molecular axis A . Hence, Eq. (96) can be written as

$$\hat{H}^{\text{rot}} = \frac{\hat{J}^2}{2I_B} + \hat{J}_A^2 \left(\frac{1}{2I_A} - \frac{1}{2I_B} \right), \quad (99)$$

and the molecular rotational energy is given by

$$\begin{aligned} E^{\text{rot}} &= J(J+1) \frac{1}{2I_B} + \Lambda^2 \left(\frac{1}{2I_A} - \frac{1}{2I_B} \right) = J(J+1)B_e + \Lambda^2(\tilde{A} - B_e) \\ &= B_e [J(J+1) - \Lambda^2] + \Lambda^2 \tilde{A}, \end{aligned} \quad (100)$$

where B_e it the rotational constants introduced earlier in (73) and

$$\tilde{A} = \frac{1}{2I_A}. \quad (101)$$

⁸This model works well for molecules like NH_3 , CH_3Cl , C_6H_6 . In a general case the asymmetric top model should be used (see Haken and Wolf 2004 for details).

Because the term $\Lambda^2 \tilde{A}$ is constant for a given electronic state, it does not depend on J . Therefore, it may be added to the electronic energy. As a result, the rotational energy may be expressed as

$$E^{\text{rot}} = B_e [J(J + 1) - \Lambda^2], \quad (102)$$

where

$$J = \Lambda, \Lambda + 1, \Lambda + 2, \dots \quad (103)$$

The energy of a diatomic molecule obtained within the adiabatic approximation i.e., neglecting terms coupling the nuclear and the electronic motion ($c_{kl}(R) = 0$ for $k \neq l$), on different levels of approximation reads:

- Molecule as a rigid rotor, assuming harmonic vibrations of the nuclei

$$E_{l\nu J}(R) = E_l^0(R) + c_{ll}(R) + \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right) + J(J + 1)B_e \quad (104)$$

- Molecule as a non-rigid rotor, assuming anharmonic vibrations of the nuclei

$$E_{l\nu J}(R) = E_l^0(R) + c_{ll}(R) + \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right) - x_e \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right)^2 + J(J + 1)B_v - J^2(J + 1)^2 D_v \quad (105)$$

- Taking into account the rotational motion of electrons in the potential of the nuclei in the harmonic approximation

$$E_{l\nu J}(R) = E_l^0(R) + c_{ll}(R) + \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right) + [J(J + 1) - \Lambda^2] B_e \quad (106)$$

- Taking into account the rotational motion of electrons in the potential of the nuclei, for the anharmonic potential

$$E_{l\nu J}(R) = E_l^0(R) + c_{ll}(R) + \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right) - x_e \bar{\nu}_{\text{vib}} \left(\nu + \frac{1}{2} \right)^2 + [J(J + 1) - \Lambda^2] B_e \quad (107)$$

The Collapse of the Adiabatic Approximation

When does the Born-Oppenheimer approximation break down? The intuitive and general answer must be: adiabatic and BO approximation collapses whenever the electrons cannot follow the nuclear motion instantaneously, when the nuclei are

moving too fast, or the electrons cannot move fast enough. The consequence of such behavior is mixing of electronic and nuclear states of the molecule. Wave functions mix when there is a small or negligible energy difference between the mixing states. A typical situation, when the adiabatic and BO approximation is collapsing, is shown in Fig. 6a.

The starting point for our discussion is the complete Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}' = \hat{H}_0 + \lambda \hat{W},$$

where $\lambda \hat{W}$, defined in Eq. (19), is a small perturbation. Parameter $\lambda < 1$ determines the magnitude of the perturbation. It depends on the ratio m/M of the electron, m , and the nuclear, M , masses. The adiabatic wave function defined in Eq. (42)

$$\Psi_{lvJ}^{\text{ad}}(\mathbf{r}, \mathbf{R}) = \psi_l^{\text{el}}(\mathbf{r}; R) \chi_{lvJ}(\mathbf{R}) \equiv \Psi^{\text{ad}}. \quad (108)$$

is an eigenfunction of the unperturbed Hamiltonian \hat{H}_0 which is a sum of two Hamiltonians. One of them acts in the electronic coordinate space where the nuclear coordinates are parameters. The eigenfunctions of this Hamiltonian are $\psi_l^{\text{el}}(\mathbf{r}; R)$. The second one acts in the nuclear coordinate space. Its eigenfunctions are $\chi_{lvJ}(\mathbf{R})$. The product functions (108) are the total unperturbed wave functions.

$$\Psi^{(0)} \equiv \Psi^{\text{ad}}. \quad (109)$$

The first-order energy is equal to the expectation value of the complete Hamiltonian \hat{H} . The first-order correction to the wave function, according to the perturbation theory, can be determined as

$$\Psi_l^{(1)} = \sum_{l \neq k} d_{lk} \Psi_k^{(0)}, \quad \text{where} \quad d_{lk} = \frac{\langle \psi_k^{\text{el}} \chi_{kvJ} | \hat{H}' | \psi_l^{\text{el}} \chi_{lv\tilde{J}} \rangle}{E_{kvJ}^{(0)} - E_{lv\tilde{J}}^{(0)}}. \quad (110)$$

The integration in Eq. (110) has to be performed over both nuclear and electronic coordinates. This expression is the basis for the formulation of the criterion for the adiabatic and for the BO approximation to be valid. It follows from Eq. (110) that, when $|d_{kl}| \ll 1$ then the non-adiabatic correction (correction to the BO energy) is only a small contribution to the adiabatic solution. However, the adiabatic approximation becomes invalid when the denominator $E_{kvJ}^{(0)} - E_{lv\tilde{J}}^{(0)}$ in d_{kl} is much smaller than the matrix element in the numerator, $\langle \psi_k^{\text{el}} \chi_{kvJ} | \hat{W} | \psi_l^{\text{el}} \chi_{lv\tilde{J}} \rangle$. From a physical standpoint, the adiabatic approximation collapses, when the electronic energies $E_{kvJ}^{(0)}$ and $E_{lv\tilde{J}}^{(0)}$ are not separated. This happens when the electronic quantum states l and k become degenerate or nearly degenerate. However, it may happen that also term (110) is zero. This situation occurs, for example, when the vibrational functions are in different regions of R and the overlap between χ_{lvJ} and $\chi_{k\tilde{v}\tilde{J}}$ is zero.

To analyze such a case, let us consider a diatomic molecule (40). Non-adiabatic coupling, which breaks the BO approximation, is in $c_{kl}(R)$. The non-adiabatic coupling arises because the electronic wave functions and the electronic Hamiltonian change with R . We can get rid of the R dependence by choosing a fixed set of basis functions for electrons. We assume that the basis is complete. Any complete fixed basis set of electronic states which describes the non-degenerate eigenstates of Hamiltonian $\hat{H}_0(R)$ (unperturbed operator) is called a diabatic basis (*dia*) (Piela 2007). The choice of such a basis set reduces (40), because, if the diabatic electronic states do not depend on R , the nuclear operators disappear. The electronic Hamiltonian is more complicated. To illustrate this problem, it is useful to work an example.

Consider a fixed point R_0 . For R_0 the electronic Hamiltonian, which depends on R , is $\hat{H}_0(R_0)$. Let Hamiltonian $\hat{H}_0(R_0)$ have two near-degenerate, but different eigenvalues $E_1(R_0)$ and $E_2(R_0)$:

$$\hat{H}_0(R_0) \psi_1^{\text{dia}}(\mathbf{r}; R_0) = E_1(R_0) \psi_1^{\text{dia}}(\mathbf{r}; R_0) \quad (111)$$

$$\hat{H}_0(R_0) \psi_2^{\text{dia}}(\mathbf{r}; R_0) = E_2(R_0) \psi_2^{\text{dia}}(\mathbf{r}; R_0). \quad (112)$$

Energies $E_1(R_0)$ and $E_2(R_0)$ are called the diabatic energies. Since $\hat{H}_0(R_0)$ is Hermitian and $E_1(R_0) \neq E_2(R_0)$, the diabatic functions $\psi_i^{\text{dia}}(\mathbf{r}; R_0)$ for $i = 1, 2$ satisfy the orthogonality condition $\langle \psi_i^{\text{dia}} | \psi_j^{\text{dia}} \rangle = \delta_{ij}$. Let us now examine whether changing the distance $|\mathbf{R}_0|$ by δR will bring a situation in which the electronic energies, E_1 and E_2 , have the same values. If we change the distance R_0 by δR , the Hamiltonian \hat{H}_0 becomes

$$\hat{H}_0(R_0 + \delta R) = \hat{H}_0(R_0) + \delta R \frac{\partial \hat{H}_0}{\partial R} \equiv \hat{H}_0(R_0) + \hat{V}(R_0 + \delta R). \quad (113)$$

$\hat{V}(R_0 + \delta R)$ is a small correction to the unperturbed Hamiltonian $\hat{H}_0(R_0)$. The eigenvalues of the new Hamiltonian $\hat{H}_0(R_0 + \delta R)$ can be calculated using the perturbation theory, where \hat{V} is regarded as a perturbation to \hat{H}_0 . To solve the eigenvalue problem with Hamiltonian (113), we cannot use the ordinary perturbation theory, because the eigenvalues E_1 and E_2 of the unperturbed problem are very close to each other. The difference $E_1 - E_2$ compared with the magnitude of the perturbation is very small and tends to zero. Hence, we need to employ the perturbation theory suitable for degenerate states (Landau and Lifschitz 1981). As the initial zero-order approximation for the eigenfunctions of (113), we take linear combinations of the diabatic states:

$$\psi^{\text{el}}(\mathbf{r}; R) = d_1(R) \psi_1^{\text{dia}}(\mathbf{r}; R_0) + d_2(R) \psi_2^{\text{dia}}(\mathbf{r}; R_0). \quad (114)$$

(Diabatic states ψ_1^{dia} and ψ_2^{dia} satisfy Eqs. (112) and (111).) Substituting expression (114) in the perturbed Schrödinger equation can be symbolically written as

$$(\hat{H}_0 + \hat{V}) \psi^{\text{el}}(\mathbf{r}; R) = E \psi^{\text{el}}(\mathbf{r}; R). \quad (115)$$

Multiplying this equation on the left by $\psi_1^{\text{dia}*}$ and $\psi_2^{\text{dia}*}$ and integrating, we obtain the following algebraic equations:

$$\begin{aligned} d_1(E_1 + V_{11} - E) + d_2 V_{12} &= 0 \\ d_1 V_{21} + d_2(E_2 + V_{22} - E) &= 0, \end{aligned} \quad (116)$$

where we used

$$\langle \psi_i^{\text{dia}} | \hat{H}_0 + \hat{V} | \psi_i^{\text{dia}} \rangle = E_i + V_{ii}, \quad \text{where} \quad V_{ij} \equiv \langle \psi_i^{\text{dia}} | \hat{V} | \psi_j^{\text{dia}} \rangle. \quad (117)$$

Since \hat{H}_0 and \hat{V} are Hermitian, $V_{12} = V_{21}^*$ and V_{11} and V_{22} are real. The condition for the nontrivial solution of Eqs. (116) is

$$\det \begin{bmatrix} E_1 + V_{11} - E & V_{12} \\ V_{21} & E_2 + V_{22} - E \end{bmatrix} = 0. \quad (118)$$

Hence we get that

$$E_{\pm}(R) = \frac{1}{2}(E_1 + E_2 + V_{11} + V_{22}) \pm \sqrt{\left(\frac{E_1 - E_2 + V_{11} - V_{22}}{2}\right)^2 + |V_{12}|^2}. \quad (119)$$

This is a quadratic equation with respect to E which has two real roots, E_+ and E_- . In general $E_1 + V_{11} \neq E_2 + V_{22}$ and $V_{12} \neq 0$. Hence the eigenvalues E_- and E_+ are different. When the curves intersect at some R , it means that $E_- = E_+$ and the term under the square root would need to be zero. Since the expression under the square root is a sum of two squares, the following two equations need to be satisfied:

$$E_1 - E_2 + V_{11} - V_{22} = 0 \quad \text{and} \quad V_{12} = 0. \quad (120)$$

There are two conditions, but there is only one degree of freedom – one arbitrary parameter – giving the perturbation \hat{V} having the magnitude proportional to δR . In general, it is not possible to satisfy two independent equations having only one parameter. The $E_- = E_+$ degeneracy may occur only if more than one parameter can be varied. One thus needs at least two degrees of freedom to get an intersection. It may happen that V_{12} vanishes identically. Then only one equation remains (120), which can be satisfied for appropriately selected δR . From a physical point of view, that is the case when the two terms have different symmetries. (By symmetry we mean all possible forms of the molecular symmetry including the symmetry of the rotational state, the electronic state, etc.) For a diatomic molecule, this means that the intersecting states correspond to different Λ 's, different parities, or multiplicity.

Referring to the groups theory, for this to happen ψ_1^{dia} and ψ_2^{dia} must belong to two different irreducible representations of Hamiltonian symmetry groups. In such a situation, we are talking about vibronic coupling, i.e., the coupled between the motions of the nuclei and the electrons. Thus, in a diatomic molecule, only terms of different symmetry can intersect⁹ and it is impossible to have an intersection of states with the same symmetry. If, as a result of some calculation, we obtain wave functions and energies of two intersecting states of the same symmetry, then the mixing of the two states, which produces two new states, can be described using the perturbation theory. The energy difference between the two new states is that the greater, the smaller is the $|E_1 - E_2|$ energy difference.

This situation is described in graphs (6). Figure 6a represents a situation when two electronic states intersect and mix. In accordance with the discussion conducted above, we know that this situation can happen only when the two states have different symmetry. In the figure it is assumed that the electronic states have energies E_{000} and E_{100} . This means that the matrix element V_{12} vanishes identically and eigenvalues of Hamiltonian $\hat{H}_0(R_0 + \delta R)$ are degenerate. In this case for $R_0 + \delta R$, “the nuclear motion” occurs on the one potential energy surface (in Fig. 6b on the dotted lines, which symbolize the vibrational levels). The total wave function of the Hamiltonian \hat{H}_0 is therefore a superposition of two BO states, the first state ($l = 0$) with an excited nuclear vibrational wave function ($\nu = 1, J = 0$ marked in Fig. 6b with the blue color) and the second state ($l = 1$) with the ground-state vibrational wave function ($\nu = 0, J = 0$) marked in Fig. 6b with the red color). At internuclear

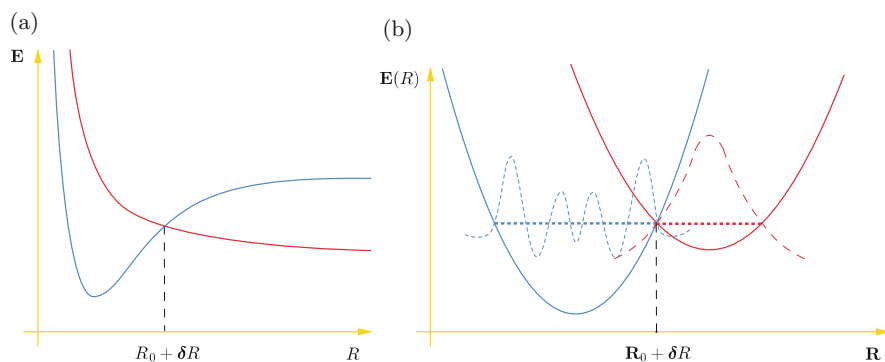


Fig. 6 Two diabatic states – a schematic view. The distance R , for which there is an intersection of the energy levels molecule, is designated as $R_0 + \delta R$. **(a)** Intersection of the electronic energy curves for states with different symmetries. **(b)** Two electronic states, in the harmonic approximation, and the wave functions corresponding to the crossing point energy, E_{000} and E_{100} . See text for details

⁹Crossing the states of the same symmetry is possible if you work within the adiabatic approximation.

distances R , the primary curves (corresponding to the ionic ground state and the covalent excited state) correspond to different diabatic functions, ψ_1^{dia} and ψ_2^{dia} .

What happens if the electronic states have the same symmetry? Then the adiabatic curves come close to each other, then nearly intersect (in Fig. 7a this happens at distance $R_0 + \delta R$) and then begin to separate (“avoided crossing”) to avoid an intersection. In the case when two states almost cross, we have near degeneracy. Classic examples of the curve non-crossing rule are the ground and first excited state potential curves for NaCl (Fig. 7). The original curves, ionic and covalent, correspond to diabatic states (see Fig. 7b). In accordance with Eq. (114), these states are represented by ψ_1^{dia} and ψ_2^{dia} , respectively. At small internuclear distances, wave function ψ_1^{dia} dominates in the total wave function (large $d_1(R)$). However, as it is apparent from Fig. 7b at larger internuclear distances ψ_2^{dia} becomes dominant (large $d_2(R)$). This means that as the atoms approach each other, a qualitative change of the interatomic interaction takes place. Thus, the adiabatic state changes its character. The predominantly ionic curve transforms to the predominantly covalent curve. This switching involves an electron jumping from Na to Cl. This is the so-called harpooning effect. The jumping of the electron takes place near 10 \AA . The time associated with the adjustment of the wave function to the non-BO perturbation is of the order of $h/2V_{12}$ which is the splitting of the adiabatic surfaces. Thus, the non-BO coupling is large in the vicinity of an avoided crossing.

This result not only applies to a diatomic molecule, but it is a general theorem of quantum mechanics. The result is true, when the Hamiltonian contains a parameter and its eigenvalues are functions of that parameter.

The number of degrees of freedom, which are the parameters that change (e.g., internuclear distances), is larger when dealing with a multi-atomic molecule. It can be shown that even for molecules with three centers, a degeneracy between states can occur (120). Even then the intersection only occurs at a point. This point is called a “conical intersection” because it has a characteristic shape of two surfaces that touch at a point. The special case of a conical intersection is also sometimes called a Jahn-Teller distortion (Jahn and Teller 1937). Jahn and Teller proved that for

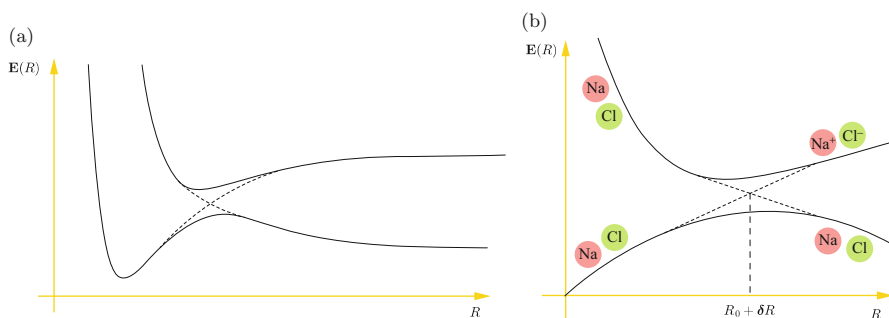


Fig. 7 Adiabatic states in the vicinity of an avoided crossing. (a) Avoided crossing. (b) The ground and the first excited state of NaCl molecule

degenerate electronic states, any symmetrical position of the nuclei, except collinear, is unstable. As a result of this instability, the nuclei move in such a way that the degeneracy of the electronic state is removed. In this way the ground state of a symmetrical, nonlinear molecule can become non-degenerate. An exception is linear molecules. For them, the linear position of the nuclei is stable even if the electronic state is degenerate. This was shown by Ruch and Schönhofer (1965). The reader can find extensive analysis of this phenomenon in Landau and Lifschitz (1981), Haken and Wolf (2004), and Piela (2007).

Non-adiabatic effects can be important in spectroscopy and photochemistry (see, e.g., Landau and Lifschitz 1981; Haken and Wolf 2004).

Examples of molecules for which the BO approximation fails are muonic molecules. In the muonic molecules, one or more electrons are replaced by muons. Muons are unstable subatomic particles. They are similar to electrons (muon has the charge of an electron), but they are about 207 times more massive. In determining bound states of the muonic molecules, it is erroneous to assume that the muon, like the electron, is much lighter than the nuclei and the nuclei can be treated as infinitely heavy. As a consequence, the BO approximation ceases to function.

Apart from the BO Approximation

In most cases, in particular in multi-atomic molecules, the BO approximation is the basic tool of quantum chemistry. However, in the case of small two- and three-center molecules containing several electrons, nonrelativistic (Kingham and Adamowicz 1997; Cafiero and Adamowicz 2002; Pavanello et al. 2012) and relativistic (Stanke et al. 2007a, 2006, 2009a, 2008) calculations, which do not use the BO approximation, have been executed. In this section I discuss this possibility and show that taking into account the coupling between electronic and nuclear motions exactly leads to very interesting conclusions.

Not assuming the Born-Oppenheimer approximation means that the theoretical model does not neglect the motion of the heavy particles – the nuclei – and thus all particles (nuclei and electrons) are treated on equal footing. Due to this equal treatment, also the description of the relativistic effects is uniformly accurate for all particles forming the system at the nonrelativistic level of quantum mechanics. Therefore, in this section, the notation is slightly different than in the previous part of the chapter. In particular, the numbering of the particles is different and their masses and charges of *all* particles are denoted by M_i and Q_i . The nonrelativistic (nr) description without the BO approximation starts with the following N -particle laboratory-frame molecular or atomic Hamiltonian:

$$H^{\text{nr}}(\mathbf{P}, \mathbf{R}) = \sum_{i=1}^N \frac{\mathbf{P}_i^2}{2M_i} + \sum_{i=1}^N \sum_{j>i}^N \frac{Q_i Q_j}{\|\mathbf{R}_i - \mathbf{R}_j\|}. \quad (121)$$

In the next step the center-of-mass motion (H_{cm}^{nr}) is separated out from the all-particle Hamiltonian (121), and internal Hamiltonian H_{int}^{nr} is obtained. This separation is possible due to a certain choice of the new coordinate system. The new coordinates are Cartesian, again a special choice, with the first three describing the position of the center of mass and the remaining $3N - 3 \equiv 3n$ being the internal coordinates. The transformation from the laboratory coordinates, \mathbf{R}_i , to the new coordinates is applied to the position vectors and the momentum vectors, \mathbf{P}_i , of all particles. Let us also denote the masses of the particles as M_i and the charges as Q_i . Let us denote the center-of-mass coordinates as \mathbf{r}_{cm} and the internal coordinates as \mathbf{r}_i , where $i = 1, \dots, n$, with $n = N - 1$. The \mathbf{r}_i vectors describe the positions of the particles $2, \dots, N$ with respect to particle 1, which is chosen to be the so-called reference particle. Let us denote the momenta corresponding to the \mathbf{r}_i coordinated as \mathbf{p}_i . Also, let us denote the reduce masses as μ_i . The coordinate transformation and the transformation of the momenta can be symbolically written as

$$\begin{aligned} O : \mathbf{R} &\longrightarrow \mathbf{r} & \mathbf{r} &= O\mathbf{R}, \\ O : \mathbf{P} &\longrightarrow \mathbf{p} & \mathbf{p} &= (O)^{-1}\mathbf{P}, \end{aligned}$$

where O denotes the following coordinate transformation:

$$O = \begin{bmatrix} \frac{M_1}{M_T} & \frac{M_2}{M_T} & \frac{M_3}{M_T} & \dots & \frac{M_N}{M_T} \\ -1 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \dots & 1 \end{bmatrix} \otimes I_3, \quad (122)$$

$$\mathbf{R} \longrightarrow [\mathbf{r}_{cm}^T, \mathbf{r}^T]^T,$$

where

$$M_T = \sum_{i=1}^N M_i.$$

The type of variables used to describe the internal state of the system and the position of the attachment of the system in space is arbitrary. The reference particle in our approach is usually the heaviest particle. The transformation of the coordinate system results in the total Hamiltonian, H_{tot}^{nr} , separating to a sum of two parts, the Hamiltonian describing the kinetic energy of the center-of-mass motion and the internal Hamiltonian:

$$H_{\text{tot}}^{\text{nr}}(\mathbf{r}_{cm}, \mathbf{p}_{cm}, \mathbf{r}, \mathbf{p}) = \left(-\frac{1}{2} \frac{1}{M_T} \nabla_{\mathbf{r}_{cm}}^T \cdot \nabla_{\mathbf{r}_{cm}} \right) + \left(-\frac{1}{2} \sum_i^n \frac{1}{\mu_i} \nabla_{\mathbf{r}_i}^T \cdot \nabla_{\mathbf{r}_i} - \frac{1}{2} \sum_{i \neq j}^n \frac{1}{M_1} \nabla_{\mathbf{r}_i}^T \cdot \nabla_{\mathbf{r}_i} + \sum_{i < j}^n \frac{q_i q_j}{r_{ij}} + \sum_{i=1}^n \frac{q_0 q_i}{r_i} \right), \quad (123)$$

where M_1 is the mass of the reference particle, and M_T is total mass of the system. $\nabla_{\mathbf{r}_i}$ denotes the gradient vector expressed in terms of the \mathbf{r}_i coordinates, and r_{ij} is the distance between the i -th and j -th particles, defined as $r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\| = \|\mathbf{R}_{i+1} - \mathbf{R}_{j+1}\|$. The relative distances with respect to the reference particle are $r_j = \|\mathbf{r}_j\| = \|\mathbf{R}_{j+1} - \mathbf{R}_1\|$. In Eq.(123) one can see the separation of the nonrelativistic Hamiltonian, $H_{\text{tot}}^{\text{nr}}$, into the kinetic energy operator of the motion of the center of mass, $H_{\text{cm}}^{\text{nr}}(\mathbf{r}_{cm}, \mathbf{p}_{cm})$, and the internal Hamiltonian, $H_{\text{int}}^{\text{nr}}(\mathbf{r}, \mathbf{p})$:

$$H_{\text{tot}}^{\text{nr}}(\mathbf{r}_{cm}, \mathbf{p}_{cm}, \mathbf{r}, \mathbf{p}) = H_{\text{cm}}^{\text{nr}}(\mathbf{r}_{cm}, \mathbf{p}_{cm}) + H_{\text{int}}^{\text{nr}}(\mathbf{r}, \mathbf{p}). \quad (124)$$

Both old and new coordinate systems provide a complete description of the positions of all particles in space. In the non-BO calculations, the internal Hamiltonian, $H_{\text{int}}^{\text{nr}}(\mathbf{r}, \mathbf{p})$, is used.

The internal Hamiltonian can be viewed as describing a system of $n - 1$ pseudoparticles with the masses equal to the reduced masses, μ_i , and the charges $q_{i-1} = Q_i$ ($i = 1, \dots, n + 1$). q_0 is the charge of the reference particle. The pseudoparticles interact with each other by the Coulomb potential and their motions are coupled through the mass-polarization terms. The internal Hamiltonian in the nonrelativistic approximation includes the operators representing the kinetic energy of the pseudoparticles:

$$-\frac{1}{2} \sum_i^n \frac{1}{\mu_i} \nabla_{\mathbf{r}_i}^T \cdot \nabla_{\mathbf{r}_i},$$

the mass-polarization term:

$$-\frac{1}{2} \sum_{i \neq j}^n \frac{1}{M_1} \nabla_{\mathbf{r}_i}^T \cdot \nabla_{\mathbf{r}_i}$$

and the Coulomb potential:

$$\sum_{i < j}^n \frac{q_i q_j}{r_{ij}} + \sum_{i=1}^{n-1} \frac{q_0 q_i}{r_i}.$$

What is accomplished by the above-described coordinate transformation for an isolated system in vacuum is also separation of the energy into a non-quantized

translational energy of the center-of-mass motion and the internal energy of the system, which for the bound states is quantized. In the new coordinates, the total wave function of the system can be written as a simple product of a wave function for the center of mass and the wave function representing an internal state of the system:

$$\Phi(\mathbf{r}_{cm}, \mathbf{r}_1, \dots, \mathbf{r}_n) = \Psi(\mathbf{r}_{cm}) \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n).$$

Recent progress in the development of high-precision laser-based spectroscopy techniques to study atomic and molecular spectra has been offering new capabilities to measure energy levels of small atoms and molecules in the gas phase with unprecedented accuracy. Due to these developments, the energy levels of atomic and molecular bound states and the corresponding transition energies are known with sub 0.001 cm^{-1} accuracy.

The experimental results can be directly compared with the theoretical calculations. This comparison not only tests the correctness of the quantum mechanical model used in the calculations, but also it sheds light on the nature of the chemical bonding, contributes to the measurements of the fundamental constants, and provides information on the chemical composition of distant interstellar objects. Very accurate experimental measurements challenge the theoreticians to improve the quantum mechanical models used in the calculations and to carry out the development of new algorithms for more accurate determination of bound states of atomic and molecular systems. The measurements also provide data for testing more approximate quantum mechanical methods developed for calculations of larger atoms and molecules.

For a full description of atoms and molecules without the BO approximation, an approach that allows to go beyond the Schrödinger description and to calculate relativistic and quantum electrodynamics corrections is needed. For systems comprising light nuclei and electrons, satisfactory accuracy can be achieved by representing the system energy in terms of powers of the fine-structure parameter $\alpha = 1/c$. In this way different effects are systematically accounted for. The terms proportional to α^2 are the leading relativistic effects (Bethe and Salpeter 1957) ($\alpha^2 \Delta E^{\text{rel}}$), and the terms proportional to α^3 ($\alpha^3 \Delta E_{\text{QED}}$) and α^4 ($\alpha^4 \Delta E_{\text{QED}}$) are the leading quantum electrodynamics (QED) effects. The basis for the calculation of relativistic effects is therefore the nonrelativistic Schrödinger Hamiltonian written without the BO approximation, which is treated as the nonperturbative operator.

The relativistic and QED effects can be determined as corrections with the use of the Rayleigh-Schrödinger perturbation theory¹⁰:

$$E_{\text{tot}} = E^{\text{nr}} + \alpha^2 \Delta E^{\text{rel}} \quad (125)$$

¹⁰From now on the symbol “tot” will be used to denote the sum of the total nonrelativistic energy and the corrections up to the certain order in terms of the hyperfine constant α .

If the relativistic calculations shall be calculated for states of the S symmetry for atoms and Σ symmetry for the diatomic molecules, the set of relativistic corrections must be included: the mass-velocity correction (mv), one- and two-particle Darwin corrections (d1,d2), and the Fermi spin-spin (ss) and orbit-orbit (oo) corrections.

The contribution of the spin-orbit correction and a part of the spin-spin correction to the energy for S states are equal to zero. The relativistic operators, before their use to calculate the relativistic energy corrections as expectation values of these operators, were transformed to the internal coordinate frame. I calculated these expectation values using the nonrelativistic wave functions obtained without assuming the Born-Oppenheimer (BO) approximation. To fulfill the symmetry requirements of the problem for the wave function and to effectively describe the correlation effects in our calculations, we have used the following explicitly correlated all-particle Gaussian (ECG) functions in calculations of atoms with s electrons

$$\psi_k(\mathbf{r}) = \exp[-\mathbf{r}^T \bar{\mathbf{A}}_k \mathbf{r}] \quad (126)$$

and the following Gaussians in calculations of diatomic molecules with σ electrons

$$\phi_k = r_1^{m_k} \exp[-\mathbf{r}^T (\mathbf{A}_k \otimes \mathbf{I}_3) \mathbf{r}] = r_1^{m_k} \exp[-\mathbf{r}^T \bar{\mathbf{A}}_k \mathbf{r}], \quad (127)$$

where

$$r_1 = |\mathbf{R}_1 - \mathbf{R}_2|.$$

is the distance between the two nuclei. Matrix $\bar{\mathbf{A}}$ occurring in the formulas (126) and (127) is a symmetric matrix of the nonlinear variational parameters.¹¹ $\bar{\mathbf{A}}$ is a positive-definite $3n \times 3n$ matrix obtained as the following product of the $n \times n$ matrix \mathbf{A} and a 3×3 unit matrix \mathbf{I}_3 :

$$\bar{\mathbf{A}} = \mathbf{A} \otimes \mathbf{I}_3.$$

As one can see, both functions (126) and (127) via their explicit dependence on the distances between the pseudoparticles can effectively describe the correlation effects associated with their coupled motion. In the case of the (126) basis set, only the correlated motion of pseudoelectrons with respect the center of the coordinate system, where the charge of the nucleus is located, needs to be described. The atomic basis functions (126) are spherically symmetric around the center of the positive charge of the nucleus as required by the symmetry of the problem. As the internal Hamiltonian for a molecular system is, alike in the atomic case, also isotropic with respect to

¹¹Optimal values of these parameters in the basis functions (126) and (127) are determined using the gradient method.

rotations about the center of the coordinate system, where the charge of the reference particle is located, the diatomic basis functions (127) are also spherically symmetric with respect to this center. However, due to the need to describe the nucleus-nucleus correlation effects in the calculation for a diatomic system the atomic ECG basis functions (126) have to be modified. As the correlation effect increases with the increase of the masses of the interacting particles, the simple ECGs, which only depend on the nucleus-nucleus distance in the exponent, are not sufficient for the molecular case. For diatomics, an additional preexponential factor needs to be included in the Gaussians in the form of a nonnegative power of the internuclear distance, $r_1^{m_k}$ (127). This factor shifts the maximum of the Gaussian away from the reference particle. The higher the power, the larger the shift. m_k has a value in the 0–250 range in our calculations. The presence of $r_1^{m_k}$ in the Gaussians also helps describe radial nodes in wave functions of higher vibrational states. Gaussian basis functions are not as efficient as the Hylleraas- or Slater-type basis functions in describing bound states of multiparticle systems interacting with Coulomb forces. The Gaussians do not have the correct behavior at very short and at very long inter-particle distances. However, in calculations of both atomic and molecular with more than three electrons, the Gaussians are the only functions which can be used, because the use of other types of basis functions leads to Hamiltonian integrals too difficult to calculate. Also, as our atomic and molecular calculations have demonstrated, properly optimized Gaussian basis sets allow to achieve in the calculations an accuracy level similar to the accuracy of the most exact experiments performed for small atomic and diatomic systems (Stanke et al. 2007a, b, 2009b).

Charge Asymmetry in Pure Vibrational States of the HD Molecule

The consequences of taking into account the coupling between electronic and nuclear motion is illustrated by the example of the molecule HD (Bubin et al. 2009). In the approach without the Born-Oppenheimer approximation, all particles are treated on equal footing. The coupled electron-nucleus motion is a part of the description of the system. The calculation of the expectation values of the relative inter-particle distance was performed in the basis set of explicitly correlated Gaussians. These functions explicitly depend on the inter-particle distance which for HD are the electron-electron, deuteron-proton, deuteron-electron, and proton-electron distances. Such dependency allows for describing very accurately the correlation effects related to the coupled motion of the nuclei and the electrons.

In consequence, the increase of the vibrational excitation (particularly for large v) results in mixing into the total wave function of some small contributions of electronically excited states. In high vibrationally excited states, the corresponding wave functions become to some (small) degree superpositions of the product of the electronic ground-state Σ_g wave function times the corresponding vibrational wave function and a higher electronic state (states) times vibrational wave functions with

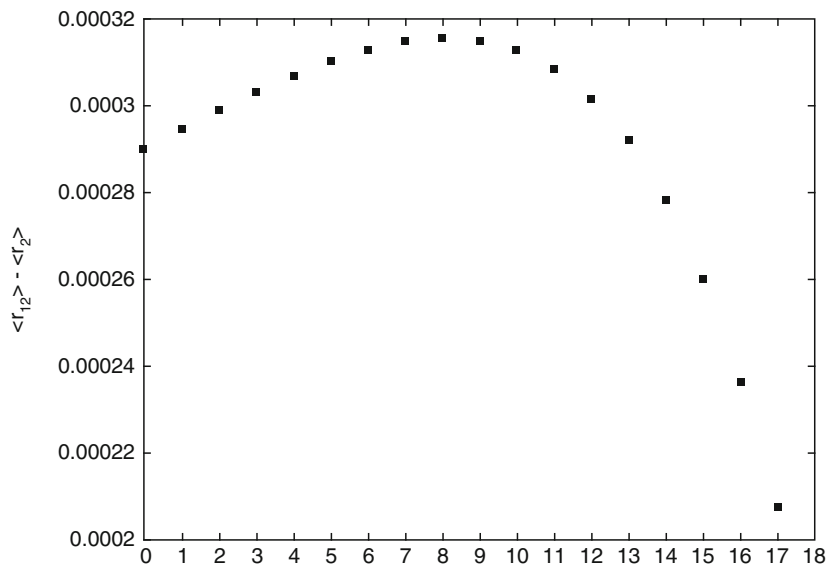


Fig. 8 The differences between the average inter-particle electron-proton and electron-deuteron distances ($\langle r_{12} \rangle - \langle r_2 \rangle$) as a function of the number of the vibrational state of the HD molecule. The calculations were done with 6,000 ECGs per state

smaller v 's. Because of this state mixing, it is difficult to ambiguously assign a single vibrational quantum number to the state the molecule is in. In consequence, the vibrational quantum number, which usually corresponds to the number of nodes in the wave function describing the nuclear motion, in the description of the molecule provided by the non-BO approach is not, strictly speaking, a good quantum number. In general, it is not possible to strictly separate the vibrational and electronic contributions to the total wave function. Thus assigning vibrational quantum numbers to consecutive states is only approximate, and it is merely a way of numbering the states in the rising-energy order. It would be more correct to use a different description of the states we calculate. Perhaps the term *rotationless* reflecting the zero value of the angular momentum for those states would be a more correct designation.

The superposition of different electronic states in high excited vibrational states introduces an interesting effect of electronic asymmetry. Using the expectation values of the deuteron-electron and proton-electron inter-particle distances, we determined these values as a function of the vibrational-state number. The results provided an evidence of the asymmetry, which is the highest for the $v = 8$ and the lowest for the $v = 17$ state. The electronic asymmetry causes that the dipole moment of the HD molecules changes by 50% in the transition from the $v = 17$ state to the $v = 8$ state. This behavior is depicted in Fig. 8.

Molecular Relativistic Corrections Determined in the Framework Where the Born-Oppenheimer Approximation Is Not Assumed

Equally interesting is the dependence of the relativistic corrections on the vibrational quantum number. Such a study has been performed for the heteronuclear isotopologues of the hydrogen molecular ion, the HD^+ ion (Stanke and Adamowicz 2013). There are two dissociation channels for HD^+ : deuter + proton ($\text{D} + \text{H}^+$) and deuteron and hydrogen ($\text{D}^+ + \text{H}$). The former is energetically lower than the latter. The point which has been examined is the convergence pattern of the relativistic corrections with the increasing vibrational quantum number. It has been demonstrated that the values of the mass-velocity (mv), Darwin (d1 + d2), and orbit-orbit (oo) corrections in the highest vibrational states approach the corresponding values for deuteron and not hydrogen. Though such a behavior of the relativistic corrections is expected, it can only be described if the corrections are calculated in the framework where the BO approximation is not assumed. In this framework the relativistic corrections are calculated for all particles comprising the system (i.e., deuteron, proton, and electron). This is different from the BO approach, where the only moving particle for which the relativistic corrections are calculated is the electron. The non-BO calculation of the relativistic corrections is performed using ECG basis functions which explicitly depend on all inter-particle distances, i.e., the deuteron-proton, deuteron-electron, and proton-electron.

In Table 1 results of Stanke et al. (Stanke and Adamowicz 2013) obtained without assuming the BO approximation as well as of Wolniewicz (1995), Wolniewicz and Orlikowski (1991), and of Howells and Kennedy (1990) obtained in the BO framework have been collected.¹² Despite the absence of the results for the $v = 22$ state, an analysis of the results reveals a definite trend, namely, the results of Wolniewicz, contrary to our non-BO results, converge to the theoretical value calculated for the system comprising of a hydrogen atom with an infinitely heavy nucleus and a proton (${}^\infty\text{H} + \text{H}^+$), i.e., to the asymptotic value one obtains in the BO approximation for the dissociation channel, $\text{D} + \text{H}^+$. In the same time, the values obtained without assuming the BO approximation converge to correct limit corresponding to the deuter+proton dissociation channel. It is worth stressing that the values of the total relativistic correction calculated by us and those calculated with assuming the BO approximation are close for all pure vibrational states of HD^+ .

It is also interesting that the values of the total relativistic correction for the different dissociation channels of HD^+ are the same. The equality of the α^2 corrections for channels $\text{H} + \text{D}^+$ and $\text{D} + \text{H}^+$ arises from the fact that, as proven by Pachucki (Pachucki and Grotch 1995), the component of the recoil correction proportional to $1/M$ for a one-electron atom disappears (although its components do not disappear).

¹²Unfortunately in the results of Wolniewicz, the values for the $v = 22$ are missing.

Table 1 A comparison of our non-BO results for the individual contributions to the relativistic correction, i.e., mass-velocity (mv), one- and two-body Darwin corrections (d1 + d2), and orbit-orbit (oo) correction, and the total relativistic correction (rel cor) for the pure vibrational states of the HD⁺ ion with the results of Wolniewicz and Orlikowski 1991 and of Howells and Kennedy (1990) obtained in the BO framework. The non-BO calculations for the $v = 0$ and $v = 1$ states were performed with 1,000 ECGs, for the $v = 2, \dots, 21$ states with 2,000 ECGs, and for the $v = 22$ state with 4,000 ECGs. All quantities in the table are in cm^{-1}

v	mv	non-BO d1+d2	oo	rel cor	mv	Wolniewicz and Orlikowski (1991) d1+d2	rel cor	Howells and Kennedy (1990) rel cor
0	-9.1988	7.6015	-0.56	-1.6029	-9.2190	7.6169	-1.6022	-1.6015
1	-8.9886	7.4244	-0.55	-1.5697	-9.0123	7.4422	-1.5702	-1.5696
2	-8.8012	7.2658	-0.53	-1.5407	-8.8199	7.2789	-1.5411	-1.5405
3	-8.6224	7.1131	-0.52	-1.5145	-8.6410	7.1264	-1.5147	-1.5141
...
18	-7.2556	5.8513	-0.0046	-1.4088	-7.2895	5.8804	-1.4092	-1.4091
19	-7.2484	5.8309	-0.0046	-1.4221	-7.2816	5.8595	-1.4222	-1.4221
20	-7.2589	5.8259	-0.0045	-1.4375	-7.2844	5.8467	-1.4378	-1.4377
21	-7.2773	5.8248	-0.0035	-1.4559	-7.2961	5.8424	-1.4537	-1.4564
22	-7.2947	5.8375	-0.0032	-1.4604				-1.4605
D + H ⁺	-7.2966	5.8389	-0.0032	-1.4609				
D ⁺ + H	-7.2887	5.8341	-0.0064	-1.4609				
∞ H + H ⁺	-7.3046	5.8437	0	-1.4609				

Acknowledgements This work has been supported by the Polish National Science Centre, grant DEC-2013/10/E/ST4/00033. I am grateful to Professor Krzysztof Strasburger, the reviewer, for his valuable comments and to Professors Jacek Karwowski and Ludwik Adamowicz for useful discussions. Thanks are extended to Ms. Ewa Palikot, MSc. for drawing the figures.

Appendix A

Detailed derivation of the Laplace operator in the new coordinate (7), (8), (9), and (10)

- First derivatives of the nuclear coordinates:
 - nucleus a

$$\begin{aligned} \frac{\partial}{\partial X'_a} &= \frac{\partial X_{CM}}{\partial X'_a} \frac{\partial}{\partial X_{CM}} + \frac{\partial Y_{CM}}{\partial X'_a} \frac{\partial}{\partial Y_{CM}} + \frac{\partial Z_{CM}}{\partial X'_a} \frac{\partial}{\partial Z_{CM}} + \frac{\partial R_x}{\partial X'_a} \frac{\partial}{\partial R_x} + \frac{\partial R_y}{\partial X'_a} \frac{\partial}{\partial R_y} + \frac{\partial R_z}{\partial X'_a} \frac{\partial}{\partial R_z} + \\ &+ \sum_{i=1}^{N_e} \left(\frac{\partial x_i}{\partial X'_a} \frac{\partial}{\partial x_i} + \frac{\partial y_i}{\partial X'_a} \frac{\partial}{\partial y_i} + \frac{\partial z_i}{\partial X'_a} \frac{\partial}{\partial z_i} \right) = \frac{M_a}{M} \frac{\partial}{\partial X_{CM}} + \frac{\partial}{\partial R_x} - \frac{1}{2} \sum_{i=1}^{N_e} \frac{\partial}{\partial x_i} \end{aligned} \quad (128)$$

analogously:

$$\begin{aligned} \frac{\partial}{\partial Y'_a} &= \frac{M_a}{M} \frac{\partial}{\partial Y_{CM}} + \frac{\partial}{\partial R_y} - \frac{1}{2} \sum_{i=1}^{N_e} \frac{\partial}{\partial y_i}, \\ \frac{\partial}{\partial Z'_a} &= \frac{M_a}{M} \frac{\partial}{\partial Z_{CM}} + \frac{\partial}{\partial R_z} - \frac{1}{2} \sum_{i=1}^{N_e} \frac{\partial}{\partial z_i} \end{aligned} \quad (129)$$

- nucleus b - analogous:

$$\frac{\partial}{\partial X'_b} = \frac{M_a}{M} \frac{\partial}{\partial X_{CM}} - \frac{\partial}{\partial R_x} - \frac{1}{2} \sum_{i=1}^{N_e} \frac{\partial}{\partial x_i} \quad (130)$$

and similarly for the nuclear coordinates \tilde{y}_b and \tilde{z}_b

- First derivatives of the electron coordinates:

$$\begin{aligned} \frac{\partial}{\partial x'_i} &= \frac{X_{CM}}{\partial x'_i} \frac{\partial}{\partial X_{CM}} + \frac{Y_{CM}}{\partial x'_i} \frac{\partial}{\partial Y_{CM}} + \frac{Z_{CM}}{\partial x'_i} \frac{\partial}{\partial Z_{CM}} + \frac{R_x}{\partial x'_i} \frac{\partial}{\partial R_x} + \frac{R_y}{\partial x'_i} \frac{\partial}{\partial R_y} + \frac{R_z}{\partial x'_i} \frac{\partial}{\partial R_z} + \\ &+ \sum_{j=1}^{N_e} \left(\frac{\partial x_j}{\partial x'_i} \frac{\partial}{\partial x_j} + \frac{\partial y_j}{\partial x'_i} \frac{\partial}{\partial y_j} + \frac{\partial z_j}{\partial x'_i} \frac{\partial}{\partial z_j} \right) = \frac{m}{M} \frac{\partial}{\partial X_{CM}} + \sum_{j=1}^{N_e} \left(\delta_{ij} \frac{\partial}{\partial x_j} \right) = \\ &= \frac{m}{M} \frac{\partial}{\partial X_{CM}} + \frac{\partial}{\partial x_i} \end{aligned} \quad (131)$$

and similarly for the electron coordinates y'_i and z'_i

- Second derivatives of the nuclear coordinates (we use the formula $[a+(b+c)]^2 = a^2 + b^2 + c^2 + 2ab + 2bc + 2ac$):
 - nucleus a

$$\begin{aligned}
 \frac{\partial^2}{\partial X'_a{}^2} &= \left(\frac{M_a}{M} \frac{\partial}{\partial X_{\text{CM}}} + \frac{\partial}{\partial R_x} - \frac{1}{2} \sum_{i=1}^{N_e} \frac{\partial}{\partial x_i} \right) \left(\frac{M_a}{M} \frac{\partial}{\partial X_{\text{CM}}} + \frac{\partial}{\partial R_x} - \frac{1}{2} \sum_{i=1}^{N_e} \frac{\partial}{\partial x_i} \right) = \\
 &= \left(\frac{M_a}{M} \right)^2 \frac{\partial^2}{\partial X_{\text{CM}}^2} + \frac{\partial^2}{\partial R_x^2} + \frac{1}{4} \sum_{i=1}^{N_e} \left(\frac{\partial}{\partial x_i} \right) + \\
 &\quad + 2 \frac{M_a}{M} \frac{\partial^2}{\partial X_{\text{CM}} \partial R_x} - \frac{\partial}{\partial X_{\text{CM}}} \sum_{i=1}^{N_e} \frac{\partial}{\partial x_i} - \frac{M_a}{M} \frac{\partial}{\partial R_x} \sum_{i=1}^{N_e} \frac{\partial}{\partial x_i} \quad (132)
 \end{aligned}$$

- nucleus b

$$\begin{aligned}
 \frac{\partial^2}{\partial X'_b{}^2} &= \left(\frac{M_b}{M} \frac{\partial}{\partial X_{\text{CM}}} - \frac{\partial}{\partial R_x} - \frac{1}{2} \sum_{i=1}^{N_e} \frac{\partial}{\partial x_i} \right) \left(\frac{M_b}{M} \frac{\partial}{\partial X_{\text{CM}}} - \frac{\partial}{\partial R_x} - \frac{1}{2} \sum_{i=1}^{N_e} \frac{\partial}{\partial x_i} \right) = \\
 &= \left(\frac{M_b}{M} \right)^2 \frac{\partial^2}{\partial X_{\text{CM}}^2} + \frac{\partial^2}{\partial R_x^2} + \frac{1}{4} \sum_{i=1}^{N_e} \left(\frac{\partial}{\partial x_i} \right) + \\
 &\quad - 2 \frac{M_b}{M} \frac{\partial^2}{\partial X_{\text{CM}} \partial R_x} + \frac{\partial}{\partial X_{\text{CM}}} \sum_{i=1}^{N_e} \frac{\partial}{\partial x_i} - \frac{M_b}{M} \frac{\partial}{\partial R_x} \sum_{i=1}^{N_e} \frac{\partial}{\partial x_i} \quad (133)
 \end{aligned}$$

- Second derivatives of the electron coordinates:

$$\begin{aligned}
 \frac{\partial^2}{\partial x_i'^2} &= \left(\frac{m}{M} \frac{\partial}{\partial X_{\text{CM}}} + \frac{\partial}{\partial x_i} \right) \left(\frac{m}{M} \frac{\partial}{\partial X_{\text{CM}}} + \frac{\partial}{\partial x_i} \right) \\
 &= \left(\frac{m}{M} \right)^2 \frac{\partial^2}{\partial X_{\text{CM}}^2} + \frac{\partial^2}{\partial x_i^2} + 2 \frac{m}{M} \frac{\partial^2}{\partial X_{\text{CM}} \partial x_i} \quad (134)
 \end{aligned}$$

Similarly, the remaining coordinates.

Appendix B

Flat rotor – a system of two particles with masses m_1 and m_2 at a constant distance R from each other. In this system there are no external forces, but it can rotate around its center of mass. Hamiltonian for this system has a form:

$$\hat{H} = -\frac{1}{2m_1} \nabla_{\mathbf{r}_1}^2 - \frac{1}{2m_2} \nabla_{\mathbf{r}_2}^2, \quad (135)$$

where \mathbf{r}_1 and \mathbf{r}_2 are vectors that describe the positions of the two bodies in the laboratory reference frame. Alternatively one can describe the same system, in the center-of-mass reference frame. The coordinates of the center of mass in the laboratory frame are given by

$$\mathbf{R} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2}{m_1 + m_2} = [X, Y, Z] \quad (136)$$

If one wants to express the Hamiltonian in the center-of-mass coordinates, one has to express the second derivatives in the new variables. Proceeding as in Appendix A one gets

$$\begin{aligned} \frac{\partial}{\partial x_1} &= \frac{\partial X}{\partial x_1} \frac{\partial}{\partial X} + \frac{\partial Y}{\partial x_1} \frac{\partial}{\partial Y} + \frac{\partial Z}{\partial x_1} \frac{\partial}{\partial Z} + \frac{\partial x}{\partial x_1} \frac{\partial}{\partial x} + \frac{\partial y}{\partial x_1} \frac{\partial}{\partial y} + \frac{\partial z}{\partial x_1} \frac{\partial}{\partial z} \\ &= \frac{\partial X}{\partial x_1} \frac{\partial}{\partial X} + \frac{\partial x}{\partial x_1} \frac{\partial}{\partial x} \\ &= \frac{m_1}{m_1 + m_2} \frac{\partial}{\partial X} + \frac{\partial}{\partial x}, \end{aligned} \quad (137)$$

$$\begin{aligned} \frac{\partial}{\partial x_2} &= \frac{\partial X}{\partial x_2} \frac{\partial}{\partial X} + \frac{\partial Y}{\partial x_2} \frac{\partial}{\partial Y} + \frac{\partial Z}{\partial x_2} \frac{\partial}{\partial Z} + \frac{\partial x}{\partial x_2} \frac{\partial}{\partial x} + \frac{\partial y}{\partial x_2} \frac{\partial}{\partial y} + \frac{\partial z}{\partial x_2} \frac{\partial}{\partial z} \\ &= \frac{\partial X}{\partial x_2} \frac{\partial}{\partial X} + \frac{\partial x}{\partial x_2} \frac{\partial}{\partial x} \\ &= \frac{m_2}{m_1 + m_2} \frac{\partial}{\partial X} - \frac{\partial}{\partial x}. \end{aligned} \quad (138)$$

Similarly, for the remaining coordinates. The second derivatives are given by

$$\begin{aligned} \frac{\partial^2}{\partial x_1^2} &= \left(\frac{m_1}{m_1 + m_2} \frac{\partial}{\partial X} + \frac{\partial}{\partial x} \right) \cdot \left(\frac{m_1}{m_1 + m_2} \frac{\partial}{\partial X} + \frac{\partial}{\partial x} \right) \\ &= \left(\frac{m_1}{m_1 + m_2} \right)^2 \frac{\partial^2}{\partial X^2} + \frac{2m_1}{m_1 + m_2} \frac{\partial^2}{\partial X \partial x} + \frac{\partial^2}{\partial x^2} \end{aligned} \quad (139)$$

$$\begin{aligned} \frac{\partial^2}{\partial x_2^2} &= \left(\frac{m_2}{m_1 + m_2} \frac{\partial}{\partial X} - \frac{\partial}{\partial x} \right) \cdot \left(\frac{m_2}{m_1 + m_2} \frac{\partial}{\partial X} - \frac{\partial}{\partial x} \right) \\ &= \left(\frac{m_2}{m_1 + m_2} \right)^2 \frac{\partial^2}{\partial X^2} - \frac{2m_2}{m_1 + m_2} \frac{\partial^2}{\partial X \partial x} + \frac{\partial^2}{\partial x^2} \end{aligned} \quad (140)$$

Thus, after substituting the second derivatives, one obtains the Hamiltonian operator (135) in the new variables

$$\begin{aligned}
\hat{H} &= -\frac{1}{2} \left[\frac{m_1}{(m_1 + m_2)^2} + \frac{m_2}{(m_1 + m_2)^2} \right] \left(\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2} \right) \\
&\quad - \frac{1}{2} \left[\frac{1}{m_1} + \frac{1}{m_2} \right] \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \\
&= -\frac{1}{2M} \Delta_{\mathbf{R}} - \frac{1}{2\mu} \Delta_{\mathbf{r}} \equiv \hat{H}_{\mathbf{R}} + \hat{H}_{\mathbf{r}},
\end{aligned} \tag{141}$$

where

$$\begin{aligned}
M &= m_1 + m_2, & \frac{1}{\mu} &= \frac{1}{m_1} + \frac{1}{m_2}, & \Delta_{\mathbf{R}} &= \frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}, \\
\Delta_{\mathbf{r}} &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},
\end{aligned} \tag{142}$$

and μ is the reduced mass. For the two particles moving freely in the space, that is assuming that $V = 0$, the eigenvalue equation of (141) can be written as $\hat{H}_{\mathbf{R}}$ and $\hat{H}_{\mathbf{r}}$:

$$(\hat{H}_{\mathbf{R}} + \hat{H}_{\mathbf{r}})\psi(\mathbf{R}, \mathbf{r}) = E\psi(\mathbf{R}, \mathbf{r}) \tag{143}$$

for

$$\psi(\mathbf{R}, \mathbf{r}) = \psi_{\mathbf{R}}(\mathbf{R}) \psi_{\mathbf{r}}(\mathbf{r}),$$

where

$$\hat{H}_{\mathbf{R}} \psi_{\mathbf{R}}(\mathbf{R}) = E_{\mathbf{R}} \psi_{\mathbf{R}}(\mathbf{R}) \quad \text{and} \quad \hat{H}_{\mathbf{r}} \psi_{\mathbf{r}}(\mathbf{r}) = E_{\mathbf{r}} \psi_{\mathbf{r}}(\mathbf{r}).$$

We focus only on the relative motion Hamiltonian $\hat{H}_{\mathbf{r}}$. The eigenfunction $\psi_{\mathbf{r}}(\mathbf{r})$ of this Hamiltonian depends on three variables. However, due to the fact that r is constant ($r^2 = x^2 + y^2 + z^2 = \text{const}$), only two angular variables are independent. In the spherical coordinates, with variables r, θ, ϕ , where $r = \text{const}$:

$$\begin{aligned}
x &= x(r, \theta, \phi) = r \sin \theta \cos \phi \\
y &= y(r, \theta, \phi) = r \sin \theta \sin \phi \\
z &= z(r, \theta, \phi) = r \cos \theta.
\end{aligned}$$

the Laplacian can be written as

$$\Delta_{\mathbf{r}} = \frac{1}{r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right].$$

Consequently, the rigid rotor Hamiltonian becomes

$$\hat{H} = -\frac{1}{2I} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right],$$

where

$$I = \mu R^2$$

is the moment of inertia. The eigenvalue problem of the rigid rotor is, thus, reduced to the equation

$$-\frac{1}{2I} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] Y(\theta, \phi) = E Y(\theta, \phi),$$

analogous to the problem of the angular momentum

$$-1 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] Y(\theta, \phi) = l(l+1)Y(\theta, \phi).$$

Thus, we can immediately conclude that the rigid rotor eigenfunctions are equal to the spherical harmonics

$$Y_J^M(\theta, \phi) = \frac{1}{\sqrt{2\pi}} N_{J,|M|} P_J^{|M|}(\cos \theta) e^{iM\phi},$$

where

$$P_l^m(x) = (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x)$$

is the associate Legendre polynomial. The eigenvalues are given by

$$E_J = \frac{1}{2I} J(J+1),$$

where $1/2I$ is the rotational constant.

Bibliography

- Bethe, H. A., & Salpeter, E. E. (1957). *Quantum mechanics of one- and two-electron systems*. Berlin: Springer.
- Born, M., & Oppenheimer, J. R. (1927). Zur Quantentheorie der Molekeln (On the quantum theory of molecules). *Annalen der Physik*, 84, 457–484.
- Bubin, S., Leonarski, F., Stanke, M., & Adamowicz, L. (2009). Charge asymmetry in pure vibrational states of the HD molecule. *The Journal of Chemical Physics*, 130, 124120.

- Cafiero, M., & Adamowicz, L. (2002). Nonadiabatic calculations of the dipole moments of LiH and LiD. *Physical Review Letters*, 88, 33002.
- Dalgarno, A., & McCarroll, R. (1956). Adiabatic coupling between electronic and nuclear motion in molecules. *Proceedings of the Royal Society (London)*, A237, 383.
- Davydov, A. S. (1965). *Quantum mechanics*. Oxford: Pergamon Press.
- Davydov, A. S. (1976). *Quantum mechanics* (2nd ed.). Oxford: Pergamon Pr.
- Demtröder, W. (2010). *Atoms, molecules and photons: An introduction to atomic-, molecular- and quantum physics* (2nd ed., 2010 ed.). Berlin: Springer, Berlin (January 19, 2011).
- Frolov, A. M. (1999). Bound-state calculations of Coulomb three-body systems. *Physical Review A*, 59, 4270.
- Haken, H., & Wolf, H. C. (2010). *Molecular physics and elements of quantum chemistry* (2nd Edn., 2004 edition). Berlin: Springer.
- Handy, N. C., & Lee, A. M. (1986). The adiabatic approximation. *Chemical Physics Letters*, 252, 425–430.
- Herzberg, G. (1951). *Spectra of diatomic molecules* (2nd ed.). D. Van Nostrand, New York.
- Howells, M. H., & Kennedy, R. A. (1990). Relativistic corrections for the ground and first excited states of H_2^+ , HD^+ and D_2^+ . *Journal of the Chemical Society Faraday Transactions*, 86, 3495.
- Hulburt, H. M., & Hirschfelder, J. O. (1941). Potential energy functions for diatomic molecules. *The Journal of Chemical Physics*, 9, 61.
- Jahn, H. A., & Teller, E. (1937). Stability of polyatomic molecules in degenerate electronic states. I. Orbital degeneracy. *Proceedings of the Royal Society of London Series A*, 161, 220.
- Kinghorn, D. B., & Adamowicz, L. (1997). *The Journal of Chemical Physics*, 106, 4589.
- Kotos, W., & Wolniewicz, L. (1963). Nonadiabatic theory for diatomic molecules and its application to the hydrogen molecule. *Reviews of Modern Physics*, 35, 473.
- Kołos, W. (1970). Adiabatic approximation and its accuracy. *Advanced in Quantum Chemistry*, 5, 99–133.
- Kołos, W., & Sadlej, J. (1998). *Atom i cząsteczka* (in Polish). Warszawa: WNT
- Kręglewski, M. (1979). *Zadania z chemii kwantowej*, Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.
- Landau, L. D., & Lifschitz, E. M. (1981). *Quantum mechanics – Non relativistic theory* (Course of theoretical physics, Vol. 3, 3rd ed.). Oxford: Pergamon Press. Butterworth-Heinemann.
- Pachucki, K., & Grotch, H. (1995). Pure recoil corrections to hydrogen energy levels. *Physical Review A*, 51, 1854.
- Pavanello, M., Adamowicz, L., Alijah, A., Zobov, N. F., Mizus, I. I., Polyansky, O. L., Tennyson, J., Szidarovszky, T., Császár, A. G., Berg, M., Petrigani, A., & Wolf, A. (2012). Precision measurements and computations of transition energies in rotationally cold triatomic hydrogen ions up to the midvisible spectral range. *Physical Review Letters*, 108, 023002.
- Piela, L. (2007). *Ideas of quantum chemistry* (1st ed.). Amsterdam: Elsevier Science., Amsterdam.
- Ruch, E., & Schönhofer, A. (1965). Ein Beweis des Jahn-Teller-Theorems mit Hilfe eines Satzes über die Induktion von Darstellungen endlicher Gruppen. *Theoretica Chimica Acta*, 3, 291–304.
- Stanke, M., & Adamowicz, L. (2013). Molecular relativistic corrections determined in the framework where the Born-Oppenheimer approximation is not assumed. *The Journal of Physical Chemistry A*, 117 (39), 10129–10137.
- Stanke, M., Kędziera, D., Molski, M., Bubin, S., Barysz, M., & Adamowicz, L. (2006). Convergence of experiment and theory on the pure vibrational spectrum of HeH^+ . *Physical Review Letters*, 96, 233002.
- Stanke, M., Kędziera, D., Bubin, S., & Adamowicz, L. (2007a). Lowest excitation energy of 9Be . *Physical Review Letters*, 99, 043001.
- Stanke, M., Kędziera, D., Bubin, S., & Adamowicz, L. (2007b). Ionization potential of 9Be calculated including nuclear motion and relativistic corrections. *Physical Review A*, 75, 052510.
- Stanke, M., Kędziera, D., Bubin, S., & Adamowicz, L. (2008). Complete α^2 relativistic corrections to the pure vibrational non-Born-Oppenheimer energies of HeH^+ . *Physical Review A*, 77, 022506.

- Stanke, M., Bubin, S., & Adamowicz, L. (2009a). Fundamental vibrational transitions of the ${}^3\text{He}^4\text{He}^+$ and ${}^7\text{LiH}^+$ ions calculated without assuming the Born-Oppenheimer approximation and with leading relativistic corrections. *Physical Review A*, *79*, 060501(R).
- Stanke, M., Komasa, J., Bubin, S., & Adamowicz, L. (2009b). Five lowest ${}^1\text{S}$ states of the Be atom calculated with a finite-nuclear-mass approach and with relativistic and QED corrections. *Physical Review A*, *80*, 022514.
- Sutcliffe, B. T. (2003). Approximate separation of electronic and nuclear motion, Part 6. In S. Wilson, P. F. Bernath, & R. McWeeny (Eds.), *Handbook of molecular physics and quantum chemistry* (Vol. 1, p. 475). Chichester: Wiley.
- Wolniewicz, L. (2011). Private consultations.
- Wolniewicz, L. (1995). Nonadiabatic energies of the ground state of the hydrogen molecule. *The Journal of Chemical Physics*, *103*, 1792.
- Wolniewicz, L., & Orlikowski, T. (1991). The $1s\sigma_g$ and $2p\sigma_u$ states of the H_2^+ , D_2^+ and HD^+ ions. *Molecular Physics*, *74*, 103–111.

Directions for Use of Density Functional Theory: A Short Instruction Manual for Chemists

6

Heiko Jacobsen and Luigi Cavallo

Contents

Introduction	226
DFT: A Paradigm Shift in Theoretical Chemistry	228
Holes and Electron Pairs	232
Climbing Jacob's Ladder	235
Practical DFT and the Density Functional Zoo	238
DFT: Computational Chemistry in Action	240
Computational Performance	241
Properties of Molecular and Electronic Structure	243
Orbitals in DFT	258
DFTips	260
B3LYP Is No Synonym for DFT	261
Choose Your DFT Method Carefully	261
Read the Fine Print	261
DFT Does Not Hold the Universal Answer to All Chemical Problems	262
Make DFT an Integral Part of Your Work	262
Follow Your Interests	262
Bibliography	263
Books on DFT	263
Reviews and Overviews of DFT	263
Conceptual Developments and Applications of DFT	264
Practical Developments and Applications of DFT	265

H. Jacobsen (✉)
KemKom, New Orleans, LA, USA
e-mail: jacobsen@kemkom.com

L. Cavallo
Physical Sciences and Engineering Division, Kaust Catalysis Center, King Abdullah University
of Science and Technology (Kaust), Thuwal, Saudi Arabia
e-mail: Luigi.Cavallo@kaust.edu.sa

Abstract

Two aspects are quintessential if one seeks to successfully perform DFT calculations: a basic understanding of how the concepts and models underlying the various manifestations of DFT are built and an essential knowledge of what can be expected from DFT calculations and how to achieve the most appropriate results. This chapter expands on the development and philosophy of DFT and aims to illustrate the essentials of DFT in a manner that is intuitively accessible. An analysis of the performance and applicability of DFT focuses on a representative selection of chemical properties, including bond lengths, bond angles, vibrational frequencies, electron affinities and ionization potentials, atomization energies, heats of formation, energy barriers, bond energies, hydrogen bonding, weak interactions, spin states, and excited states.

Introduction

Density functional theory (DFT) is an enticing subject. It appeals to chemists and physicists alike, and it is entrancing for those who like to work on mathematical physical aspects of problems, for those who relish computing observable properties from theory, and for those who most enjoy developing correct qualitative descriptions of phenomena. It is this combination of a qualitative model that at the same time furnishes quantitative reliable estimates that makes DFT particularly attractive for chemists.

DFT is an alternative, and complementary, to wave function theory (WFT). Both approaches are variations of the basic theme of electronic structure theory, and both methods originated during the late years of the 1920s. Whereas WFT evolved rapidly and gained general popularity, DFT found itself in a state of shadowy existence. It was the appearance of the key papers by Hohenberg and Kohn (1964) and by Kohn and Sham (1965), generally perceived as the beginning of modern DFT, which changed the perception and level of acceptance of DFT. With the evolution of reliable computational technologies for DFT chemistry, and with the advent of the generalized gradient approximation (GGA) during the 1980s, DFT emerged as powerful tool in computational chemistry, and without exaggeration the 1990s can be called the decade of DFT in electronic structure theory. During this time period, despite the lack of a complete development, DFT was already competitive with the best WFT methods. Furthermore, the advancement of computational hardware as well as software has progressed to a state where DFT calculations of “real molecules” can be performed with high efficiency and without major technical hurdles.

But at the end of the first decade of the new millennium, it appeared that DFT might have become a victim of its own success. DFT has transformed into an off-the-shelf technology and ready-to-crunch component and often was and still is used as such. Still and all, it became clear that the happy days of black-box DFT were over and that not all the promises of DFT came to fruition. DFT has its

own limitations and shortcomings, and the numbers obtained from DFT calculations began to lose some of their awe-inspiring admiration they enjoyed at the end of the first millennium. At the same time, DFT has matured into a standard research tool, used routinely by many experimental chemists to support their work.

Despite the fact that DFT has evolved into a first-choice approach for a stunning potpourri of applications throughout chemistry and materials science, it is all too well known that DFT in its current form still has serious limitations; Cohen et al. (2012) identify and analyze current challenges to DFT. Moreover, it appears that the philosophical, theoretical, and computational framework of DFT might lose its prominent position in electronic structure theory. Fifty years after its formal inception, Becke (2014) reviewed the development and current status of Kohn–Sham density functional theory and raises the concern that increasing pressure to deliver higher and higher accuracy and to adapt to ever more challenging problems “may submerge the theory in the wave-function sea” (Becke 2014). But in whatever direction the path of extended development will guide users and developers of DFT, it seems evident that DFT will continue its role as research tool not only for computational but also for theoretical inclined chemists.

We have composed our short instruction manual for chemists in view of the ontogeny of DFT. A chemist using DFT calculations should be aware of the fact that all approximations and simplifications of any general theory may lead to failures in computed data. Every principally correct theory, if not executed with specific care, may produce essentially wrong results and therefore erroneous predictions. Two aspects are quintessential if one seeks to successfully perform DFT calculations: a basic understanding of how the concepts and models underlying the various manifestations of DFT are built and an essential knowledge of what can be expected from DFT calculations and how to achieve the most appropriate results. Thus, we have divided the main body of our directions into two parts.

In section “[DFT: A Paradigm Shift in Theoretical Chemistry](#),” we expand on the development and philosophy of DFT. We do not present a course or textbook work on DFT; the interested reader will find a selection of references to the literature for more elaborate and detailed descriptions. Rather, we aim to illustrate the essentials of DFT in a manner that is intuitively accessible. For this reason, we will avoid mathematical equations as much as possible, incorporate formulas into the flow of the text, and only on occasion add the odd-numbered equation to the elaboration. As a consequence, we have to abandon the rigor that unambiguous and well-defined derivations require and introduce a certain degree of sloppiness. We hope that such a treatment will facilitate the flow of our arguments and emphasize that what we think are indispensable aspects of DFT.

In section “[DFT: Computational Chemistry in Action](#),” we present an analysis of the performance and applicability of DFT. We focus on a representative selection of chemical properties and system types and base our review on a selection of representing benchmarking studies, which encompass several well-established density functionals together with the most recent efforts in the field. Due to the multitude of papers that report DFT applications, our analysis is far from complete,

but we aim to present a representative snapshot of the current situation of DFT at the beginning of the second decade of the new millennium.

We will close our work with two additional short sections “DFTips” and “A Concise Guide to the Literature,” where we present the reader with a few tips how to use DFT, and with a collection of selected references as a concise guide to the by now vast literature of DFT.

DFT: A Paradigm Shift in Theoretical Chemistry

Density functional theory is primarily a theory of electronic ground-state structure, which is based on the electron density distribution $\rho(r)$. In contrast to DFT, wave function theory is an approach to electronic structure, which is based on the many-electron wave function $\Psi(\mathbf{r}_n)$. In order to put the innovation of DFT into proper perspective, we begin with a brief overview of WFT, before we illustrate the essentials and growth of DFT.

The objective of WFT is the exact solution of the time-independent Schrödinger equation (TISE), $H\Psi = E\Psi$, for a system of interest. [We recall that in quantum mechanics, associated with each measurable parameter in a physical system is an operator, and the operator associated with the energy of a system is called the Hamiltonian H . The Hamiltonian contains the operations associated with the kinetic and potential energies of all particles that comprise a system. We further note that the terms function, operator, and functional are to be understood such that a function is a prescription which maps one or more numbers to another number, an operator is a prescription which maps one function to another function, and a functional takes a function and provides a number.] The solution to the TISE yields the wave function Ψ as well as the energy E for the system of interest. In a systematic, variational search, one looks for the wave function that produces the lowest energy and arrives at a description for the system in its ground state.

If we consider a system of nuclei and N electrons, solving the TISE – within the Born–Oppenheimer separation between the apathetic ambulation of nuclei and the rapid ramble of electrons – yields the electronic molecular wave function $\Psi_{\text{el}}(\mathbf{r})$. This wave function depends explicitly on the $3N$ coordinates of all N electrons, all of which might undergo positional permutation due to repulsive Coulomb interaction. A first approximation to the challenging task of solving the TISE slackens the interaction between electrons by any exchange and thus reduces the function $\Psi_{\text{el}}(\mathbf{r})$ of $3N$ variables to a product of N functions ϕ each depending only on three variables, $\Psi_{\text{el}}(\mathbf{r}) = \prod_{i=1,N} \phi_i(\mathbf{r}_i)$. Atomic orbitals are conveniently chosen to represent the functions ϕ . However, such a *Hartree* product of atomic orbitals violates the Pauli exclusion principle due to the fermion nature of electrons, and hence the appropriate form for a system of noninteracting electrons is a single determinant $|\phi_1(\mathbf{r}_1), \dots, \phi_n(\mathbf{r}_n)|$, known as *Slater* determinant. Such a wave function, from the mathematical properties of determinants, is antisymmetric with respect to exchange of two sets of electronic variables as it should be.

One row of a *Slater* determinant carries contributions from all atomic orbitals ϕ and is commonly referred to as molecular orbital (MO) $\psi\psi^{\text{MO}}$. This approximate method for the determination of the ground-state wave function and ground-state energy is the well-known Hartree–Fock (HF) method.

Although in the HF method the electrons obey exchange as required by the Pauli exclusion principle, the electrons are noninteracting, and the movement of one electron within the system is independent from the movement of all other electrons. However, as the presence of Coulomb repulsion between electrons would suggest, the electrons move in a correlated fashion. In order to allow for electron correlation, configuration interaction (CI) is introduced in that the wave function is constructed as a linear combination of several *Slater* determinants, obtained from a permutation of electron occupancies among all MOs available. Increasing the number of *Slater* determinants increases the accuracy of the calculations, although the added accuracy comes with the price of added computational cost that often becomes the limiting factor for WFT calculations.

This brief exposition brings about two main differences between DFT and WFT. A WFT calculation in general, and increasing the accuracy of WFT calculations in particular, is computationally demanding. DFT seems to be more cost-efficient; after all, the simplest HF wave function $\Psi_{\text{el}}(\mathbf{r})$ depends on $3N$ spatial coordinates, whereas the probability distribution of electrons in space $\rho(\mathbf{r})$ depends only on three coordinates. But there exist strategies how the result of WFT can be systematically improved, whereas there is no methodical, standardized scheme to improve DFT calculations. In the following, we will explore reasons for the WFT–DFT differences.

In WFT, atoms and molecules constitute the basic systems of interest. Since the distribution and redistribution of electrons within atoms and molecules are central to chemical properties and reactivity, we now limit ourselves to systems comprised of N electrons in motion, with some two-particle interaction. The Hamiltonian of such a system reads $H = T + U + V$, where T and U denote the operators for kinetic energy and for electron–electron interaction energy, respectively. Whereas U results in the *internal* potential energy of our system, the moving electrons might in addition interact with an *external* potential, and the operation V recovers the potential energy due to this extra interaction. For systems of electrons that move in a field of fixed nuclei, the external potential V is always just the nuclear field. Systems of electrons in combination with fields of fixed nuclei represent the essential building blocks of matter such as molecules or solids. For a chemist, it might appear counterintuitive that nuclei, being an essential part of a molecule, represent an *external* potential, but the nuclear potential – although internal to a molecule – is external to the density of the moving electrons.

If we consider the external potential to be a uniformly distributed background positive charge, we arrive at the uniform or homogeneous electron gas (UEG or HEG), also known as jellium. At zero temperature, the properties of jellium depend solely upon the constant electron density distribution $\rho(\mathbf{r}) \{= \text{const.}\}$. Such a treatment of electronic density, the Thomas–Fermi (TF) model, constitutes the origins of

DFT (Parr and Yang 1989). The TF model is able to describe the kinetic energy of the UEG as functional of the electronic density, and later on Dirac added a density functional for the exchange energy as a conclusion of the Pauli principle. The UEG formalism itself provides the basis for the local density approximation (LDA).

Before we proceed, we take a small step back to WFT. At the beginning of the 1950s, Slater (1951) described the then current situation in WFT as follows: “The Hartree-Fock equations furnish the best set of one-electron wave functions for use in a self-consistent approximation to the problem of the motion of electrons in the field of atomic nuclei. However, they are so complicated to use that they have not been employed except in relatively simple cases.” Facing the decision “Do you want to calculate it, or do you want it to be accurate?,” Slater decided to replace the peculiar exchange term in the HF equations by something equivalent, yet easier to calculate. Slater used the free-electron approximation for the exchange potential, which, as Dirac has shown, could be expressed as a density functional. His new method, termed HFS, “was easy enough to apply so that we can look forward to using it even for heavy atoms” (Slater 1951), and in order to check its applicability, he performed calculations for the transition metal ion Cu^+ . The exchange potential functional derived from the exchange energy functional contains one additional, scalable parameter α , which led to the development of the $X\alpha$ – method. This model enjoyed a significant amount of popularity among physicists and is still a topic of ongoing research activities (Zope and Dunlap 2006). The HFS – or the $X\alpha$ – method – became the first practically used DFT method in chemistry.

Two points that were fundamental for the progress of DFT were already anticipated within the advancements of the HFS – or $X\alpha$ – method: (1) Every density functional method to some degree contains one or more empirical parameters. Therefore, DFT has often been regarded as “Yet another Semi-Empirical Method” (YaSEM). The Hartree–Fock–Slater model, which can be regarded as ancestor of modern DFT, is such an example. But whereas the HFS method is *intrinsically* approximate, modern DFT is in principle exact (Kohn et al. 1996). (2) Transition metal chemistry has played and continues to play a major role in the progression of DFT. The work of Baerends and Ros (1978) is representative of the transition-metal–HFS era, and Ziegler’s contributions, summarized in his early review article (Ziegler 1991), have provided the impetus that changed the perception of calculations based on densities from YaSEM to DFT, “Das Future Tool.” A review article by Cramer and Truhlar (2009) is dedicated solely to developments and progress of DFT for transition metal chemistry and summarizes the state of affairs at the end of the first decade of the new millennium.

At the heart of modern DFT is the rigorous, simple lemma of Hohenberg and Kohn (1964), which states that the specification of the ground-state density, $\rho(\mathbf{r})$, determines the external potential $V(\mathbf{r})$ uniquely (Eq. 1):

$$\rho(\mathbf{r}) \rightarrow V(\mathbf{r}) \text{ (unique)} \quad (1)$$

This first theorem by Hohenberg and Kohn (HK-I) is not difficult to prove (Parr and Yang 1989), but for a chemist, the essentials of HK-I that given a

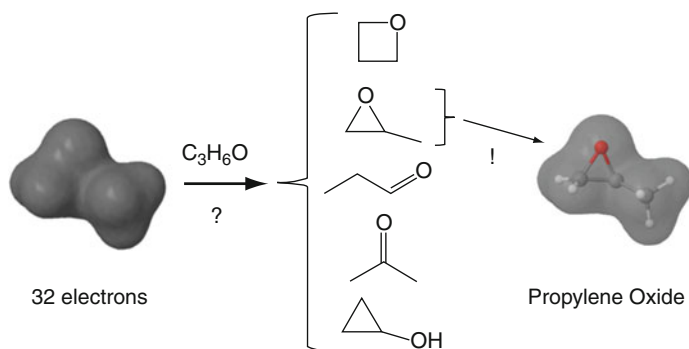


Fig. 1 Visualization of the first Hohenberg–Kohn theorem (density map drawn for a contour envelope of 0.01 a.u.)

density, only one external potential corresponds to that density, are intuitively clear. A pictorial representation of HK-I is shown in Fig. 1; we consider the density created by 32 electrons that move around the external potential created by one oxygen, three carbon, and six hydrogen nuclei. By inspection of a density map of a certain density value, it becomes obvious that of all the atomic constellations considered, only one seems to be consistent with the shape of the density map. Such a consideration reflects ideas developed in the context of “conceptual DFT” (Geerlings et al. 2003).

HK-I also expresses the fact that there is a one-to-one mapping between the potential $V(r)$, the particle density $\rho(r)$, and the ground-state wave function Ψ_0 (Eq. 2):

$$\rho(r) \longleftrightarrow V(r) \longleftrightarrow \Psi_0 \quad (2)$$

This implies that all properties of a system are functionals of the ground-state density, since any property may be determined as the expectation value of the corresponding operator. With the help of this lemma, a minimal principle for the energy as functional of $\rho(r)$ can be derived. The second Hohenberg–Kohn theorem (HK-II) provides the necessary guidelines to obtain the ground-state energy. Following HK-II, a variational principle is established, according to which the ground-state density of a system of interest can be determined.

In order to put the promise of the HK theorems that all properties of a system can be obtained from its ground-state density, into reality, one would need a construction that is computationally accessible while maintaining the formal exactness of HK-I and HK-II. To this end, Kohn and Sham (1965) introduced a fictitious system of N noninteracting electrons that have for their overall ground-state density the same density as some real system of interest where the electrons do interact. Using some aspects of HF theory, the ground-state wave function Ψ_0 of such a noninteracting system is described by a single *Slater* determinant. The orbitals, which form this

Slater determinant, known as Kohn–Sham (KS) orbitals φ^{KS} , are solutions of N single-particle equations. Following the variational principle, the ground-state energy and the ground-state density are determined from variations in φ^{KS} .

The essential contribution to the KS energy comes from the so-called exchange–correlation energy E_{XC} . It incorporates corrections to the kinetic energy due to the interacting nature of the electrons of the real system, *all* nonclassical corrections to the electron–electron repulsion, as well as electron self-interaction corrections. If E_{XC} is ignored, the physical content of the theory becomes identical to that of the *Hartree* approximation. Thus, within the KS formalism, the electronic energy of the ground state of a system of N electrons moving within an external potential of nuclei is expressed – without approximations – as a functional of the ground-state density (Eq. 3):

$$E[\rho] = T_s[\rho] + U[[\rho(1), \rho(2)]] + V_{\text{ne}}[\rho] + E_{\text{XC}}[\rho] \quad (3)$$

In Eq. 3, the first term represents the kinetic energy of the system of N noninteracting electrons, the second term corresponds to the Coulombic repulsions between the total charge distributions at two different positions within the system, and the third term accounts for nuclear–electron interactions, due to the presence of an external potential. It is the fourth term, the functional for the exchange–correlation energy E_{XC} , which is responsible for the power and magic of DFT. What makes current DFT applications approximate is the unknown analytic expression of E_{XC} , for which an approximation is needed.

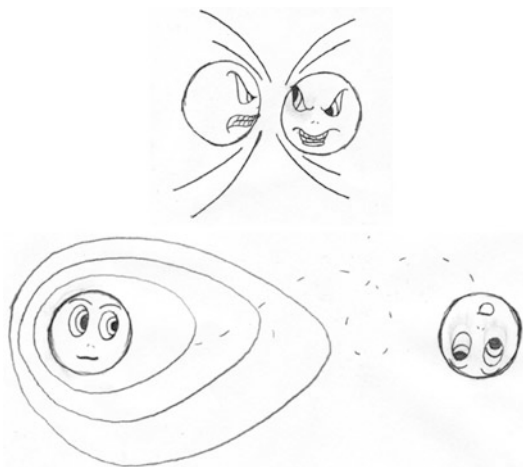
The KS formalism is closely related to the HF formalism. What differentiates the KS operator from the HF operator is the exchange–correlation potential V_{XC} . V_{XC} in turn is a functional derivative of the exchange–correlation energy E_{XC} . Furthermore, the Hamiltonian H operating on the wave function that is associated with the density of a fictitious system of N noninteracting electrons can be expressed as sum of one-electron operators.

Whereas increasing the accuracy of HF calculations is accompanied with a steep increase in computational cost, increasing the accuracy of DFT calculations apparently requires modifications in V_{XC} , which – if at all – only lead to a moderate increase in computational cost. However, whereas it is well known how to systematically improve the accuracy and quality of HF calculations, no comparable strategy exists for DFT calculations within the confines of a KS approach. It appears that a detailed knowledge of the exchange–correlation energy E_{XC} is essential for designing more accurate density functionals.

Holes and Electron Pairs

The exchange–correlation energy E_{XC} is a relatively small part of the total energy of a typical system, although it is by far the largest part of “nature’s glue” that binds atoms together (Kurth and Perdew 2000). It arises because the electrons do not

Fig. 2 Electrons in distress: While trying to maximize the attraction from the nuclei, an electron experiences enhanced repulsion from the other electron as it moves around in the molecular framework (*top*). To minimize the repulsion, each electron creates an exclusion area around itself, into which no other electron can penetrate (*bottom*) (Cartoon by Lauren Bertolino)



move randomly through the density but avoid one another. Ziegler (1995) illustrates the situation as follows: An electron will try to maximize the attraction from the nuclei and minimize the repulsion from the other electrons, as it moves around in the molecular framework. To do so, it creates an exclusion area or “no-fly zone” around itself into which no other electron can penetrate, as pictorial exemplified in Fig. 2. The exclusion zone is referred to as the exchange and correlation (XC) hole, and it is the way in which the XC hole is modeled that distinguishes one electronic theory from another. Each density functional has its own characteristic XC fingerprint.

The XC hole also determines to a large part E_{XC} , which however contains *three* different contributions. The first is the potential energy of exchange, which also should include corrections for self-exchange or self-interaction. The second is the potential energy of correlation due to the effect of Coulomb repulsion. Both potential energies are negative and determined by the nature of the XC hole. The third contribution to E_{XC} is a smaller positive kinetic energy of correlation due to the extra swerving motion of the electrons as they avoid one another (Perdew et al. 2009).

The XC hole arises from an extension of the concept of the unconditional one-electron probability density $\rho(1)$ by considering pairs of electrons and a resulting conditional probability. When a reference electron is known to be at position 1, the conditional probability $\rho^{\text{cond}}(2, 1)$ of the other electron to be at position 2 can be written as the sum of the unconditional probability $\rho(2)$ of the other electron and the XC-hole density $\rho^{\text{XC-hole}}(2, 1)$. Thus, the hole $\rho^{\text{XC-hole}}(2, 1)$ describes how the conditional density of the other electron deviates from its unconditional density $\rho(2)$.

It is instructive to have a closer look at hole profiles, and as simple example, we will consider the hydrogen molecule H_2 with only two electrons or one electron pair. In Fig. 3, hole densities for H_2 are shown; the two nuclei H_A and H_B are separated by 72 pm, and the reference electron is placed 15 pm to the left of nucleus H_B .

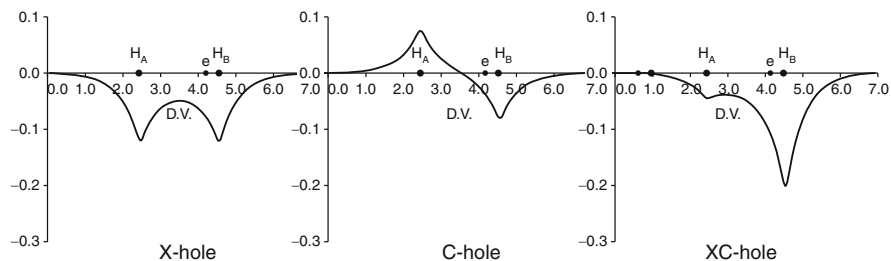


Fig. 3 Hole densities in the hydrogen molecule: Only the full XC-hole $\rho^{\text{XC-hole}}(2, 1)$ has physical meaning (Adapted from Baerends and Gritsenko (1997), with permission by the American Chemical Society)

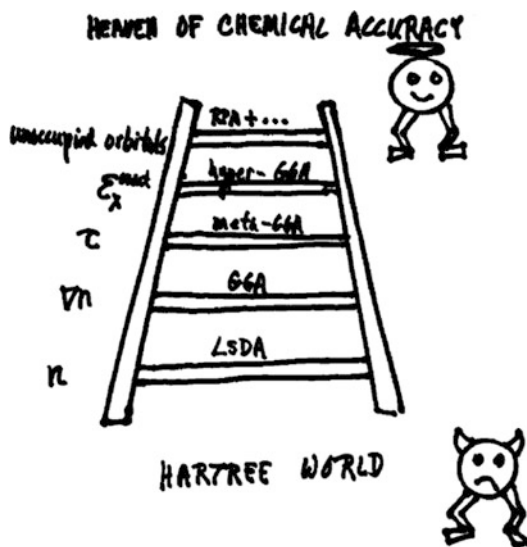
The XC hole can be split into contributions from the exchange or X hole, which arises from the fermion nature of an electron obeying the Pauli principle and the correlation or C hole due to Coulomb repulsion within the pair of electrons. [The X hole and C hole are often referred to as Fermi hole and Coulomb hole, respectively.]

The X hole puts an emphasis on the reference electron. It creates a taboo zone for the other electron with a negative $\rho^{\text{X-hole}}(2, 1)$ probability density not only around the region in space where the reference electron currently *is* but also where it *might be*. Regions in the vicinity of both nuclei H_A and H_B are declared as “no-fly zone.” The C hole on the other hand puts an emphasis on the other electron. It excludes regions where the other electron might experience Coulomb repulsion with the reference electron, but it also maps out regions where the other electron can benefit from attractive Coulomb interactions with nuclei. We see a negative $\rho^{\text{C-hole}}(2, 1)$ probability density around H_B but a positive $\rho^{\text{C-hole}}(2, 1)$ probability, a buildup of density, around H_A far away from the reference electron. Whereas the X hole and C hole illustrate exchange and correlation, only the combined XC hole has physical meaning.

Before we continue, a short remark on the use of some language is in order. Since the terms *local* and *nonlocal* are often recurred to in the context of DFT, and often with different meanings, we briefly define how the terms *local* and *nonlocal* are used in this work. An approximation is said to be *local* if its energy density and related properties at any position of interest depend only on the electron density neighborhood of the given position. Otherwise, an approximation is said to be fully *nonlocal*. [We note that some physicists separate local approximations into “strictly local” and “semilocal.”]

From an inspection of Fig. 3, it appears that both the C hole and the X hole are inherently nonlocal (the same as the HF exchange energy). The XC hole, too, must therefore be nonlocal, but its dominant contributions arise from the region around its reference electron – the XC hole appears to be more local and less nonlocal than the X or C hole. This observation already anticipates that local density functionals might be able to produce approximate models for the nonlocal XC hole.

Fig. 4 The Jacob's ladder of density functional approximations to the exchange-correlation energy (Reprinted from Perdew *et al.* (2009), with permission by the American Chemical Society)



Although it seems that there exists no systematic approach like in WFT to improve the accuracy of DFT, the advancement of density functionals depends on more precise descriptions not only of the XC hole but also of the exchange-correlation energy E_{XC} . This task can be approached in a methodical manner.

Climbing Jacob's Ladder

Perdew and Schmidt (2001) compare the development of enhanced density functionals to a climb of Jacob's ladder, leading the way from the *Hartree* world to the heaven of chemical accuracy, illustrated in Fig. 4. Each rung of the ladder adds a refinement to the approximation of the exchange-correlation energy.

First Rung: The Local Density Approximation

The first rung employs only the local densities in the description of the exchange-correlation energy. This method is known as the local density approximation (LDA). [Although vital to the fermion nature of electrons, so far we have treated spin rather nonvercal. But the issue of spin can be treated as well in density functional theory, and the local spin density approximation (LSD or LSDA) replaces the spin-averaged energy density with the energy density for a polarized homogeneous electron gas. LDA and LSDA are now often used synonymously.]

LDA takes its densities from the uniform electron gas (UEG), and an analytical form of a density functional for the exchange energy of the UEG can be derived (the same exchange energy as used in the HFS method). No such expression exists for the correlation energy, but the UEG correlation energy can be calculated numerically

and fit in various ways. One successful and popular parameterization comes from the work of Vosko, Wilk, and Nussair, referred to as VWN (Vosko et al. 1980).

LDA performs surprisingly well in predicting molecular properties that are based on relative energy differences within a given density. Molecular geometries for representative main group compounds could be reproduced in close agreement to the experiment (Versluis and Ziegler 1988). For transition metal complexes, metal–ligand separations are slightly underestimated, but still within acceptable conformity with X-ray data (Ziegler 1995). However, properties that are based on absolute energy differences between densities, such as bond energies, are not well described by LDA, where a clear overbinding tendency emerged. LDA is therefore a remarkably useful structural, though not thermochemical, tool. The disappointing performance of LDA in estimating thermochemical properties spawned the development of gradient-based methods, the second rung on Jacob’s ladder.

Second Rung: The Generalized Gradient Approximation

The second rung or generalized gradient approximation (GGA) adds the gradients of the local densities to the exchange–correlation picture. It became clear that the homogeneous electron gas is only of limited use as a model of the inhomogeneous electron density within molecules, and approximations for exchange and correlation energy were augmented by density gradients. [In the older literature, GGAs are sometimes called nonlocal (LDA/NL), since the gradient implies a directional change within the density.] Two early models for correlation (Perdew 1986) and exchange (Becke 1988a) in combination resulted in the BP86 functional, the GGA that was most influential in the early developments of transition metal DFT (Ziegler 1991). Gradient corrections are essential for a quantitative estimate of bond energies as well as metal–ligand bond distances (Ziegler 1995).

Third Rung: Meta-Functionals

The third rung adds the kinetic energy density to the description of density functionals (Tao et al. 2003) and addresses the smaller third contribution to E_{XC} . Such functionals are referred to as meta-functionals and, when built on second rung functionals, as meta-GGA (MGGA).

The first three rungs of Jacob’s ladder all represent local functionals. They often work because of proper accuracy for a slowly varying density or because of error cancellation between exchange and correlation. Error cancellation can occur because the exact XC hole is usually more localized around its reference electron than the exact X hole (compare Fig. 3). Regions in which no error cancellation is expected are regions where exchange dominates correlation (Perdew et al. 2009).

Climbing up the ladder, the approximations become more complicated, more sophisticated, and typically more accurate. Computation times increase modestly from the first to the third rungs and much more steeply after that. The added ingredients on each higher rung of the ladder can be used to satisfy more exact constraints or to achieve better agreement with experimental data (or both). These two strategies define the nonempirical and semiempirical approaches commonly

used to improve density functionals. Beginning with the fourth rung, the nature of the density functionals changes from local to nonlocal.

Fourth Rung: Hyper-Functionals

The fourth rung, which also represents the first fully nonlocal rung, adds the exact exchange energy density. Such a functional is termed hyper-GGA (HGGA). After reaching the second rung, DFT progressed rapidly and took one giant step from the second to the fourth rung, omitting meta-GGAs. Following the idea of adiabatic connection, Becke derived a functional for exchange, which contained contributions from the exact HF exchange (1993a). He then designed an advanced functional for the exchange-correlation contribution containing three parameters for its various parts, including gradient corrections for correlation, gradient corrects for exchange, as well as an exact exchange contribution (1993b). These semiempirical coefficients have been determined by a linear least-square fit to 56 atomization energies, 42 ionization potentials, eight proton affinities, and ten first-row total atomic energies. Becke's functional, combining HF theory and DFT with the use of three empirical coefficients, was the first example and initiated the evolution of so-called hybrid functionals. The hybrid functional B3LYP, based on Becke's parameterization, was to a large part responsible for the meteoric ascent of DFT during the 1990s.

While the first three rungs of Jacob's ladder require no fitting of experimental data, empiricism seems unavoidable on the fourth rung. This has caused some skepticism, and it appeared that the success of "empirical DFT" would eventually be responsible for the death of "true DFT." Gill humorously described the situation at the beginning of the new millennium in his obituary to DFT (Gill 2001). The Jacob's letter metaphor puts the addition of exact exchange to density functionals into proper perspective.

The step from the third rung to the fourth rung results in a new class of functionals, so-called HMGGAs. HMGGA functionals are currently a field of active development and appear to produce promising results. M06, for example, is a HMGGA with good accuracy for a variety of different chemical applications ranging from transition metals over main-group thermochemistry to barrier heights of chemical reactions. Thus, HMGGAs might be considered as a class of density functionals with broad applicability in chemistry (Zhao and Truhlar 2008a). Whether HMGGA is read as hybrid meta-GGA or hyper meta-GGA is a matter of taste; fortunately, both specifications result in the same acronym.

Fifth Rung

The fifth rung of Jacob's ladder adds exact correlation as new ingredient. One might think of this as an expansion of the density space of a system by adding virtual densities into the picture. One approach to this problem is the use of the random-phase approximation (RPA). RPA in DFT in turn is closely related to time-dependent DFT (TD-DFT). The essence of RPA might be described as constructing the excited states of a system as a superposition of particle-hole excitations.

When building a fifth-rung density functional for the exchange-correlation energy, the RPA utilizes full exact exchange and constructs the correlation with the

help of the unoccupied Kohn–Sham orbitals. Like the first three rungs of Jacob’s ladder, the fifth rung requires no fitting. Although the essentials of RPA originated in the 1950s, fifth-rung methodologies were considered too computationally expensive for widespread use and application. Yet improvements in hardware and algorithms also entered the development of RPA technology, and it seems that “RPA has the potential to become a building block of future generations of electronic structure methods” (Furche 2008).

In the early days of modern density functional theory, hazy clouds of ambiguity that enfolded the XC hole obscured the view of Jacob’s ladder. The existence of the third rung of Jacob’s ladder was recognized before the fourth rung entered the Jacob’s ladder picture (Becke and Roussel 1989), but at first it did not appear as a safe and secure stage for the ascent toward the heaven of chemical accuracy. Thus, although MGGAs predate HGGAs, the computational development of HGGAs predates that of MGGAs. Only after the clouds of ambiguity lifted, MGGAs became a recognized DFT approach in computational chemistry, and HMGGAs began to appear.

The Jacob’s ladder scheme is not the only way to arrive at exact functionals. When leaving the confines of ordinary KS-DFT methods, and using ideas from WFT, one arrives at *ab initio* density functional theory, the seamless connection of DFT and WFT (Bartlett et al. 2005). While these methods have not yet positively established themselves as standard approaches in computational chemistry, the fundamental conception that guided the way to the fourth rung – capturing the best of both worlds – marks the beginning of *ab initio* DFT and led to the notion of double-hybrid theory (Grimme 2006a), in which DFT calculations are supported not only by exact exchange but also by HF correlation. Along the same lines, range-separated functionals (Leininger et al. 1997) base their decision on how to deal with electron–electron correlation – the DFT or WFT way – on the central variable that describes a pair of electrons. The idea also found its entry into the realm of exchange (Tsuneda and Hirao 2014), and at the time of writing, it appears that the offspring of parent functionals is ever more growing in number.

Practical DFT and the Density Functional Zoo

A practical DFT-based calculation is in many ways similar to a traditional HF treatment in that the final outcome is a set of orbitals, the Kohn–Sham orbitals φ^{KS} . The KS orbitals are often expanded in terms of a basis set as in the traditional linear combination of atomic orbital (LCAO) approach of traditional HF methods. Most often, Gaussian-type basis functions are used to construct atomic orbitals (GTO). A major exception is the Amsterdam Density Functional (ADF) suite of programs, where Slater-type basis functions (STO) are used. ADF constitutes one of the first programs developed essentially for applications of DFT.

The evaluation of matrix elements of the Kohn–Sham exchange–correlation potential always requires at some step a 3-day numerical integration. Solutions to the problem of carrying out 3D numerical integration for polyatomic systems to

arbitrary precision (Becke 1988b) provided a major thrust for computational DFT, and proficient improvements were made in connection with developments of the ADF computer code (Boerrigter et al. 1988). The availability of economical numerical integration schemes made the choice of STOs over GTOs computationally compatible.

The local nature of the effective potential in the one-electron Kohn–Sham equations affords efficient computational schemes. During the development of ADF, the remaining Coulomb problem, the two-electron-integral “bottleneck,” has been addressed by the introduction of auxiliary basis sets, the so-called density fitting (Baerends et al. 1973).

Many of the pioneering improvements made during the development of the ADF suite of programs have become standard tools in density functional calculations, and as a result, DFT calculations perform compatible to, if not better than HF methods. We note that a density fit is not possible, when the chosen functional utilizes exact exchange.

By now, a plethora of density functionals is available for electronic structure calculations. Toward the end of the first decade of the new millennium, Sousa and coworkers have presented an authoritative review, in which they evaluate the performance of over 50 different density functionals (Sousa et al. 2007). The authors also report the percentage of occurrences of the names of different functionals in journal titles and abstracts; we interpret these numbers as measure for usage and popularity of the corresponding functional. Although new functionals appear every year, the popularity ranking seems to possess some stability within a time interval of several years. Thus, in Fig. 5, we have compiled data for the seven most popular functionals, taken from the work of Sousa and coworkers, and include a popularity pie chart as well.

FUNCTIONAL (YEAR)	TYPE	USAGE
B3LYP (1994)	HGGA	80 %
BLYP (1988)	GGA	5 %
B3PW91 (1993)	HGGA	4 %
BP86 (1988)	GGA	3 %
PBE (1996)	GGA	2 %
BPW91 (1991)	GGA	1 %
TPSS (2003)	MGGA	1 %

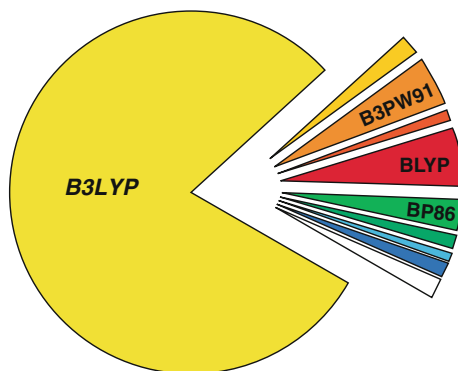


Fig. 5 Most popular density functionals at the end of the first decade of the new millennium (Data based on results of extensive literature searches (Sousa et al. 2007))

The key information conveyed in Fig. 5 is the fact that B3LYP is by far the most popular density functional in chemistry, representing 80 % of the total of occurrences of density functionals in the literature, in the period 1990–2006 (Sousa et al. 2007). Other popular density functionals such as BLYP, B3PW91, and BP86 acquire usage shares of 5 % and less and can only be considered as also-rans in the functional race.

It seems advisable to briefly talk about how to decipher the density functional code. With the advent of GGAs at the end of the 1980s, the abbreviation for each GGA functional usually consisted of two parts: the first for exchange and the second for correlation. As an example, BP86 takes its exchange contribution from the work of Becke (1988a) and correlation from the work of Perdew (1986). Similarly, BLYP breaks down into B exchange and LYP correlation. Later, when gradient corrections for exchange and correlation were often taken from the same work, the functional is usually referred to by one combined code only. The GGA functional PBE takes its gradients for exchange as well as for correlation from the work of Perdew, Burke, and Ernzerhof. The same holds true for the MGGA TPSS. Strictly spoken, the PBE functional should be referred to as PBEPBE. It is also possible that the individual parts are combined with other functionals, for example, PBELYP or BPBE. HGGAs usually contain one number that indicates the degree of parameterization when building the hybrid: B3LYP refers to a three-parameter mixing of B exchange, LYP correlation, and exact exchange, whereas B1LYP refers to a one-parameter hybrid density functional. As density functionals get more elaborate and more complex, the XC coding is not always strictly followed. The HMGGA M06 and its variations (M06-L, M06-2X, M06-HF), for example, refer to a set of functionals developed at the University of Minnesota in 2006.

The pie chart presented in Fig. 5 bears a striking resemblance to *Pac-Man*. Like the arcade game *Pac-Man*, often credited with being a landmark in video game history and virtually synonymous with video games, B3LYP has to be considered a landmark in electronic structure theory and is often used as a synonym for DFT. However, the same as there is more to video games than just *Pac-Man*, there is more to density functional theory than just B3LYP. In section “[DFT: Computational Chemistry in Action](#),” we will explore the characteristics and capabilities of density functional in more detail.

The reader, who would like to know more about the details and derivations of DFT, will find valuable information about the basics in the *A Chemist's Guide to Density Functional Theory* (Koch & Holthausen 2002) and will learn more about advanced aspects in the *A Primer in Density Functional Theory* (Fiolhais et al. 2003).

DFT: Computational Chemistry in Action

The breakthrough of DFT coincided with a rise of computational power at the end of the last millennium. CPU architectures advanced from CISC to RISC designs, supercomputers transformed from single vector processors to computing clusters.

However, as the computational power increased, the problems too became more and more demanding, and the molecules that found their way into input files for density functional programs grew bigger and bigger. Computing time remains to be a crucial factor when assessing the performance of computational methods, and linear-scaling approaches are one of the great strengths of DFT (Yang 1991). However, these techniques fall out of the scope of our review and assessment, and we begin with a comparison of computational demands of representative density functionals, following standard approaches.

Computational Performance

We start this section remarking that there is not something like “the best functional and basis set for all properties.” Rather, the specific methodological approach to be used depends from the specific problem at hand. Nevertheless, many functionals are robust enough to give rather reasonable results in a large series of chemical properties, and the scope of this section is to provide an overview of the performances of typical functionals and basis sets, trying to highlight which ones perform remarkably better or remarkably worse than the average, if this is known. Further, as a practical *vademecum*, the scope of this section is to give an overview of performances under “standard working conditions.” Thus, the focus will be on performances that can be expected when working with real-size systems (50–100 atoms including a transition metal), which requires a compromise between the computer resources available and the combination of functional and basis set used, rather than peak performances that can be reached with a sophisticated last-generation functional in combination with a very extended basis. On the other hand, there are several excellent reviews that provide an accurate and critical assessment of the various methods, with a particular focus on the best performances that can be achieved, independently from the computational cost (Sousa et al. 2007; Cramer and Truhlar 2009). Methods that currently are computationally too expensive might become the standard computational tools in the future.

Finally, the number of possible functionals is very large, so that it is more confusing than enlightening to review all of them. In addition, it might well be that the best functional for a specific problem has not been tested in the several benchmarking studies published in the literature. As a general rule, before wasting huge amounts of computer power with the wrong functional and/or basis set, it is wiser to invest some time to read the literature to find which computational method works better (or acceptably well) for a given problem and possibly have a feeling of the accuracy through test calculations on small selected systems.

To give an idea of the relative cost of the various functionals, the relative computational time required by some functionals in two standard applications such as the calculations of the energy and of the first derivatives of the energy with respect to the atomic coordinates (which must be calculated at each geometry optimization step) and the calculation of the second derivatives of the energy with respect to the atomic coordinates (which must be calculated for vibrational analysis) is reported

Table 1 Relative performance of various functionals, as implemented in the Gaussian 09 package, in the calculation of energy and gradients, or second derivatives, for an organometallic complex of formula $\text{RuC}_{31}\text{H}_{33}\text{Cl}_2\text{N}_3$

Method		E + gradients	2nd derivatives	E + gradients
		SVP		TZVP
BP86 (density fit)	GGA	1.0	1.0	1.6
BP86	GGA	1.5	1.9	4.1
PBE	GGA	1.5	1.8	4.2
B3LYP	HGGA	2.6	2.9	8.9
PBE1PBE	HGGA	3.0	3.2	9.2
TPSS (density fit)	MGGA	1.6	1.9	2.6
TPSS	MGGA	2.1	2.3	5.1
TPSSh	HMGGA	3.5	3.5	9.4
M06	HMGGA	3.6	4.0	10.6

in Table 1. The system considered in these calculations is a 70 atom Ru complex whose brute formula is $\text{RuC}_{31}\text{H}_{33}\text{Cl}_2\text{N}_3$.

The data reported in Table 1 clearly indicate that GGA calculations are computationally very effective. In addition, for DFT methods that do not rely on exact exchange, the performance can be further improved by using an auxiliary basis set to fit the electron density (usually called density fit or resolution of identity). Indeed, without this technical setup, the same GGA or MGGA functional – compare the BP86 (density fit) and the simple BP86 values in Table 1 – is roughly 50 % slower. The same consideration applies when MGGA functionals are considered; compare the TPSS (density fit) and the simple TPSS values in Table 1. We note that this technical acceleration is not possible when the Hartree–Fock exchange must be evaluated, and thus HGGA and HMGGA calculations cannot benefit from it. Generally speaking, there are marginal differences within a family of functionals, and HGGA functionals are roughly 2–3 times slower than GGA functionals. MGGA functionals, particularly when the resolution of identity technique is invoked, are roughly 50 % slower than GGA functionals and thus are quite faster than HGGA functionals. HMGGA functionals are roughly 3–4 times slower than GGA functionals. This relative speed between the different families of functionals is maintained when second derivatives are evaluated. Moving to the effect of the size of the basis set, calculations performed with a triple- ζ plus one polarization function of main group atoms results in an increase of the required computational time by a factor of 2–3 roughly. Thus, on going from an accelerated GGA functional in combination with a split valence plus one polarization function basis set, to a HMGGA functional in combination with a triple- ζ plus one polarization function, basis set result in an increase of the computational time by a factor of 10 roughly. In other words, a GGA/SVP calculation that would take one day would require more than one week, roughly, if performed at the HMGGA/TZVP level. This indicates that the selection of the most appropriate computational method (both functional and

basis set) must be a trade between the accuracy needed and the computational time (or power) available. Of course, degradation of accuracy below the level required by the specific problem at hand is not possible.

Properties of Molecular and Electronic Structure

To explore the capabilities of various density functionals, we have selected twelve representative properties of atoms and molecules. We begin with molecular structure (bond lengths, bond angles, vibrational frequencies) and basics of electronic structure (electron affinities and ionization potentials). We then proceed to the energetics of transformations of molecules (atomization energies, heats of formation, energy barriers), which will carry us to chemical bonding. The nature of the chemical bond remains a central theme in theoretical chemistry, and we discuss regular bonds as well as weak bonds, all being at the focus of ongoing research activities (bond energies, hydrogen bonding, weak interactions). The issue of spin in DFT deserves particular attention (spin states). Although DFT essentially is a ground-state theory, excited states too can be treated with density functional theory, and with our last property we briefly touch this topic (excited states). Time-dependent density functional theory (TD-DFT) is a topic in its own right, and an appropriate coverage of TD-DFT falls out of the scope of the present work. The reader will find an entry into this excited field when studying the articles compiled by Marques and coworkers (2006).

The twelve topics we selected in no way exhaust the capabilities of DFT, and any property that can be treated with WFT is in principle accessible with DFT as well. As an example, we refer the reader to a most instructive review by Neese (2009) that illustrates the capabilities of DFT in the field of molecular spectroscopy.

When evaluating the performance of computational methods, benchmarking is an essential procedure in which calculated properties are evaluated against accurate experimental data. By now, a large number of problem-specific databases have been established, which cover a wide variety of different physicochemical properties, such as proton affinities, atomization energies, barrier heights, reaction energies, and spectroscopic properties. However, these databases are not free from chemical biases and often narrowed by the structural space of chemical intuition. There is always the risk that when following established procedures, benchmark studies might lose some of their general appeal. As to avoid the dangers of casual benchmarking, Korth and Grimme have developed a “mindless” DFT benchmarking protocol. Here, the databases consist of randomly generated molecules that rely on systematic constraints (Korth and Grimme 2009) rather than on what is supposed to be chemical insight.

In the following, we will make extensive reference to published benchmark studies. However, it might well be that a particular molecule of interest to the reader is not covered within one of the existing benchmark databases, and benchmark studies in general provide good starting points for calculations, but no guarantee for correctness.

One: Bond Lengths

It is well established that almost any DFT approach, beyond LDA, is able to reproduce correctly the geometry of molecular systems composed of main group atoms. With the increase of computer resources, it is becoming customary to test the performances of various methods through calculations on a rather large set of molecules and to report statistical values. In one of such comparative studies, 17 closed-shell molecules composed by first-row atoms, for which accurate experimental geometries determined in the gas phase were available, confirmed that the performance of commonly used GGA (BLYP, BPW91, and BP86) and HGGA (B3LYP, B3PW91, and B3P86) is quite accurate, although the mean unsigned error (MUE) on the bond length obtained with the GGA functionals, between 0.015 and 0.020 Å roughly, is slightly larger than that calculated with the HGGA functionals, usually below 0.010 Å (Wang and Wilson 2004). The convergence of the geometry was tested with respect to increasing basis set size from cc-pVDZ to aug-cc-pV5Z and was shown to occur quickly. Convergence is typically reached at the triple- ζ level, and beyond this level minor fluctuations, in the order of 0.002 Å, were observed. Thus, excellent performances require that at least a triple- ζ basis set is used. Similar conclusions were reached in a different benchmark study on a dataset of 44 small molecules (Riley et al. 2007). Again, GGA functionals in combination with Pople basis sets of the 6-31G family result in MUE between 0.015 and 0.020 Å, while HGGA functional results in MUE below 0.010 Å. The MGGA functionals tested resulted in a minor improvement relative to GGA functionals, while the tested HMGGA functionals substantially reproduce the performance of HGGA functionals. This indicates that the advantage of meta-functionals certainly is not in bond distances.

To give an idea of the performance of some popular functionals, and also to show the effect of the basis set, the dependence of the O–H bond length in water is reported in Table 2 as an exemplary case. The data indicate that reasonably accurate bond lengths (within 0.02 Å from the experimental value) can be achieved with computationally cheap GGA functionals and that HGGA performs slightly better with modest basis sets. As a general trend, the HGGA bond lengths are slightly shorter than the corresponding GGA value and independent of the computational approach; slightly shorter bond lengths are predicted with basis sets of increasing quality. It is noteworthy that with the extended aug-cc-pV5Z basis set, the GGA values slightly overestimate the experimental value, whereas the HGGA values slightly underestimate it. Importantly, rather good results can be achieved also with relatively small basis sets, which allow calculating geometries for fairly large systems with a reasonable accuracy.

The very good performance of almost any functional to calculate accurately bond lengths of molecular systems composed by main group atoms is not replicated when bonds to transition metals are considered. Focusing on an extensive benchmark of 42 functionals on a database of 13 metal–ligand bond lengths, the MLBL13/05 database, all functionals provide rather good results, with MUE normally between 0.01 and 0.02 Å when a basis set of triple- ζ quality is used (Schultz et al. 2005a).

Table 2 Performance of selected functionals and basis sets in predicting the experimental value of the O–H bond length of water, 0.956 Å

	BP86	revPBE	B3LYP	TPSS	TPSSh	M06
6-31G	0.985	0.984	0.976	0.983	0.978	0.970
SVP	0.976	0.974	0.967	0.974	0.969	0.963
6-31G(d,p)	0.974	0.973	0.965	0.972	0.967	0.960
TZVP	0.972	0.971	0.962	0.969	0.965	0.958
cc-pVDZ	0.978	0.977	0.969	0.977	0.972	0.964
cc-pVTZ	0.971	0.970	0.961	0.968	0.964	0.958
cc-pVQZ	0.970	0.968	0.960	0.968	0.963	0.956
cc-pV5Z	0.970	0.968	0.960	0.967	0.963	0.956
aug-cc-pVDZ	0.974	0.974	0.965	0.973	0.970	0.961
aug-cc-pVTZ	0.971	0.970	0.962	0.969	0.964	0.958
aug-cc-pVQZ	0.970	0.968	0.961	0.967	0.963	0.956
aug-cc-pV5Z	0.970	0.968	0.960	0.967	0.963	0.956

Extending the benchmark to a database containing the bond length of eight metal–metal dimers, the TMBL8/05 databases, the situation deteriorates. GGA functionals, including the popular BP86, BLYP, and PBE functionals, still provide rather good results, with MUE between 0.02 and 0.03 Å when a basis set of triple- ζ quality is used, while reducing the quality of the basis set to double- ζ deteriorates performances remarkably, with MUE between 0.06 and 0.09 Å (Schultz et al. 2005b). In the GGA family, the HTCH and OLYP functionals, with MUE greater than 0.03 Å, should be avoided. The rather good performance of GGA functionals is not replicated by HGGA functionals, including the popular B3LYP and PBE1PBE functionals, with MUE between 0.08 and 0.09 Å. In the HGGA family, the BH&HLYP and MPW1K functionals, with MUE greater than 0.12 Å, should be avoided. Interestingly, MGGA functionals do not perform as or better than GGA functionals in predicting bond lengths, but rather worse. Indeed, including also the popular BB95 and TPSS functionals, they result in MUE greater than 0.06 Å. Introduction of HF exchange partially improves the performance of MGGA functionals, and the tested HMGGA functionals, including the B1B95 and the TPSSh functionals, result in MUE between 0.03 and 0.07 Å (Schultz et al. 2005a). Finally, the M06 functional performs particularly poor, with a MUE of 0.131 Å (Zhao and Truhlar 2008b).

To give an idea of the performance of some popular functionals in the calculation of the M–ligand distances and of the effect of the metal on the bond distance of the ubiquitous CO ligand, analysis of these distances in three typical binary carbonyl complexes involving first-row transition metals is reported in Table 3.

Basically, all the functionals reproduce the experimental M–CO distances well within 0.02 Å, but many of the functionals tested underestimate the difference in the axial and equatorial Fe–CO distances. In this respect, HGGA and HMGGA functionals seem to perform slightly better, although there is quite a debate on the exact

Table 3 Performance of selected functionals, in combination with the TZVP basis set on all the atoms, in predicting the experimental value of the M–C and C–O bond length (in Å) in three first-row M(CO)_n complexes

		Ni(CO) ₄		Fe(CO) ₅		Cr(CO) ₆	
		M–C	C–O	M–C _{eq} , M–C _{ax}	C–O _{eq} , C–O _{ax}	M–C	C–O
Exp.		1.838	1.141	1.803, 1.811	1.133, 1.117	1.918	1.141
BP86	GGA	1.828	1.151	1.809, 1.810	1.157, 1.153	1.910	1.155
PW91	GGA	1.824	1.149	1.805, 1.806	1.154, 1.151	1.906	1.153
revPBE	GGA	1.829	1.150	1.809, 1.810	1.155, 1.153	1.911	1.154
B3LYP	HGGA	1.845	1.137	1.820, 1.828	1.142, 1.138	1.926	1.141
PBE1PBE	HGGA	1.822	1.134	1.796, 1.805	1.140, 1.136	1.900	1.138
B98	HGGA	1.839	1.137	1.813, 1.820	1.142, 1.138	1.915	1.141
TPSS	MGGA	1.830	1.149	1.813, 1.816	1.154, 1.151	1.918	1.152
mPWKCIS	MGGA	1.832	1.150	1.810, 1.811	1.156, 1.153	1.911	1.154
BB95	MGGA	1.833	1.150	1.811, 1.811	1.156, 1.153	1.912	1.154
TPSSh	HMGGA	1.828	1.142	1.809, 1.815	1.147, 1.144	1.915	1.146
M06	HMGGA	1.848	1.133	1.820, 1.821	1.139, 1.135	1.920	1.138

assignment of the Fe–CO distances in Fe(CO)₅. Similar good behavior is shown by all the functionals in the prediction of the CO distance when bonded to a transition metal, although the GGA and MGGA functionals tested yield systematically longer CO distances. In this respect, HGGA and HMGGA functionals do perform slightly better.

Two: Bond Angles

The good performance of almost every functional to predict correctly bond lengths of molecular systems composed by main group atoms is confirmed in the case of bond angles. Again, a benchmark study on 17 closed-shell molecules composed by first-row atoms, for which accurate experimental geometries determined in the gas phase was available, confirmed that the performance of commonly used GGA (BLYP, BPW91, and BP86) and HGGA (B3LYP, B3PW91, and B3P86) functionals is quite accurate, with a MUE between 1.0° and 1.5° (Wang and Wilson 2004). Differently from bond lengths, GGA and HGGA methods perform rather similarly on bond angles. Also for bond angles, the convergence was tested with respect to increasing basis set size from cc-pVDZ to aug-cc-pV5Z and was shown to occur quickly, and again convergence is typically reached at the triple- ζ level. Similar conclusions were reached in a different benchmark study on a dataset of 44 small molecules (Riley et al. 2007). All the functionals considered resulted in MUE between 1.0° and 1.5°, independent of the functional used.

To give an idea of the performance of some popular GGA and HGGA functionals, and also to show the effect of the basis set, the dependence of the H–O–H angle in water is reported in Table 4 as an exemplary case. Accurate bond angles (within 1.0° from the experimental value) can be achieved with all functionals and moderate basis sets.

Table 4 Performance of selected functionals and basis sets in predicting the experimental value of the H–O–H angle of water, 105.2°

	BP86	revPBE	B3LYP	TPSS	TPSSh	M06
6-31G	107.2	107.1	108.3	107.1	107.7	109.4
SVP	102.2	102.1	103.1	102.3	102.7	103.3
6-31G(d,p)	103.1	103.0	104.0	103.3	103.7	104.5
TZVP	104.1	104.1	105.1	104.3	104.6	104.9
cc-pVDZ	101.7	101.7	102.7	104.9	102.3	102.8
cc-pVTZ	103.6	103.5	104.5	103.7	104.0	104.4
cc-pVQZ	103.9	104.0	104.9	104.0	104.3	104.8
cc-pV5Z	104.2	104.3	105.1	104.3	104.5	105.0
aug-cc-pVDZ	103.8	103.8	104.8	103.8	104.1	104.7
aug-cc-pVTZ	104.2	104.2	105.0	104.3	104.6	104.9
aug-cc-pVQZ	104.2	104.3	105.1	104.3	104.6	104.9
aug-cc-pV5Z	104.2	104.3	105.1	104.3	104.6	105.1

Three: Vibrational Frequencies

Benchmarking various DFT methods to reproduce accurately vibrational frequencies of 35 molecular systems composed by main group atoms revealed that GGA methods, with a MUE of roughly 40 cm^{-1} , are among the most accurate functionals (Riley et al. 2007). Indeed, the performance of several HGGA methods was at least 20 cm^{-1} worse, with MUE between 60 and 80 cm^{-1} , and meta-functionals are not an improvement. As for other geometrical properties, accurate performance requires that a triple- ζ basis set be used. The GGA functionals also performed better than HGGA functionals in the prediction of the vibrational frequency in nine homonuclear 3d metal dimers, with a MUE around 100 cm^{-1} for BLYP and BP86 and around 120 cm^{-1} for B3LYP and B3P86. Nevertheless, both families of functionals resulted in a rather large deviation from accurate data (Barden et al. 2000).

The performance of various functionals in the prediction of the CO stretching frequency in typical binary carbonyl complexes with first-row transition metals is exemplified in Table 5. Simple GGA and also MGGA functionals perform better and are able to capture the experimental value with an accuracy of roughly $50\text{--}100\text{ cm}^{-1}$, while Hartree–Fock exchange seems to deteriorate results, since the HGGA and HMGGA functionals reproduce the experimental value with an accuracy of roughly $150\text{--}200\text{ cm}^{-1}$. In terms of percent, the GGA and MGGA functionals overestimate the experimental values by 3–4 %, while the HGGA and HMGGA functional by 6–10 %. While these results may seem quite accurate, almost all the functionals considered are unable differentiate too little between metals. In fact, the experimental value decreases by 58.1 cm^{-1} on going from $\text{Ni}(\text{CO})_4$ to $\text{Cr}(\text{CO})_6$, but the functionals examined are unable to capture this difference. The best performing are the B98 functional, with a difference of merely 15.0 cm^{-1} , and the HGGA and HMGGA with differences slightly smaller than 10 cm^{-1} .

On the other hand, the simple BP86 GGA functional has been also tested in the prediction of the CO stretching frequency in rather large organometallic complexes.

Table 5 Performance of selected functionals, in combination with the TZVP basis set on all the atoms, in predicting the symmetric frequency of the CO stretching mode in selected first-row transition metal binary carbonyl complexes. Next to each calculated frequency is reported the %error calculated as $\%err = 100 * (v_{Exp.} / v_{DFT})$. The final column reports the difference between the $Ni(CO)_4$ and the $Cr(CO)_6$ frequencies

Method		$Ni(CO)_4$		$Fe(CO)_5$		$Cr(CO)_6$		$v_{Ni} - v_{Cr}$
Exp.		2061.3	%err	2038.1	%err	2003.0	%err	58.1
BP86	GGA	2104.0	2.1	2103.0	3.2	2102.0	4.9	2.0
PW91	GGA	2116.0	2.7	2115.0	3.8	2115.0	5.6	1.0
revPBE	GGA	2110.0	2.4	2109.0	3.5	2109.0	5.3	1.0
B3LYP	HGGA	2194.0	6.5	2184.0	7.2	2185.0	9.1	9.0
PBE1PBE	HGGA	2232.0	8.3	2223.0	9.1	2223.0	11.0	9.0
B98	HGGA	2212.0	7.3	2202.0	8.0	2197.0	9.7	15.0
TPSS	MGGA	2121.0	2.9	2116.0	3.8	2117.0	5.7	4.0
mPWKCIS	MGGA	2102.0	2.0	2102.0	3.1	2102.0	4.9	0.0
BB95	MGGA	2100.0	1.9	2100.0	3.0	2100.0	4.8	0.0
TPSSh	HMGGA	2169.0	5.2	2162.0	6.1	2162.0	7.9	7.0
M06	HMGGA	2234.0	8.4	2224.0	9.1	2226.0	11.1	8.0

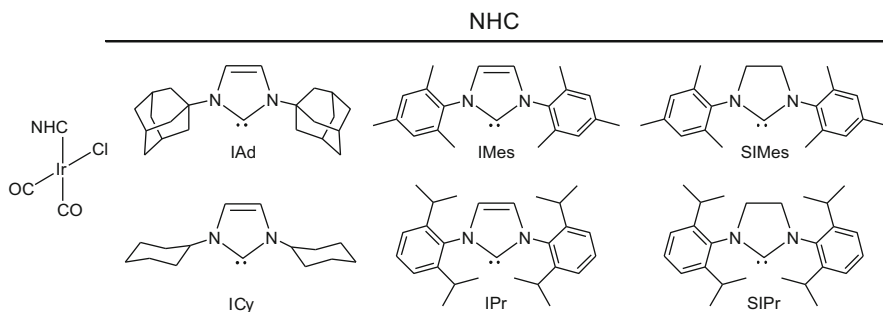


Fig. 6 Iridium complexes that bear N-heterocyclic carbene ligands

For these large and computationally demanding systems, which are displayed in Fig. 6, computationally effective methods are needed.

The data reported in Table 6 clearly show the very good performance of the BP86 functional, which is able to reproduce the higher frequency symmetric CO stretching with an error of roughly 20 cm^{-1} only and the lower frequency asymmetric CO stretching with an error of roughly 30 cm^{-1} . Further, despite the poor performances discussed above in the ability of GGA functionals to differentiate between different binary $M(CO)_n$ complexes, comparison of the saturated N-heterocyclic carbene complexes (SIPr and SIMes) with their unsaturated counterparts (IPr and IMes) indicates that the DFT values reproduce the experimental finding that both CO stretches are about 1 or 2 cm^{-1} smaller in the complexes with the unsaturated N-heterocyclic carbene ligand (Kelly et al. 2008).

Table 6 Experimental and DFT calculated CO stretching frequencies, in cm^{-1} , in several N-heterocyclic carbene complexes

Method	Experimental		BP86	
	(IAd)Ir(CO) ₂ Cl	2063.4	1979.8	2082
(ICy)Ir(CO) ₂ Cl	2064.8	1981.2	2083	2015
(IPr)Ir(CO) ₂ Cl	2066.8	1981.0	2083	2005
(SIPr)Ir(CO) ₂ Cl	2068.0	1981.8	2084	2007
(IMes)Ir(CO) ₂ Cl	2066.4	1979.8	2084	2007
(SIMes)Ir(CO) ₂ Cl	2068.0	1981.8	2085	2008

Table 7 Calculated and experimental electron affinities, in kcal/mol, for neutral late transition metal M(CO)_n complexes

System	Experimental	BP86/6-311G(d)	B3LYP/6-311G(d)
Mn(CO)	25.7	25.0	24.0
Fe(CO)	26.7	25.7	21.8
Ni(CO)	18.5	19.1	13.8
Fe(CO) ₂	28.1	32.6	–
Ni(CO) ₂	14.8	19.9	16.0
Ni(CO) ₃	24.8	35.3	27.9

Four: Electron Affinities and Ionization Potentials

A benchmark study of 24 molecules from the G2/97 dataset, with the addition of PO₂, indicated that, with some exceptions, DFT methods reproduce the electron affinity of molecular systems composed by main group atoms with a MUE close to 4 kcal/mol. Some hybrid functionals, such as the B98 functional, performs slightly better (MUE = 3.15 kcal/mol), while GGA functionals with the P86 correlation term usually perform rather poorly, with MUE around 6 kcal/mol (Riley et al. 2007). Finally, meta-functionals are not an improvement. As for any molecular system with a negative charge, the calculation of electron affinity requires that basis set containing diffuse functions are used. Moving to ionization potentials, the performance of the various methods on 36 molecules from the G2/97 dataset, with the addition of PO₂, substantially replicates that found for electron affinities, although the MUE, between 5 and 6 kcal/mol, is slightly larger (Wang and Wilson 2004). The B1B95 MHGGA functional provided the best performance, with a MUE of 4.25 kcal/mol.

Moving to selected cases (see Table 7), the BP86 GGA functional seems quite more accurate than the HGGA B3LYP functional in predicting the electron affinity of highly unsaturated late transition metal M(CO)_n complexes, but the HGGA functional seems to be more accurate when the unsaturation at the metal is reduced (Zhou et al. 2001). These results exemplify the difficulty to extract trends from benchmarks, if the specific case at hand has not been included in the testing dataset.

Five: Atomization Energies

A benchmark study on 17 first-row closed-shell molecules indicated that in the calculation of the atomization energy with the Dunning correlation-consistent basis sets, the HGGA functionals, with a MUE of roughly 2.2 kcal/mol, outperform GGA functionals, such as BLYP (MUE = 7.2 kcal/mol), with BP86 performing particularly bad (MUE = 16.2 kcal/mol) (Wang and Wilson 2004). In another benchmark study on the atomization energy of a dataset composed by 109 main groups of organic and inorganic molecules, calculated with the MG3S basis set, HGGA functionals were again confirmed to perform well, with MUE smaller than 1.0 kcal/mol, while GGA functionals, such as the PBE, with a MUE of 3.0 kcal/mol, again performed poorly. The peak performance of 0.40 kcal/mol was produced with the B1B95 functional, while the classic B3LYP resulted in a MUE of 0.91 kcal/mol. Finally, HMGGA functionals were shown to perform similarly to HGGA functionals (Zhao and Truhlar 2005a). A set of range-corrected functionals, including the CAM-B3LYP, the LC- ω PBE, and the ω B97 family, also performed particularly well, with MUE between 0.51 and 0.89 kcal/mol (Peverati and Truhlar 2014). On the other hand, GGA functionals perform well when the atomization energy of 9 metal dimers in the TMAE9/05 database is calculated, with MUE in the range of 5–8 kcal/mol. HGGA functionals, instead, resulted in MUE around 15–30 kcal/mol, with only the B97-1, B97-2, and B98 functionals with MUE below 10 kcal/mol. MGGA and HMGGA functionals substantially replicate the results obtained with non-meta-functionals, although the M05 and M06 functionals result in the very low MUE of 6.9 and 4.7 kcal/mol, respectively (Zhao and Truhlar 2008b). These results are another indication that hybrid functionals usually perform better for molecules composed by main group atoms, whereas GGA and MGGA functionals perform better for transition metal chemistry.

Six: Heats of Formation

The accurate calculation of this property, even for rather simple molecular systems composed by main group atoms, still represents a challenge for several functionals. Indeed, the MUE on the heat of formation in a benchmark study on 24 molecules from the G2/97 dataset, with the addition of PO₂, spans the rather broad range of 10–30 kcal/mol, even if basis set of rather good quality, such as the 6-31++G* basis set, are used. Functionals to be avoided are those containing the P98 correlation term among the GGA, the B1LYP and the B98 among the MGGA, the PBEKICIS among the MGGA, and the BBIK among the HMGGA. On the other side, good performances were obtained with the PBELYP, PW91LYP, MPWPW91, and MPWPBE GGA functionals, the PBE1PBE and B3PW91 HGGA functionals, and finally the TPSS and the TPSSKICIS MGGA functionals. With these functionals, the MUE on the heats of formation usually is below 10 kcal/mol (Riley et al. 2007). Augmented correlation-consistent Dunning-type basis sets usually lead to slightly better performances. Nevertheless, some computational results have cast a shadow on the common procedure to test heats of formation on small molecules. Indeed, it has been shown that almost all the functionals, including the popular

B3LYP functional, are unable to predict correctly the heats of formation of *n*-alkanes (Curtiss et al. 1997, 2000). Due to the inability to describe properly long-range attractive dispersion interaction, practically all functionals introduce a systematic error in the calculation of the isodesmic stabilization energy (Eq. 4). This systematic error ranges between 0.90 kcal/mol for the HMGGGA MPWB1K functional and 1.82 kcal/mol for the OLYP functional. The BP86 and PBE GGA functionals result in an error of 1.25 and 1.08 kcal/mol, respectively, while the HGGA functional B3LYP results in an error of 1.33 kcal/mol. While these errors may seem to be not dramatic, they are per CH₂ unit, so that the errors are between 10 and 20 kcal/mol for a simple molecule such as *n*-decane (Wodrich et al. 2006).



Seven: Energy Barriers

The performance of a series of functionals was tested to reproduce the barrier height for a series of reactions. Starting from hydrogen transfer reactions, the forward and backward barrier for the following 3 reactions OH • + CH₄ → CH₃• + H₂O, OH • + H • → O (³P) + H₂, and H • + H₂S → HS • + H₂, which constitute the BH6 database (Lynch and Truhlar 2003), GGA functionals systematically underestimate the barriers, with a MUE between 7.8 and 9.4 kcal/mol. Better results were obtained with HGGA functionals, with MUE around 4.5–5.0 kcal/mol. The mPWPW91 MGGA functional is not a clear improvement, with a MUE of 8.5 kcal/mol, while HMGGGA functionals such as the mPW1PW91 and the MPW1K perform better, with MUE of 3.9 and 1.4 kcal/mol, respectively (Zhao et al. 2004). Testing the functional on the larger BH42/04 database, consisting of 42 transition-state barrier heights of hydrogen transfer reactions in mostly open-shell systems, gave substantially similar results (Zhao & Truhlar 2004). The HGGA functionals underestimate barriers by a MUE of roughly 4 kcal/mol, while HMGGGA functionals give better results, with peak performances from the BB1K, XB1K, and MPWB1K resulting in MUE around 1.2–1.3 kcal/mol. Similar conclusions are achieved when the HTBH38 database, which contains 38 transition-state barrier heights for 19 hydrogen transfer reactions, 18 of which involve radicals as reactant and product, is considered. In this case, 10 examined HGGA functionals result in an average MUE of 3.7 kcal/mol, while 7 range-corrected functionals slightly decrease this value to 2.7 kcal/mol (Peverati and Truhlar 2014). Moving to 38 transition-state barrier heights for non-hydrogen transfer reactions constituting the NHTBH38/04 database, and comprising 12 barrier heights for heavy-atom transfer reactions, 16 barrier heights for nucleophilic substitution reactions, and 10 barrier heights for non nucleophilic substitution unimolecular and association reactions, GGA functionals, such as the PBE, still approximate barriers severely, with a MUE of 8.64 kcal/mol. HGGA functionals, such as the B3LYP and the PBE1PBE functionals, perform better, although the MUE, around 4 kcal/mol, still is quite large (Zhao et al. 2005). MGGA functionals perform similarly to the GGA functionals, while the performance of HMGGGA is rather scattered. For example, the TPSSh and the MPW1KCIS functionals do not perform impressively, with MUE of 6.6 and

Table 8 Selected mean errors for HTBH38 and NHTBH38 database

Method		Hydrogen transfer	Heavy atom transfer	Nucleophilic substitution	Unimolecular and association
PBE	GGA	9.32	14.93	6.97	3.35
BLYP	GGA	7.52	14.66	8.40	3.51
B3LYP	HGGA	4.23	8.49	3.25	2.02
PBE1PBE	HGGA	4.22	6.62	2.05	2.16
TPSS	MGGA	7.71	14.65	7.75	4.04
TPSSK CIS	MGGA	4.69	9.26	4.88	2.12
TPSSh	HMGGA	5.97	11.51	5.78	3.23
BB1K	HMGGA	1.16	1.58	1.30	1.44

4.6 kcal/mol, respectively, while the MPW1K and the BB1K are very accurate, with MUE of 1.3 and 1.2 kcal/mol, respectively (Zhao and Truhlar 2005a). As for hydrogen transfer reactions, comparison between 7 range-separated functionals and 10 HGGA functionals indicated that the range-separated functionals, with an average MUE of 2.9 kcal/mol, perform slightly better than the HGGA functional, with an average MUE of 3.7 kcal/mol. However, the peak performance of the MPW1K and B97-3 HGGA functionals, with MUE around 1.5 kcal/mol, is not reached by any of the range-separated functionals, with a minimum MUE of 2.41 kcal/mol achieved by the ω B97 functional (Peverati and Truhlar 2014).

The breakdown of these cumulative MUEs for selected functionals is reported in Table 8. Clearly, the most problematic cases are the transfer reaction of both hydrogen and heavy atoms. With the exception of the BB1K, all the other functionals fail to a large and embarrassing extent. For the HTBH38 dataset, the M06 and M06-2X HMGGA functionals, with MUE of 2.00 and 1.13 kcal/mol, also perform well (Zhao and Truhlar 2005a). Barrier heights of nucleophilic substitutions are predicted better, although GGA functionals still result in a too large MUE. Finally, the barrier heights of unimolecular and association reactions are predicted with reasonable accuracy by all functionals. Again, the performance of the BB1K functional is extremely good for the four classes of reactions considered. Similarly good and well-balanced performances are also obtained with other HMGGA functionals such as the PW6BK and the MPWB1K functionals.

Similar extensive tests on another dataset, comprising the barrier height for 23 reactions of small systems with a radical transition state, also highlighted the good performances of the HMGGA BB1K functional, with a MUE of 1.05 kcal/mol, in predicting energy barriers. However, when the same functionals were tested on a dataset of 6 barrier heights of larger systems with a singlet transition state, the best performance was obtained with the simple HGGA B1LYP functional, with a MUE of 2.58 kcal/mol, and the performance of the B3LYP functional, with a MUE of 3.10 kcal/mol, was slightly worse, while all the HMGGA functionals tested, including the BB1K functional, resulted in MUE greater than 4 kcal/mol (Riley et al. 2007).

Table 9 First metal–carbonyl dissociation energy, in kcal/mol, for selected first-row transition metal systems. DFT values, calculated with the TZVP basis set on the metal, and the SVP or TZVP basis sets on CO

		Ni(CO) ₄		Fe(CO) ₅		Cr(CO) ₆	
Exp.		24.9±2 ^a		41.5±2 ^b		36.8±2 ^b	
Method		SVP	TZVP	SVP	TZVP	SVP	TZVP
BP86	GGA	32.3	29.4	50.1	46.8	46.0	42.8
PW91	GGA	34.3	31.4	52.8	49.4	48.6	45.5
revPBE	GGA	33.9	30.9	52.1	48.6	48.2	44.9
B3LYP	HGGA	24.2	21.3	42.1	38.8	39.8	36.5
PBE1PBE	HGGA	29.2	26.4	50.1	47.1	46.1	43.4
B98	HGGA	25.6	22.8	44.7	41.8	41.5	38.6
TPSS	MGGA	32.9	30.5	50.8	48.1	46.2	43.8
mPWKCIS	MGGA	30.3	27.3	48.4	44.8	44.7	41.3
BB95	MGGA	30.4	27.4	48.8	45.2	45.4	42.4
TPSSh	HMGGA	31.2	28.6	50.0	47.4	45.5	43.0
M06	HMGGA	25.6	22.5	44.2	40.3	45.0	41.7

Eight: Bond Energies

The performance of several functionals to predict bond energies on a database of 21 metal–ligand bond energies, the MBL21/05 database (Schultz et al. 2005a), indicated that GGA functionals predict metal–ligand bond energies with a MUE between 6 and 12 kcal/mol, with the BLYP, PBE, and BP86 among the worst, and that performances can be quite better at the HGGA level, with MUE normally in the 5–7 kcal/mol range. MGGA and HMGGA functionals do not offer an improvement, as exemplified by the MUE of the TPSS and TPSSh functionals, 7.9 and 5.5 kcal/mol, respectively, and by the M06 functional, with a MUE of 5.4 kcal/mol (Zhao and Truhlar 2008b).

Taking again some selected binary carbonyl complexes of first-row transition metals as examples, the performance of various functionals in the prediction of the first bond dissociation energy is reported in Table 9. The first clean result is that the binding energy is strongly affected by the basis set quality, and a triple- ζ basis set yields binding energies that are roughly 3–4 kcal/mol lower. Focusing on the TZVP results, the GGA functionals examined consistently overestimate the CO binding energy in the three complexes by roughly 3–6 kcal/mol. The PBE1PBE HGGA functional performs like the GGA functionals examined, whereas the B3LYP, and particularly the B98 functional, is among the best performing functionals. Meta-functionals only offer a marginal improvement.

Nine: Hydrogen Bonding

Benchmarking various DFT methods to reproduce the H–bond interaction energy of ten systems composed by main group atoms revealed that HGGA functionals generally perform better than GGA functionals, with MUE around 0.5 kcal/mol, and the B3LYP functional among the best. GGA functionals, instead, result in

Table 10 H-bond interaction energy, in kcal/mol, of selected nucleic acid base pairs

Base pair	MP2/aug-cc-pVQZ	B3LYP/6-31G(d,p)	PW91/6-31G(d,p)
G:C Watson–Crick	−27.7	−25.5	−27.7
A:T Watson–Crick	−15.1	−12.3	−14.5
G:U Wobble	−15.7	−13.4	−14.8
U:U Calcutta	−9.6	−7.5	−8.7

MUE between 0.5 and 2.0 kcal/mol, with the BLYP, MPWPW91, and MPWPBE among the best performing (Riley et al. 2007). MGGA and HMGGA functionals perform close to the B3LYP functional. In all cases, the quality of the basis sets had a strong impact on the quality of the results, and at least a 6-31G(d,p) basis set should be used. Further tests were performed on a database consisting of the binding energies of 6 hydrogen bonding dimers, the HB6/04 database (Zhao and Truhlar 2005b). Also these tests indicated that the performance of GGA functionals is generally not very impressive, with MUE normally between 0.5 and 4 kcal/mol, with the exception of the very well-performing PBE functional, with a MUE of 0.25 kcal/mol only. On the average HGGGA performs better, with MUE between 0.3 and 1.0 kcal/mol, with the PBE1PBE functional, with a MUE of 0.27 kcal/mol, among the best, while the B3LYP functional, with a MUE of 0.87 kcal/mol, does not perform impressively. Meta-GGA and meta-HGGGA functionals are not a great improvement. These results have been achieved with the rather large aug-cc-pVTZ basis set, using the 6-31+G(d,p) basis set leading to slightly lower performances.

Extensive tests of the performance of the B3LYP and PW91 functionals in the case of H-bonded nucleic acid–base pairs indicated that the PW91 GGA functional replicates with better accuracy the MP2/aug-cc-pVQZ values than the B3LYP HGGGA functional (Sponer et al. 2004). The B3LYP calculations underestimate the interaction energies by few kcal/mol with relative error of 2.2 kcal/mol. Representative values for Watson–Crick (G–C, A–T) purine–pyrimidine (G–U) and pyrimidine–pyrimidine (U–U) pairs are presented in Table 10.

Ten: Weak Interactions

For a long time, one of the well-known failures of DFT was in the description of long-range dispersion forces, which are the glue that keeps together Van der Waals complexes and are at the basis of the well-known $\pi - \pi$ stacking interactions. Members of the former family are the rare-gas dimers that almost invariably are predicted to be unbound (or bound by effect of the basis set superposition error) by many GGA and HGGGA functionals. Indeed, an extensive benchmark on the WI4/04 datasets, composed by four rare-gas dimers (HeNe, NeNe, HeAr, and NeAr) indicated that the B3LYP functional simply does not predict a energy well for three out of the four dimers if the basis set superposition error is corrected by the counterpoise procedure and that even including the basis set superposition error the interaction energy is underestimated by 0.23 kcal/mol, and equilibrium distances are overestimated by more than 1 Å. On the other hand, HMGGA functionals were shown to reproduce with good accuracy both the interaction energy and the

equilibrium distance of these dimers, with MUE on the counterpoise-corrected binding energies below 0.05 kcal/mol by several functionals, such as the X1B95 and the MPWB1K, and with MUE on the equilibrium distances as low as 0.10 Å, which is a remarkable result for weakly interacting systems (Zhao and Truhlar 2004).

Enlarging the dataset to include also organic molecules, such as the C₆H₆-Ne, and the (C₂H₄)₂ and (CH₄)₂ dimers, indicated that standard GGA functionals systematically underestimate the interaction energy by a MUE of roughly 0.4–1.0 kcal/mol, with the exception of the PBE functional, with a MUE of 0.30 kcal/mol. MGGA functionals do not perform much better, while a small improvement is shown by HGGA functionals, with the best results from B97-1 with a MUE of 0.20 kcal/mol. HMGGA functionals do not offer better performances, although the largest MUE is reduced to 0.64 kcal/mol (Zhao and Truhlar 2005b). Finally, better performances were shown by the M05-2X and M06-2X functionals, with a MUE of only 0.03 and 0.09 kcal/mol, respectively, when applied to a dataset formed by the four rare-gas dimers above, C₆H₆ – Ne, CH₄ – Ne, and the (CH₄)₂ dimer. Surprisingly, on this reduced dataset, the well-performing B97-1 functional was not tested (Zhao and Truhlar 2005b).

Moving to π – π stacking interactions, the performance of a series of functionals was tested on the PPS5/05 database, which consists of binding energies of five π – π stacking complexes, namely, (C₂H₂)₂, (C₂H₄)₂, and sandwich, T-shaped, and parallel-displaced (C₆H₆)₂ (Zhao and Truhlar 2005a). Basically, all the functionals tested underestimated the binding energies in the PPS5/05 database, with MUE in the range of 1.0–3.8 kcal/mol. This is not a minor failure, since the average binding energy in the PPS5/05 dataset amounts to 2.02 kcal/mol only. The PBE GGA functional, with a MUE of 2.1 kcal/mol, performs better than the B3LYP HGGA functional, which results in a MUE of 3.2 kcal/mol. The best performances, slightly above 1 kcal/mol, are obtained with the PWB6K, PW6B95, and MPWB1K. A clear improvement is obtained with the M06-2X functional, with a MUE of 0.30 kcal/mol only. Reducing the amount of Hartree–Fock exchange deteriorates somewhat performances, as evidenced by the MUE of 0.60 kcal/mol yielded by the M06 functional (Zhao and Truhlar 2008b).

Focusing on π – π stacking interactions between nucleic acid bases, which are fundamental to describe properly the base–base stacking in nucleic acids, the popular B3LYP is simply unable to find a minimum corresponding to the A...T as well as the G...C base pairs in a stacked geometry. Of the six functionals tested, only the MPWB1K and the PWB6K functionals resulted in stable stacked base pairs, with an underestimation of the stacking energy of roughly 2–3 kcal/mol (Zhao & Truhlar 2005c).

Range-separated functionals, which were developed with the particular intent of describing properly the exchange term at long distances, offer a clear improvement over the performance of classic HGGA functionals on weak interacting systems. For example, when the performance of different families of functionals is compared through the NCCE31/05 database, which comprises the various databases described above, seven range-separated functionals, including the CAM-B3LYP, the LC- ω PBE, and the ω B97 functionals, with an average MUE of

0.61 kcal/mol, perform clearly better than 10 HGGA functionals, with a MUE of 0.92 kcal/mol. However, it is important to remark that on the same NCCE31/05 database, functionals fitted to reproduce non-covalent interactions, as the M06 functional, with a MUE of 0.41 kcal/mol, also perform remarkably well (Peeverati and Truhlar 2014).

As a final remark, we note that, based on an idea earlier proposed for Hartree–Fock calculations (Hepburn et al. 1975; Ahlrichs et al. 1977), the addition of an empirical $C_6 \cdot R^{-6}$ dispersion term represents a very simple cure to the failure of standard density functionals to predict dispersion interactions (Grimme 2006b). This approach is treated extensively in Chapter 12, so we do not discuss it here in details. We only remark the very good performance of a range-separated functional that also include an empirical dispersion term, such as the wB97X-D functional, on the NCCE31/05 database, with a MUE of only 0.32 kcal/mol.

Eleven: Spin States

The problem of a reliable prediction of the relative ordering of different spin states in transition metal complexes remains a tough challenge for DFT – not only for a quantitative judgment (energy separation between the different spin states) but even for a qualitatively assessment (correct prediction of the spin ground state). There is general consensus that GGA functionals overestimate the stability of low-spin states, whereas the inclusion of Hartree–Fock exchange in the HGGA functionals results in an overestimation of the stability of high spin states (Ghosh 2006; Swart 2008). This led to the development of the B3LYP* functional, in which the amount of the Hartree–Fock exchange is reduced from 20 % of the original B3LYP functional to 15 % (Reiher et al. 2001). Furthermore, it has been suggested that results obtained with Slater-type basis sets converge rapidly with the basis set size, while this convergence in case of Gaussian-type basis sets is much slower, and demanding basis sets like Dunning’s correlation-consistent basis are needed to achieve good results (Güell et al. 2008).

Besides the above general comment, benchmark tests of the different functionals in this case are often hampered by the problem of accurate reference data and by the problem that functionals that seem to behave properly for a metal or system simply fail if the system changes (Ghosh 2006). Focusing on Fe complexes, a benchmark study versus CASPT2 values (see Table 11) indicated that the OPBE functional performs definitely better than commonly used GGA and HGGA functionals. In a similar study, the performance of some DFT functionals to describe CASPT2 results for a series of five Fe complexes indicated that the OLYP functional performed remarkably well in these cases, whereas the success of the HGGA PBE1PBE, B3LYP, and B3LYP* functionals varied from case to case (Pierloot and Vancoillie 2008). Finally, another benchmark study, in which various properties of Fe_2 , Fe_2^- , and FeO^+ , as obtained from a series of GGA, HGGA, MGGA, and HMGGA functionals, were calculated, indicated that no single functional was found to yield a satisfactory description of all characteristics for all states of these species (Sorkin et al. 2008). These results clearly indicate that the spin-state problem still is an open challenge for DFT.

Table 11 Singlet–quintet splitting, $E_{\text{singlet}}-E_{\text{quintet}}$ in kcal/mol, for selected Fe complexes (Swart 2008)

	$\text{Fe}(\text{H}_2\text{O})_6^{2+}$	$\text{Fe}(\text{NH}_3)_6^{2+}$	$\text{Fe}(\text{bpy})_3^{2+}$	MAD
CASPT2	46.6	20.3	-13.2	
BP86	28.4	5.1	-23.2	14.5
RPBE	34.3	6.3	-29.9	14.3
B3LYP	33.1	14.1	0.6	11.2
PBE1PBE	46.0	24.7	9.0	9.1
OPBE	49.3	19.0	-14.9	1.9

MAD mean absolute deviation

Twelve: Excited States

Testing the performance of several functionals in the TD-DFT prediction of 21 valence and 20 Rydberg excitation energies in N_2 , O_2 , HCOOH , and tetracene indicated that valence excitations are easier to predict than Rydberg excitation. Indeed, valence excitations were predicted by 15 functionals with a MUE of 0.36 eV, while Rydberg excitations were predicted with a MUE of 1.13 eV. Focusing on valence excitations, HMGGGA functionals such as TPSSh, B98, and B97-3 are the best performers, with MUE as small as 0.25 eV. Nevertheless, GGA and HGGA functionals, with a MUE of 0.32 and 0.28 eV for the PBE and B3LYP functionals, respectively, also perform reasonably well. Differently, for the accurate prediction of Rydberg excitations, a high amount of Hartree–Fock exchange, as in the M06-2X and M06-HF functionals, with MUE lower than 0.4 eV, is beneficial. With the exception of the also well-performing BMK functional, almost all the other functionals tested results in MUE greater than 0.78 eV, with the PBE and B3LYP functionals resulting in a MUE of 1.95 and 1.07 eV, respectively. When Rydberg and valence excitations are combined into a single database, the best average performance is of the M06-2X and BMK functionals, with a MUE around 0.35 eV (Zhao and Truhlar 2008b).

With regards to charge-transfer excitations, the performance of 16 functionals was tested in the prediction of three charge-transfer excitation energies in tetracene and in the $\text{NH}_3 \cdots \text{F}_2$ and $\text{C}_2\text{H}_4 \cdots \text{C}_2\text{F}_4$ complexes. The average MUE over the 16 functionals examined, 3.86 eV, is depressive. The only working functional is M06-HF, with a MUE of 0.09 eV. However, the inclusion of Hartree–Fock exchange is not the only reason for this surprisingly good performance, since simple HF and the HFLYP functional results in MUE around 1 eV. With the exception of the M05-2X and M06-2X, with a MUE around 2.5 eV, all the other functionals tested resulted in a MUE greater than 3 eV (Zhao and Truhlar 2008b). As a final note, it would be interesting to test the performance of the M06-HF on a larger database.

Moving to more complex organic molecules, the $\pi \rightarrow \pi^*$ transitions of more than 100 dyes from the major classes of chromophores have been investigated using a TD-DFT with the PBE, PBE1PBE, and long-range corrected hybrid functionals. The PBE1PBE and CAMB3LYP were shown to outperform all other approaches, with the latter functional especially adequate to treat molecules with delocalized

excited states. The PBE1PBE functional resulted in a MUE of 22 nm (0.14 eV) with no deviation exceeding 100 nm (0.50 eV), thus delivering reasonable estimates of the color of most organic dyes of practical or industrial interest. Long-range functionals allowed a better description of the low-lying excited-state energies than HGGA functionals, and linearly corrected long-range approaches yield an average error of 10 nm (Jacquemin et al. 2008).

An extended test of the performance of 41 functionals in the calculation of the electronic absorption spectra of Cu and Zn complexes by TD-DFT methods indicated that HGGA functionals outperform GGA functionals. In the case of the spin-unrestricted calculations on the Cu^{II}(thiosemicarbazonato) complex, the functional best performing in the reproduction of the experimental spectra and geometry was the B1LYP, while the B3LYP functional was ranked 8. This order was not replicated in case of the spin-restricted Zn^{II}(thiosemicarbazonato) complex, where the best functional was PBE1PBE, with the B3LYP ranked 10. In both cases, HGGA functionals did not offer an improvement. In almost all the cases, the calculations underestimated the experimental excitation energies (Holland and Green 2010). Nevertheless, it may be worth noting that the OPBE and the OBLYP functionals, which were shown to perform well in other cases, were not considered.

Orbitals in DFT

Since chemists long have used and continue to use orbitals as natural language to explain and rationalize the complex reality of molecules that define the realms of inorganic and organic chemistry, we conclude with a few remarks on orbitals obtained from density functional calculations.

Originally, chemists have built their understanding of orbitals on constructs that resulted from WFT analyses, and such an orbital is usually referred to as molecular orbital (MO). One important aspect of an MO analysis is the investigation of orbital overlap. A simplified wave function theory that emphasizes this particular feature of orbital interaction, the extended Hückel theory (EHT), has revolutionized the general perception of molecular structure and reaction mechanisms.

Since the Kohn–Sham orbitals, introduced and used in DFT, serve a different purpose than creating a reasonable single determinantal wave function Ψ , chemists were seeking answers to the question “what do the Kohn–Sham orbitals and eigenvalues mean?” as DFT moved into the spotlight of electronic structure theory. A simple answer was based on a comparison of orbitals of small molecules (H₂O, N₂, PdCl₄²⁻) obtained from WFT (Hartree–Fock, EHT) and DFT calculations: The shape and symmetry properties of the KS orbitals are very similar to those calculated by HF and EHT methods. The energy order of the occupied orbitals is in most cases in agreement between WFT and DFT methods. Overall, the KS orbitals are a good basis for qualitative interpretation of molecular orbitals (Stowasser and Hoffmann 1999). This simple conception of the meaning and the use of KS orbitals by now has gained general acceptance, and chemists often use KS orbitals in ways that are familiar to them from MO analysis.

Baerends and Gritsenko (1997) have provided an answer to the same question based on fundamental concepts of density functional theory. Among other aspects, the authors identify the following two important characteristics of KS orbitals:

1. The highest occupied Kohn–Sham orbital energy is equal to the exact first ionization energy. This is a property that is very desirable in qualitative MO theory in general and is often simply assumed in such theories.
2. The lowest unoccupied Kohn–Sham orbital energy and all other virtual orbital energies are solutions in exactly the same potential as the occupied orbitals. They are therefore not shifted toward higher energies in the same way as Hartree–Fock virtual orbitals are – HF orbital energy differences are not estimates of excitation energies. Further, it has been observed empirically for a long time that KS orbital energy differences are good approximations to excitation energies, and the KS orbital energy differences play a role as a first approximation to the excitation energy in the treatment of excitation energies using time-dependent DFT.

The authors therefore recommend KS orbitals and one-electron energies as tools in the traditional qualitative MO considerations on which much of the rationalizations of contemporary chemistry are based. The Kohn–Sham one-electron model and KS orbitals provide an ideal MO theoretical context to apply concepts such as “charge control” and “orbital control.” KS orbitals usually follow the expected behavior, in terms of bonding and antibonding character in terms of geometrical distortions and in terms of interaction with other atoms or molecules.

About 10 years later, Cramer and Truhlar (2009) presented a more conservative view of the use of KS orbitals. One should be careful not to stretch the interpretation of KS orbitals beyond its limits, since KS orbitals correspond to a fictitious noninteracting system with the same electron density as the correct many-body function. Since the density computed from KS orbitals is an approximation to the exact density, properties that depend on individual orbitals, with the exception of the energy of the highest occupied orbitals, should be interpreted with care. Nonetheless, many studies published in the literature do employ DFT molecular orbitals to interpret the electronic origins of chemical bonding and reactivity.

The quality of the KS orbitals depends to a large part on the ability of a chosen density functional to correctly represent the ground-state density of a given molecule. In most cases, different density functionals produce qualitatively identical orbitals, which also agree with WFT orbitals. For molecules that might possess a spin-polarized ground-state density, different methods of electronic structure calculation not only produce *quantitatively* different results but also lead to *qualitatively* contrastive conclusions. One such case is illustrated in Fig. 7.

Figure 7 Energies and shapes of the two highest occupied (red-blue) and the two lowest unoccupied (yellow-green) KS orbitals (contour envelope 0.05 a.u.) of $\text{Fe}(\text{S}_2\text{C}_2\text{H}_2)_2$ extracted from an unpolarized ground-state density according to GGA and HGGA DFT calculations. The molecule was constructed according to D_{2h} symmetry with a Fe–S bond length of 220 pm. Also shown for the sake of comparison are orbitals obtained from an extended Hückel calculation.

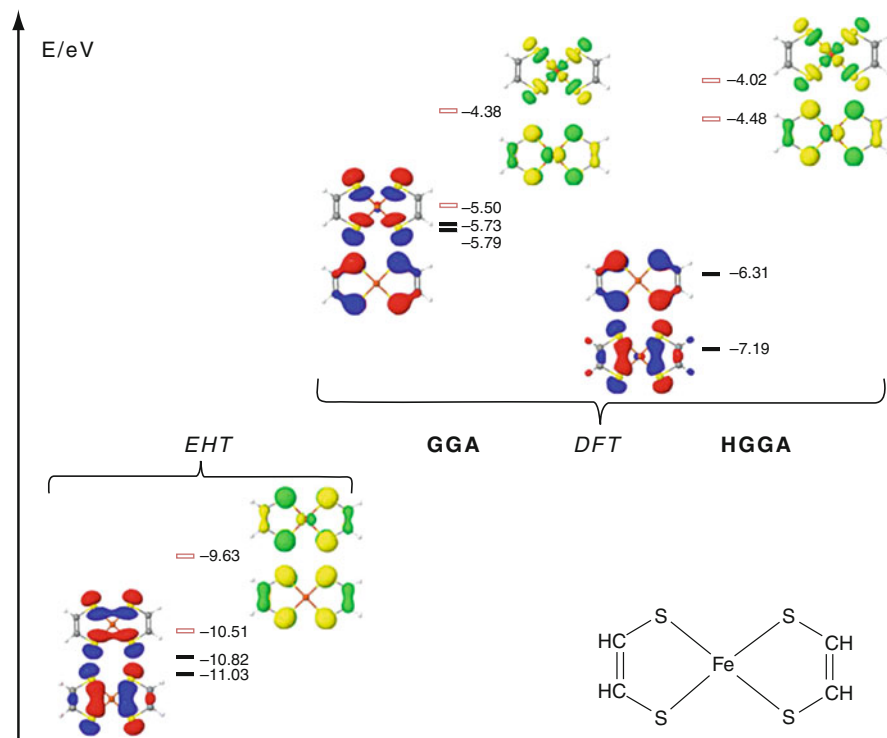


Fig. 7 Energies and shapes of the two highest occupied (*red-blue*) and the two lowest unoccupied (*yellow-green*) KS-orbitals (contour envelope: 0.05 a.u.) of $\text{Fe}(\text{S}_2\text{C}_2\text{H}_2)_2$ extracted from an unpolarized ground state density according to GGA and HGGA DFT calculations. The molecule was constructed according to D_{2h} symmetry with a Fe-S bond length of 220 pm. Also shown for the sake of comparison are orbitals obtained from an extended Hückel calculation

But the essential understanding of chemical orbitals itself is open to interpretation; Autschbach (2012) elaborates on the use and misuse of orbitals in quantum theory and chemistry. Naturally, KS orbitals play an essential role in this challenging controversy. We leave it to the reader to decide whether or not chemically meaningful information can be extracted from the orbital picture as displayed in Fig. 7.

DFTips

The advertent reader has noticed the many technical recommendations that we have given within the discussion of the twelve selected scenarios for properties of atoms and molecules. Here, we will close our instruction manual with a few pieces of general advice.

B3LYP Is No Synonym for DFT

Some researchers hold the opinion that given the fact that the B3LYP functional has been identified as the most successful DFT method in an overwhelming number of systematic investigations in very many areas of chemical research, there is no persuasive motivation to recommend its replacement by one of the other functionals. We do not agree with such a judgment. Although the B3LYP functional is largely responsible for DFT becoming one of the most popular tools in computational chemistry, it does have unsatisfactory performance issues, notably the unreliable results obtained for transition metal chemistry (Zhao and Truhlar 2008a). One should become aware of the capabilities and shortcomings of particular density functionals. It might well be that B3LYP is the proper approach to many chemical problems, but choosing a functional for its previous success while ignoring its potential failures cannot be the right strategy to approach DFT calculations.

Choose Your DFT Method Carefully

We agree with Burke's interpretation of "The good, the bad and the ugly": It is *good* to choose one functional of each kind and stick with it. It is *bad* to run several functionals and pick the "best" answer. And it is *ugly* to design your own functional with 2300 parameters.

In view of the many functionals available, it is likely that one will find a functional that fits one's own particular problems. In view of possibilities offered by computational programs that make it fairly easy to create new sets of parameterized hybrid functionals, it is almost certain that for each chemical problem, the right functional can be designed.

Following such an approach, density functional calculations lose their generality and meaning. One should aim for consistency within one's calculations, rather than for the best agreement with experiment. If one choose a particular functional for its characteristics and capabilities, the results of DFT calculations gain a predictive quality.

Read the Fine Print

During the early days of HGGA development, it became clear that B3LYP is not B3LYP (Hertwig and Koch 1997). Different programs based their implementation of the B3LYP functional on different models of the underlying LDA and produced slightly different results for the same chemical problem. The same considerations apply to the gradient corrections for correlation; P86 requires a different choice of LDA than LYP. Some computer programs automatically make the right choice, while other programs rely on correct input instructions given by the user.

Familiarize yourself with the default values and basic implementations of your favorite DFT code. Default values are chosen with much consideration and, most of the time, are just adequate. However, for subtle problems, you might find it necessary, for example, to change the accuracy of the numerical integration routines.

DFT Does Not Hold the Universal Answer to All Chemical Problems

You can expect that not all of your DFT calculations will produce satisfying results. An increasing number of problem molecules are currently identified for which standard density functionals fail to produce satisfactory results. Seemingly simple stereoelectronic effects in alkane isomers provided a serious challenge for many standard functionals, such as B3LYP, BLYP, or PBE (Grimme 2006c). However, such failures are not bad news for DFT but rather good news. The origin of the dissatisfying DFT performance has been carefully analyzed and led to the design of improved functionals (Zhao et al. 2006). Be honest with your DFT results and accept apparent failures – it might be just another small step climbing Jacob's ladder.

Make DFT an Integral Part of Your Work

Although in some areas of chemical research DFT is at the verge of becoming a standard research tool, necessary for the complete characterization of new molecules, not every chemical problem warrants a full-fledged DFT investigation. However, if you structure your computational approach beginning with an elaboration of qualitative aspects before addressing a problem in a quantitative way, a few quick DFT calculations might provide you with valuable insights how to further pursue your line of work.

Follow Your Interests

As final advice we will leave you with the words of Nobel Prize laureate Harold Kroto: Do something which interests you or which you enjoy, and do it to the absolute best of your ability. If it interests you, however mundane it might seem on the surface, still explore it because something unexpected often turns up just when you least expect it.

We hope that our work could satisfy some old interests as well as perk some new interests. We wish that our short instruction manual serves as a valuable guide for the perplexed and provides some food for thought for the enlightened, be it in agreement or in disagreement. If you follow your interests, keep an open mind, and maintain a broad perspective, good things will happen.

Bibliography

The literature on DFT is large and rich in excellent reviews and overviews. At the same time, while thousands of papers on DFT have been published, most of them will become out of date in the future, as the picture and perception of DFT, chemistry, and science in general is in a state of constant flux. It was necessary to make a conscious selection that of course is not unbiased – some references represent the most recent developments in DFT, while others have been part of our DFT folder for a long time. The reader will find references to books, review articles, and articles that illustrate developments and applications of DFT.

Books on DFT

- Fiolhais, C., Nogueira, F., & Marques, M. (Eds.). (2003). *A primer in density functional theory (lecture notes in physics)*. Berlin/New York: Springer.
- Koch, W., & Holzhäusen, M. C. (2002). *A chemist's guide to density functional theory* (2nd ed.). Weinheim/New York: Wiley.
- Marques, M. A. L., Ullrich, C. A., Nogueira, F., Rubio, A., Burke, K., & Gross, E. K. U. (Eds.). (2006). *Time-dependent density functional theory (lecture notes in physics)*. Berlin: Springer.
- Parr, R. G., & Yang, W. (1989). *Density functional theory of atoms and molecules*. New York: Oxford University Press.

Reviews and Overviews of DFT

- Baerends, E. J., & Gritsenko, O. V. (1997). A quantum chemical view of density functional theory. *Journal of Physical Chemistry A*, *101*, 5383–5403.
- Becke, A. D. (2014). Perspective: Fifty years of density-functional theory in chemical physics. *Journal of Chemical Physics*, *140*, art. 18A301.
- Cohen, A. J., Mori-Sánchez, P., & Yang, W. (2012). Challenges for density functional theory. *Chemical Reviews*, *112*, 289–320.
- Cramer, C. J., & Truhlar, D. G. (2009). Density functional theory for transition metals and transition metal chemistry. *Physical Chemistry Chemical Physics*, *11*, 10757–10816.
- Geerlings, P., De Proft, F., & Langenaeker, W. (2003). Conceptual density functional theory. *Chemical Reviews*, *103*, 1793–1873.
- Kohn, W., Becke, A. D., & Parr, R. G. (1996). Density functional theory of electronic structure. *Journal of Physical Chemistry*, *100*, 12974–12980.
- Neese, F. (2009). Prediction of molecular properties and molecular spectroscopy with density functional theory: From fundamental theory to exchange-coupling. *Coordination Chemistry Reviews*, *253*, 526–563.
- Perdew, J. P., Ruzsinszky, A., Constantin, L. A., Sun, J. W., & Csonka, G. I. (2009). Some fundamental issues in ground-state density functional theory: A guide for the perplexed. *Journal of Chemical Theory and Computation*, *5*, 902–908.
- Sousa, S. F., Fernandes, P. A., & Ramos, M. J. (2007). General performance of density functionals. *Journal of Physical Chemistry A*, *111*, 10439–10452.
- Ziegler, T. (1991). Approximate density functional theory as practical tool in molecular energetics and dynamics. *Chemical Reviews*, *91*, 651–667.

- Ziegler, T. (1995). Density functional theory as practical tool in studies of organometallic energetics and kinetics. Beating the heavy metal blues with DFT. *Canadian Journal of Chemistry*, *73*, 743–761.
- Zhao, Y., & Truhlar, D. G. (2008a). Density functionals with broad applicability in chemistry. *Accounts of Chemical Research*, *41*, 157–167.

Conceptual Developments and Applications of DFT

- Ahlrichs, R., Penco, R., & Scoles, G. (1977). Intermolecular forces in simple systems. *Chemical Physics*, *19*, 119–130.
- Baerends, E. J., Ellis, D. E., & Ros, P. (1973). Self-consistent molecular Hartree-Fock-Slater calculations – I. The computational procedure. *Chemical Physics*, *2*, 41–47.
- Baerends, E. J., & Ros, P. (1978). Evaluation of the LCAO Hartree-Fock-Slater method – Applications to transition-metal complexes. *International Journal of Quantum Chemistry*, *12*, 169–190.
- Bartlett, R. J., Lotrich, V. F., & Schweigert, I. V. (2005). Ab initio density functional theory: The best of both worlds? *Journal of Chemical Physics*, *123*, art. 062205.
- Becke, A. D. (1988a). Density-functional exchange-energy Approximation with correct asymptotic behavior. *Physical Review A*, *38*, 3098–3100.
- Becke, A. D. (1988b). A multicenter numerical-integration scheme for polyatomic molecules. *Journal of Chemical Physics*, *88*, 2547–2553.
- Becke, A. D., & Roussel, M. R. (1989). Exchange holes in inhomogeneous systems – A coordinate-space model. *Physical Review A*, *39*, 3761–3767.
- Becke, A. D. (1993a). A new mixing of Hartree-Fock and local density-functional theories. *Journal of Chemical Physics*, *98*, 1372–1377.
- Becke, A. D. (1993b). Density-functional thermochemistry: 3. The role of exact exchange. *Journal of Chemical Physics*, *98*, 5648–5652.
- Boerrigter, P. M., te Velde, G., & Baerends, E. J. (1988). 3-dimensional numerical-integration for electronic-structure calculations. *International Journal of Quantum Chemistry*, *33*, 87–113.
- Furche, F. (2008). Developing the random phase approximation into a practical post-Kohn–Sham correlation model. *Journal of Chemical Physics*, *129*, art. 114105.
- Gill, P. M. W. (2001). Obituary: Density functional theory (1927–1993). *Australian Journal of Chemistry*, *54*, 661–662.
- Grimme, S. (2006a). Semiempirical hybrid density functional with perturbative second-order correlation. *Journal of Chemical Physics*, *124*, art. 034108.
- Hepburn, J., Scoles, G., & Penco, R. (1975). Simple but reliable method for prediction of intermolecular potentials. *Chemical Physics Letters*, *36*, 451–456.
- Hertwig, R. H., & Koch, W. (1997). On the parameterization of the local correlation functional. What is Becke-3-LYP? *Chemical Physics Letters*, *268*, 345–351.
- Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. *Physics Reviews*, *136*, B646–B871.
- Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Physics Reviews*, *140*, A1133–A1138.
- Kurth, S., & Perdew, J. P. (2000). Role of the exchange-correlation energy: Nature’s glue. *International Journal of Quantum Chemistry*, *77*, 814–818.
- Leininger, T., Stoll, H., Werner, H. J., & Savin, A. (1997). Combining long-range configuration interaction with short-range density functionals. *Chemical Physics Letters*, *275*, 151–160.
- Perdew, J. P. (1986). Density-functional approximation for the correlation-energy of the inhomogeneous electron gas. *Physical Review B*, *33*, 8822–8824.
- Perdew, J. P., Burke, K., & Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Physical Review Letters*, *77*, 3865–3868.

- Perdew, J. P., & Schmidt, K. (2001). Jacob's ladder of density functional approximations for the exchange-correlation energy. *Density Functional Theory and Its Applications to Materials*, 577, 1–20.
- Peeverati, R., & Truhlar, D. G. (2014). The quest for a universal density functional: The accuracy of density functionals across a broad spectrum of databases in chemistry and physics. *Philosophical Transactions of the Royal Society A*, 372, art. 20120476.
- Slater, J. C. (1951). A simplification of the Hartree-Fock method. *Physics Reviews*, 81, 385–390.
- Tao, J. M., Perdew, J. P., Staroverov, V. N., & Scuseria, G. E. (2003). Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids, *Physical Review Letters*, 91, art. 146401.
- Tsuneda, T., & Hirao, K. (2014). Long-range correction for density functional theory. *WIREs Computational Molecular Science*, 4, 375–390.
- Versluis, L., & Ziegler, T. (1988). The determination of molecular structures by density functional theory: The evaluation of analytical energy gradients by numerical integration. *Journal of Chemical Physics*, 88, 322–328.
- Vosko, S. H., Wilk, L., & Nusair, M. (1980). Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics*, 58, 1200–1211.
- Zope, R. R., & Dunlap, B. I. (2006). The limitations of Slater's element-dependent exchange functional from analytic density-functional theory, *Journal of Chemical Physics*, 124, art. 044107.

Practical Developments and Applications of DFT

- Autschbach, J. (2012). Orbitals: Some fiction and some facts. *Journal of Chemical Education*, 89, 1032–1040.
- Barden, C. J., Rienstra-Kiracofe, J. C., & Schaefer, H. F. (2000). Homonuclear 3d transition-metal diatomics: A systematic density functional theory study. *Journal of Chemical Physics*, 113, 690–700.
- Curtiss, L. A., Raghavachari, K., Redfern, P. C., & Pople, J. A. (1997). Assessment of Gaussian-2 and density functional theories for the computation of enthalpies of formation. *Journal of Chemical Physics*, 106, 1063–1079.
- Curtiss, L. A., Raghavachari, K., Redfern, P. C., & Pople, J. A. (2000). Assessment of Gaussian-3 and density functional theories for a larger experimental test set. *Journal of Chemical Physics*, 112, 7374–7383.
- Ghosh, A. (2006). Transition metal spin state energetics and noninnocent systems: Challenges for DFT in the bioinorganic arena. *Journal of Biological Inorganic Chemistry*, 11, 712–714.
- Grimme, S. (2006b). Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry*, 27, 1787–1799.
- Grimme, S. (2006c). Seemingly simple stereoelectronic effects in alkane isomers and the implications for Kohn-Sham density functional theory. *Angewandte Chemie International Edition*, 45, 4460–4464.
- Güell, M., Luis, J. M., Solà, M., & Swart, M. (2008). Importance of the basis set for the spin-state energetics of iron complexes. *Journal of Physical Chemistry A*, 112, 6384–6391.
- Holland, J. P., & Green, J. C. (2010). Evaluation of exchange-correlation functionals for time-dependent density functional theory calculations on metal complexes. *Journal of Computational Chemistry*, 31, 1008–1014.
- Jacquemin, D., Perpète, E. A., Scuseria, G. E., Ciofini, I., & Adamo, C. (2008). TD-DFT performance for the visible absorption spectra of organic dyes: Conventional versus long-range hybrids. *Journal of Chemical Theory and Computation*, 4, 123–135.
- Kelly, R. A., Clavier, H., Giudice, S., Scott, N. M., Stevens, E. D., Bordner, J., Samardjiev, I., Hoff, C. D., Cavallo, L., & Nolan, S. P. (2008). Determination of N-heterocyclic carbene (NHC)

- steric and electronic parameters using the [(NHC)Ir(CO)(2)Cl] system. *Organometallics*, *27*, 202–210.
- Korth, M., & Grimme, S. (2009). “Mindless” DFT Benchmarking. *Journal of Chemical Theory and Computation*, *5*, 993–1003.
- Lynch, B. J., & Truhlar, D. G. (2003). Small representative benchmarks for thermochemical calculations. *Journal of Physical Chemistry A*, *107*, 8996–8999.
- Pierloot, K., & Vancoillie, S. J. (2008). Relative energy of the high-(T-5(2g)) and low-((1)A(1g)) spin states of the ferrous complexes [Fe(L)(NHS4)]: CASPT2 versus density functional theory. *Journal of Chemical Physics*, *128*, art. 034104.
- Reiher, M., Salomon, O., & Hess, B. A. (2001). Reparameterization of hybrid functionals based on energy differences of states of different multiplicity. *Theoretical Chemistry Accounts*, *107*, 48–55.
- Riley, K. E., Op’t Holt, B. T., & Merz, K. M. (2007). Critical assessment of the performance of density functional methods for several atomic and molecular properties. *Journal of Chemical Theory and Computation*, *3*, 407–433.
- Schultz, N. E., Zhao, Y., & Truhlar, D. G. (2005a). Density functionals for inorganometallic and organometallic chemistry. *Journal of Physical Chemistry A*, *109*, 11127–11143.
- Schultz, N. E., Zhao, Y., & Truhlar, D. G. (2005b). Databases for transition element bonding: Metal-metal bond energies and bond lengths and their use to test hybrid, hybrid meta, and meta density functionals and generalized gradient approximations. *Journal of Physical Chemistry A*, *109*, 4388–4403.
- Sorkin, A., Iron, M. A., & Truhlar, D. G. (2008). Density functional theory in transition-metal chemistry: Relative energies of low-lying states of iron compounds and the effect of spatial symmetry breaking. *Journal of Chemical Theory and Computation*, *4*, 307–315.
- Sponer, J., Jurecka, P., & Hobza, P. (2004). Accurate interaction energies of hydrogen-bonded nucleic acid base pairs. *Journal of the American Chemical Society*, *126*, 10142–10151.
- Stowasser, R., & Hoffmann, R. (1999). What do the Kohn-Sham orbitals and eigenvalues mean? *Journal of the American Chemical Society*, *121*, 3414–3420.
- Swart, M. (2008). Accurate spin-state energies for iron complexes. *Journal of Chemical Theory and Computation*, *4*, 2057–2066.
- Wang, N. X., & Wilson, A. K. (2004). The behavior of density functionals with respect to basis set. I. The correlation consistent basis sets. *Journal of Chemical Physics*, *121*, 7632–7646.
- Wodrich, M. D., Corminboeuf, C., & Schleyer, P. v. R. (2006). Systematic errors in computed alkane energies using B3LYP and other popular DFT functionals. *Organic Letters*, *8*, 3631–3634.
- Yang, W. (1991). Direct calculation of electron density in density functional theory. *Physical Review Letters*, *66*, 1438–1441.
- Zhao, Y., Pu, J., Lynch, B. J., & Truhlar, D. G. (2004). Tests of second-generation and third-generation density functionals for thermochemical kinetics. *Physical Chemistry Chemical Physics*, *6*, 673–676.
- Zhao, Y., & Truhlar, D. G. (2004). Hybrid meta density functional theory methods for thermochemistry, thermochemical kinetics, and noncovalent interactions: The MPW1B95 and MPWB1K models and comparative assessments for hydrogen bonding and van der Waals interactions. *Journal of Physical Chemistry A*, *108*, 6908–6918.
- Zhao, Y., & Truhlar, D. G. (2005a). Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *Journal of Physical Chemistry A*, *109*, 5656–5667.
- Zhao, Y., & Truhlar, D. G. (2005b). Benchmark databases for nonbonded interactions and their use to test density functional theory. *Journal of Chemical Theory and Computation*, *1*, 415–432.
- Zhao, Y., Gonzalez-Garcia, N., & Truhlar, D. G. (2005). Benchmark database of barrier heights for heavy atom transfer, nucleophilic substitution, association, and unimolecular reactions and its use to test theoretical methods. *Journal of Physical Chemistry A*, *109*, 2012–2018.

- Zhao, Y., & Truhlar, D. G. (2005c). How well can new-generation density functional methods describe stacking interactions in biological systems? *Physical Chemistry Chemical Physics*, *7*, 2701–2705.
- Zhao, Y., Schultz, N. E., & Truhlar, D. G. (2006). Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *Journal of Chemical Theory and Computation*, *2*, 364–382.
- Zhao, Y., & Truhlar, D. G. (2008b). The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theoretical Chemistry Accounts*, *120*, 215–241.
- Zhou, M., Andrews, L., & Bauschlicher, C. W. (2001). Spectroscopic and theoretical investigations of vibrational frequencies in binary unsaturated transition-metal carbonyl cations, neutrals, and anions. *Chemical Reviews*, *101*, 1931–1961.

Thomas Bondo Pedersen

Contents

Introduction	270
Response Theory	272
The Linear Response Function	275
Quadratic and Higher-Order Response Functions	281
Static Response Functions	283
Gauge and Origin Invariance	285
Effects of Nuclear Motion	288
Further Reading	292
Bibliography	293

Abstract

This chapter provides a concise introduction to quantum chemical response theory as implemented in a number of widely used electronic structure software packages. While avoiding technical derivations of response functions, the fundamental idea of response theory, namely, the calculation of field-induced molecular properties through changes in expectation values, is explained in a manner equally valid for approximate wave function and density functional theories. Contrasting response theory to textbook treatments of perturbation theory, key computational concepts such as iterative solution of response equations and the identification and calculation of electronic excitation energies are elucidated. The wealth of information that can be extracted from approximate linear, quadratic, and higher-order response functions is discussed on the basis of the corresponding exact response functions. Static response functions and

T.B. Pedersen (✉)

Department of Chemistry, Centre for Theoretical and Computational Chemistry, University of Oslo, Oslo, Norway

e-mail: thomas.b.pedersen@gmail.com

their identification and numerical calculation as energy derivatives are discussed separately. Practical issues related to the lack of gauge and origin invariance in approximate calculations are discussed without going into too much theoretical detail regarding the sources of these problems. Finally, the effects of nuclear motion (molecular vibrations, in particular) and how to include them in computational studies are treated in some detail.

Introduction

The ultimate vision of computational chemistry is to provide a virtual laboratory for chemical explorations. The goal is not to replace but rather complement the real lab. In the virtual lab, it is easy to change “experimental” conditions and to experiment with hypothetical what-if scenarios. Virtual synthesis of chemical compounds can be done with a few mouse clicks, whereas the virtual measurement of physical and chemical properties is the often burdensome task of a computational engine based on the laws of classical and/or quantum physics. Most research efforts, therefore, are aimed at improving the algorithms and approximations implemented in the computational engines. The main role of quantum chemical response theory, which is the subject of this chapter, is the virtual measurement of optical properties of molecules.

The scope of this chapter is to provide a rudimentary understanding of response theory as implemented in a number of molecular electronic structure packages based on wave function models or density functional theory. Only the general structure of response theory and its computer implementation are discussed, leaving out the often complicated details of advanced wave function and density functional models. For these details the reader is referred to the literature mentioned in the last section and references therein. Although the discussion of this chapter is restricted to nonrelativistic theory, it is the same line of reasoning that is applied in relativistic response theory. In conjunction with the chapter on applications of response theory, the reader should become sufficiently familiar with the concepts and practices of response theory to allow educated use of quantum chemistry software packages.

As the name indicates, response theory describes the response of a molecular system to external potentials such as electromagnetic fields. The meaning of *external* depends on the definition of the molecular system. For example, the magnetic field due to nuclear spins would typically be considered an external potential. The fundamental assumption is that the external potentials are weak compared to the internal potentials of the molecule. That is, the external potentials can be regarded as perturbations of the isolated molecule. This assumption can be justified by noting that the electric field strength in the hydrogen atom is on the order of 10^{11} Vm^{-1} , which is well above the external field strengths applied in almost all experimental setups. Experimental observations are then rationalized by means of response functions, which may be expressed entirely in terms of states and energies of the unperturbed (isolated) molecule. Response functions are thus characteristic properties of a molecule. The objective of quantum chemical response theory is the computation of response functions from first principles.

Response functions quantify the field-induced change in a given observable such as the electric dipole moment. It is therefore necessary to know the state of the molecule before the field is applied. This unperturbed state is normally taken to be the molecular ground state, which is usually assumed to be nondegenerate. Calculation of the molecular ground state is, however, a daunting task and approximations are invoked in practice. We will use the clamped-nucleus Born–Oppenheimer approximation to separate the electronic degrees of freedom from the nuclear ones. Focusing on the response of the electronic ground state, the coupling between the nuclei and the external field is neglected. This is justified for field frequencies much larger than vibrational and rotational frequencies.

The electronic ground state $|0\rangle$ is computed as an approximate solution to the time-independent Schrödinger equation, that is,

$$H_0 |0\rangle \approx E_0 |0\rangle \quad (1)$$

where H_0 is the electronic Hamiltonian (see below) and E_0 is the ground state energy in the absence of external fields. The approximate solution (energy and wave function or density) can be obtained using standard methods such as density functional theory (DFT), Hartree–Fock (HF) theory, multiconfigurational self-consistent field (MCSCF) theory, configuration interaction (CI), or coupled cluster (CC) theory. For fixed nuclear positions (clamped-nucleus approximation), the electronic Hamiltonian is given by

$$H_0 = \frac{1}{2m_e} \sum_{s=1}^{N_e} (\mathbf{p}^s)^2 - \sum_{s=1}^{N_e} \sum_{K=1}^{N_n} \frac{e^2 Z_K}{4\pi\epsilon_0 |\mathbf{r}^s - \mathbf{R}^K|} + \sum_{K=2}^{N_n} \sum_{L=1}^{K-1} \frac{e^2 Z_K Z_L}{4\pi\epsilon_0 |\mathbf{R}^K - \mathbf{R}^L|} + \sum_{s=2}^{N_e} \sum_{t=1}^{s-1} \frac{e^2}{4\pi\epsilon_0 |\mathbf{r}^s - \mathbf{r}^t|} \quad (2)$$

Here, N_e is the number of electrons, N_n is the number of nuclei, m_e is the mass of the electron, e is the elementary charge, ϵ_0 is the vacuum permittivity, \mathbf{r}^s and \mathbf{p}^s are the position and momentum operators of electron s , and \mathbf{R}^K and Z_K are the position and atomic number of nucleus K . Starting from the Schrödinger equation (Eq. 1), relativistic effects as described by the Breit–Pauli Hamiltonian can be treated as perturbations on an equal footing with external fields. Effects of nuclear motion (vibrations and rotations) can be estimated once the electronic response functions have been calculated.

Response functions are functions of the frequencies of the external fields, and although response theory is based on time-dependent perturbation theory, it is more appropriate to describe response theory as frequency dependent rather than time dependent. Nevertheless, response theory for approximate quantum chemical models is sometimes referred to as time dependent, for example, time-dependent DFT (TDDFT) and time-dependent HF (TDHF) theory. Related by Fourier transformation, the time and frequency domains are two sides of the same coin. In practice, however, response functions are computed for a limited number of perturbation frequencies without any explicit reference to the time evolution of the quantum state.

Response Theory

Approximate wave function and density functional theories provide information about the electronic structure of molecules in their electronic ground state. The information includes the electronic charge density, total energy, electric multipole moments (dipole, quadrupole, octupole, etc.), forces on the nuclei, and vibrational frequencies, which is sufficient to model a wide range of chemical phenomena. For example, equilibrium structures and transition states can be calculated from the forces, and vibrational frequencies are not only useful for the interpretation of vibrational spectra but also enable the calculation of thermochemical data from first principles. These theories are sufficient to model experimental conditions where only the electronic ground state is significantly populated.

When a molecule is subjected to external fields, however, quantum mechanics tells us that the electronic ground state responds by becoming a superposition of many electronic states. Response theory takes this wave function change into account and thus facilitates calculations of molecular properties beyond the reach of the ground state theory. Computation of the ground state response allows extraction of information about transitions from the ground state to excited states induced by one or more photons and even allows us to calculate properties of the excited states without explicitly computing the excited state wave functions.

Response theory describes the change in observable quantities such as electric and magnetic multipole moments due to external fields. The starting point of response theory therefore is the time evolution of the expectation value $\langle A \rangle$ of a Hermitian operator A representing the observable quantity (e.g., the electric or magnetic dipole moment). In order to monitor the change in an observable quantity, we need to know the state of the molecule before the external field is switched on. As mentioned above, the electronic ground state $|0\rangle$ is the proper choice under most experimental conditions. With this initial condition, the time evolution can be written as the perturbation expansion

$$\begin{aligned}
 \langle A \rangle &= \langle 0 | A | 0 \rangle \\
 &+ \int_{-\infty}^{\infty} \langle \langle A; V(\omega) \rangle \rangle_{\omega} \exp(-i\omega t) d\omega \\
 &+ \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \langle A; V(\omega), V(\omega') \rangle \rangle_{\omega, \omega'} \exp(-i(\omega + \omega')t) d\omega d\omega' \\
 &+ \dots
 \end{aligned} \tag{3}$$

where the operator $V(\omega)$ describes the interaction between the electrons and the external field of frequency ω . It should be noted that the perturbation expansion cannot generally be guaranteed to converge. Moreover, the perturbation expansion does not make sense when the external potential is comparable to or larger than the internal potential of the molecule. In most cases of practical interest, including the interaction of molecular electrons with laser fields, the perturbation expansion may be used.

Frequency-dependent response functions can only be computed within approximate electronic structure models that allow definition of the time-dependent expectation value. Hence, frequency-dependent response functions are not defined for approximate methods that provide an energy but no wave function. Such methods include Møller–Plesset (MP) perturbation theory, multiconfigurational second-order perturbation theory (CASPT2), and coupled cluster singles and doubles with noniterative perturbative triples [CCSD(T)]. As we shall see later, it is possible to derive static response functions for such methods.

Common methods that do allow calculation of frequency-dependent response functions include DFT, HF, MCSCF, and members of the CC hierarchy of wave functions (CCSD(T) is not a member of this hierarchy, as it does not provide a wave function). In these models, the wave function (or density in the case of DFT) is written in terms of time-dependent parameters which are determined from the time-dependent Schrödinger equation with the initial condition that the approximate ground state wave function (or density) is recovered in the limit ($t \rightarrow -\infty$), that is, before the external fields are switched on. The zeroth-order parameters are thus determined in the same way as in the ground state model, and response theory calculations are naturally performed on top of a standard ground state calculation. First-, second-, and higher-order parameters are determined from response equations that take into account the coupling with the external field in a manner consistent with the time-dependent Schrödinger equation. The response equations have the general matrix form (the detailed form of these matrices and vectors depends on the approximate theory used)

$$[\mathcal{H} - \omega\mathcal{S}] \lambda^{(n)}(\omega) = -\mathcal{V}^{(n)}(\omega) \quad (4)$$

where $\lambda^{(n)}(\omega)$, the unknown quantity in this equation, is a vector containing the n th-order ($n > 0$) wave function parameters in a suitable order. While the matrices \mathcal{H} and \mathcal{S} only depend on the unperturbed ground state wave function, the vector on the right-hand side additionally depends on the lower-order wave function parameters ($\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)}, \dots, \lambda^{(n-1)}$) as well as on one or more perturbation operators $V(\omega)$. This implies that the solution of the first-order equations must be known before the second-order equations can be solved, and so on. The frequency ω represents the sum of the frequencies corresponding to the perturbations included in the vector on the right-hand side of Eq. 4.

Solving the response equations (Eq. 4) is formally simple, namely,

$$\lambda^{(n)}(\omega) = -[\mathcal{H} - \omega\mathcal{S}]^{-1} \mathcal{V}^{(n)}(\omega) \quad (5)$$

where we assume that the matrix $[\mathcal{H} - \omega\mathcal{S}]$ is nonsingular. Those frequencies ω for which the matrix is singular have a special significance, as will be discussed below. Since the matrices \mathcal{H} and \mathcal{S} only depend on the ground state wave function or density, they may in principle be computed once and for all. For each frequency and right-hand side vector we may then calculate the inverse $[\mathcal{H} - \omega\mathcal{S}]^{-1}$ followed by a simple matrix–vector multiplication to obtain the response parameters according to

Eq. 5. Unfortunately, this simple solution strategy is too computationally demanding to be useful in practice. The large number of matrix elements in \mathcal{H} and \mathcal{S} makes it expensive to compute the matrices and to compute the inverse which, in addition, needs to be done for each frequency ω . The number of matrix elements is given by the square of the number of ground state wave function parameters. For DFT or HF, for example, the number of ground state parameters is given by OV where O and V are the number of occupied and virtual (Kohn–Sham or HF) orbitals, respectively. For the more complicated coupled cluster singles and doubles (CCSD) wave function, the number of ground state wave function parameters is roughly $O^2V^2/2$. To get an impression of the orders of magnitude, consider a simple benzene molecule with the aug-cc-pVDZ basis set. In this case, there are approximately 3,600 DFT or HF ground state parameters and over six million CCSD parameters. The number of elements in the \mathcal{H} and \mathcal{S} matrices would therefore be approximately 13 million for DFT and HF and 36×10^{12} for CCSD. For two benzene molecules, these numbers would increase to approximately 207 million matrix elements for DFT and HF and 10^{16} for CCSD. Just storing the matrices in computer memory would quickly pose an insurmountable problem, and a different solution strategy is needed.

Rather than straightforward application of Eq. 5, the response equations (Eq. 4) are solved in an iterative procedure in which the matrices and vectors are projected onto a small subspace. In this procedure, an initial guess (a trial vector) is generated according to Eq. 5 by neglecting all off-diagonal elements of \mathcal{H} and \mathcal{S} such that the inverse matrix $[\mathcal{H} - \omega\mathcal{S}]^{-1}$ is readily computed. The iterations proceed by refining the trial vector until the norm of the residual vector

$$\mathbf{R} = (\mathcal{H} - \omega\mathcal{S})\lambda^{(n)} + \mathcal{V}^{(n)} \quad (6)$$

falls below a specified tolerance. Although quantum chemistry software packages provide reasonable default values for the tolerance, it may occasionally be necessary to lower the value to obtain sufficient accuracy in the computed response functions.

From the perturbative corrections to the wave function parameters we obtain perturbative corrections to the expectation value of the operator A , that is,

$$\langle A \rangle = \langle A \rangle^{(0)} + \langle A \rangle^{(1)} + \langle A \rangle^{(2)} + \dots \quad (7)$$

We may then identify response functions for each of the approximate electronic structure models by comparing with Eq. 3:

$$\langle A \rangle^{(0)} = \langle 0 | A | 0 \rangle \quad (8)$$

$$\langle A \rangle^{(1)} = \int_{-\infty}^{\infty} \langle \langle A; V(\omega) \rangle \rangle_{\omega} \exp(-i\omega t) d\omega \quad (9)$$

$$\begin{aligned} \langle A \rangle^{(2)} &= \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle\langle A; V(\omega), V(\omega') \rangle\rangle_{\omega, \omega'} \exp(-i(\omega + \omega')t) d\omega d\omega' \\ &\vdots \end{aligned} \quad (10)$$

Response theory deviates in one crucial aspect from the formulation of time-dependent perturbation theory in most textbooks on quantum mechanics: the response parameters $\lambda^{(n)}$ are not explicitly expressed in terms of the excited states. As a consequence, knowledge of the excited state wave functions is not needed in response theory. Instead, we must solve the response equations (Eq. 4) for each set of perturbation operators $V(\omega)$. This is a tremendous computational advantage as there are significantly fewer perturbation operators, and hence response equations to solve, than excited states in virtually all cases of practical interest.

In order to illuminate the wealth of information that can be extracted from response functions for approximate wave functions, it is instructive to study the details of the exact response functions expressed in terms of electronic excited states.

The Linear Response Function

The linear response function describing the first-order induced electric dipole moment due to an oscillating and spatially uniform electric field is related to the frequency-dependent electric dipole polarizability as

$$\alpha_{ij}(-\omega; \omega) = -\langle\langle \mu_i; \mu_j \rangle\rangle_{\omega} \quad (11)$$

The i th Cartesian component of the electronic electric dipole operator is given by

$$\mu_i = -e \sum_s r_i^s \quad (12)$$

where the sum is over all electrons in the molecule. This identification is obtained by substituting $A \rightarrow \mu_i$ and $V(\omega) \rightarrow -\mu \cdot \mathbf{F}(\omega)$ in Eq. 3. The uniform electric field amplitude vector at frequency ω is here represented by $\mathbf{F}(\omega)$. In an approximate calculation, the electric dipole polarizability is thus obtained by computing (minus) the linear response function from first-order wave function parameters determined by Eq. 4 with the three Cartesian components of the electric dipole operator as perturbation operators. Note that the response equations must be solved for each frequency.

More information can be extracted from the linear response function, however. This is most easily seen by studying the exact linear response function. Unlike the approximate theories, the exact linear response function is expressed in terms of the ground and excited states satisfying the time-independent Schrödinger equation

$$H_0 |k\rangle = E_k |k\rangle \quad (13)$$

where $k \geq 0$ and H_0 is the molecular electronic Hamiltonian in the clamped-nucleus Born–Oppenheimer approximation, Eq. 2. The exact linear response function is given by

$$\langle\langle A; B \rangle\rangle_\omega = \frac{1}{\hbar} \sum_{k \neq 0} \left(\frac{\langle 0 | A | k \rangle \langle k | B | 0 \rangle}{\omega - \omega_{k0}} - \frac{\langle 0 | B | k \rangle \langle k | A | 0 \rangle}{\omega + \omega_{k0}} \right) \quad (14)$$

where $\hbar\omega_{k0} = E_k - E_0$ is the excitation energy for the transition from the ground state to the k th excited state. This formulation of the linear response function is often referred to as the sum-over-states expression or spectral resolution.

The exact linear response function is singular at the molecular excitation energies, that is, it has poles at $\omega = \pm\omega_{k0}$. This is exploited in approximate theories to identify excitation energies as those frequencies for which the approximate linear response function is singular. In principle, the search for excitation energies could be done by scanning over frequencies in a manner analogous to measurements of absorption spectra. In practice, however, it is straightforward to identify the poles of the approximate linear response function as the frequencies for which the response equations (Eq. 4) are singular. As discussed above, this occurs at frequencies where the matrix $\mathcal{H} - \omega\mathcal{S}$ cannot be inverted. This, in turn, occurs when ω equals one of the eigenvalues w_n of the generalized eigenvalue problem

$$(\mathcal{H} - w_n\mathcal{S}) \mathbf{V}_n = 0 \quad (15)$$

where \mathbf{V}_n is the eigenvector corresponding to the eigenvalue w_n . The generalized eigenvalue problem does not depend on the perturbation, and it can therefore be solved to directly obtain the excitation energies, identifying $\omega_{k0} = w_k, k = 1, 2, 3, \dots$, without explicit calculation of neither the excited state wave functions nor their energies. The eigenvectors \mathbf{V}_n are the closest we get to an explicit wave function representation of the excited states in response theory. In practice, therefore, analysis of these eigenvectors is used to extract the (approximate) orbital nature of an excitation, for example, as a $\pi - \pi^*$ orbital transition. Since we know the ground state energy E_0 from the prerequisite unperturbed calculation, the excited state energy may be deduced from the excitation energy as

$$E_k = E_0 + \hbar\omega_{k0} \quad (16)$$

Solving the eigenvalue problem, Eq. 15, does not necessarily yield all excited states. With a spin-adapted wave function, only excited states with the proper spin symmetry would be produced. For example, only spin-singlet excited states can be found with a spin-singlet wave function parameterization. Triplet states require a suitable spin-triplet adaptation of the wave function response. In unrestricted

theories, the ground and excited states need not be spin eigenfunctions and the labeling of states as spin-singlet, spin-triplet, etc., is not well defined.

It is important to realize, however, that the number of eigenvalues w_n equals the number of excited states within a given approximate theory. Calculation of all eigenvalues is very computationally demanding. The number of excited states in an approximate theory equals the number of ground state wave function parameters, which grows quickly with the size of the molecule, as discussed above. One would most often be interested in just a few, typically on the order of 10, of the lowest excitation energies. Iterative techniques have been devised that allow the solution of Eq. 15 for a given number of eigenvalues and eigenvectors in ascending order. In order to calculate the excitation energy ω_{k0} , it is therefore necessary to compute all excitation energies below this one as well, that is, $\omega_{10}, \omega_{20}, \omega_{30}, \dots, \omega_{k-1,0}$.¹ The iterative solution of the generalized eigenvalue problem Eq. 15 is carried out by refining trial vectors and eigenvalues until the norm of the residuals

$$\mathbf{R}_n = (\mathcal{H} - w_n \mathcal{S}) \mathbf{V}_n \quad (17)$$

falls below a given tolerance.

Another important piece of information can be extracted from the pole structure of the linear response function. The residue of the linear response function at an excitation energy provides the one-photon transition strength:

$$\langle 0 | A | k \rangle \langle k | B | 0 \rangle = \lim_{\omega \rightarrow \omega_{k0}} \hbar (\omega - \omega_{k0}) \langle \langle A; B \rangle \rangle_{\omega} \quad (18)$$

This observation allows us to identify transition moments between the ground state and the k th excited state, $\langle 0 | A | k \rangle$, in approximate theories by computing the residue of the linear response function at the k th excitation energy. Again, this is achieved without explicitly computing the excited state wave function. In practice, the transition moment is expressed in terms of the k th eigenvector of Eq. 15 and the matrix representation of the operator A in the molecular orbital (or spin orbital) basis. It is therefore possible, in principle, to compute the linear response function for an approximate wave function using the sum-over-states expression with excitation energies calculated according to Eq. 15 and transition moments from Eq. 18. If all excitation energies and residues are included, the result is identical to the response function computed from the response equations (Eq. 4). From a computational point of view, of course, this approach is not advisable unless the combined number of perturbation operators and frequencies is larger than the number of excited states. This is practically never the case.

¹Although techniques exist which solve an eigenvalue equation around a specific energy, these techniques are not used for quantum chemical calculations of excitation energies, as the energy range is rarely known in advance.

Within approximate wave function or density functional theories, it thus becomes possible to calculate one-photon absorption intensities such as the electric dipole oscillator strength of a transition from the ground state to the k th excited state from the following residue of the frequency-dependent electric dipole polarizability:

$$\begin{aligned} f_{k0} &= \frac{2m_e}{3e^2\hbar} \omega_{k0} \langle 0 | \mu | k \rangle \cdot \langle k | \mu | 0 \rangle \\ &= \frac{2m_e}{3e^2\hbar} \omega_{k0} \sum_i \lim_{\omega \rightarrow \omega_{k0}} \hbar (\omega - \omega_{k0}) \alpha_{ii}(-\omega; \omega) \end{aligned} \quad (19)$$

where m_e is the mass of the electron.

In approximate theories, the iterative solution of the response equations (Eq. 4) becomes difficult and may even fail to converge when the frequency is too close to one of the excitation energies. As a consequence, linear response functions should only be calculated in transparent spectral regions. In absorptive spectral regions (“close” to an excitation energy), so-called anomalous dispersion is observed experimentally. Anomalous dispersion can be understood theoretically by taking into account the finite lifetime of the excited electronic states, leading to damped response theory. The exact damped linear response function is given by

$$\langle\langle A; B \rangle\rangle_\omega = \frac{1}{\hbar} \sum_{k \neq 0} \left(\frac{\langle 0 | A | k \rangle \langle k | B | 0 \rangle}{\omega - \omega_{k0} + i\Gamma_k/2} - \frac{\langle 0 | B | k \rangle \langle k | A | 0 \rangle}{\omega + \omega_{k0} - i\Gamma_k/2} \right) \quad (20)$$

The broadening Γ_k is proportional to the probability of the excited state $|k\rangle$ decaying into any of the other states, and it is related to the lifetime of the excited state as $\tau_k = 1/\Gamma_k$. For $\Gamma_k = 0$, the lifetime is infinite and Eq. 14 is recovered from Eq. 20. Unfortunately, it is not possible to account for the finite lifetime of each individual excited state in approximate theories based on the response equations (Eq. 4). We would be forced to use a sum-over-states expression, which is computationally intractable. Moreover, the lifetimes cannot be adequately determined within a semiclassical radiation theory as employed here and a fully quantized description of the electromagnetic field is required. In addition, all decay mechanisms would have to be taken into account, for example, radiative decay, thermal excitations, and collision-induced transitions. Damped response theory for approximate electronic wave functions is therefore based on two simplifying assumptions: (1) all broadening parameters are assumed to be identical, $\Gamma_1 = \Gamma_2 = \dots = \Gamma$, and (2) the value of Γ is treated as an empirical parameter. With a single empirical broadening parameter, the response equations take the same form as in Eq. 4 with the substitution $\omega \rightarrow \omega + i\Gamma/2$, and the damped linear response function can be calculated from first-order wave function parameters, which are now inherently complex. For absorption spectra, this leads to a Lorentzian line-shape function which is identical for all transitions.

Figure 1 shows the damped linear response function calculated from the real part of Eq. 20 with different values of the broadening. Only a section of the frequency

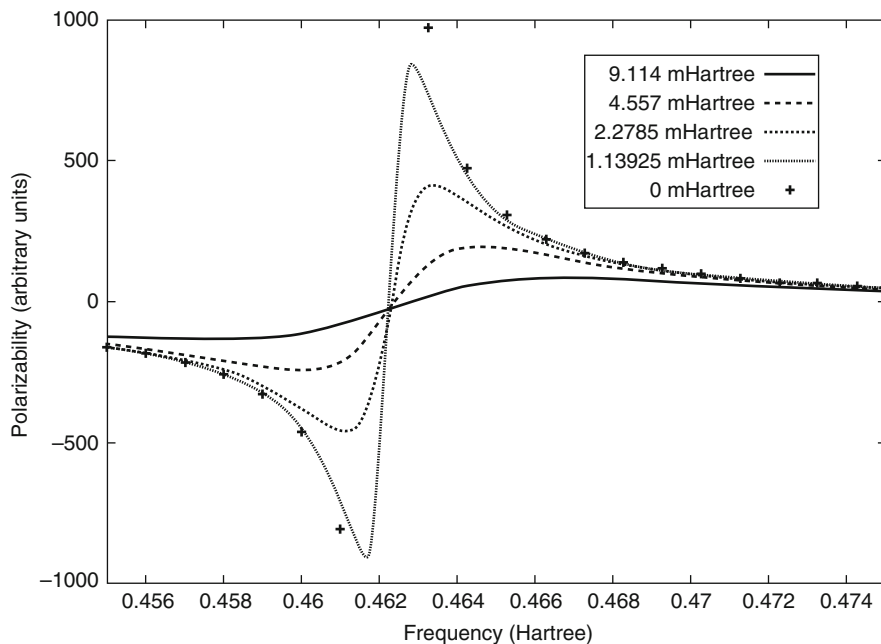


Fig. 1 The damped linear response function computed from the real part of Eq. 20 with the indicated values of the broadening Γ (corresponding to 0, 250, 500, 1,000, and 2,000 cm^{-1}). Electric dipole transition moments and excitation energies were obtained using CCSD for the hydrogen molecule with the aug-cc-pVDZ basis set

spectrum around the lowest excitation energy is shown. With infinite lifetime ($\Gamma = 0$), the linear response function goes to $\pm\infty$ as the frequency approaches the excitation. This is avoided when the finite lifetime ($\Gamma > 0$) is taken into account. In the absorptive region, the behavior of the damped linear response function depends strongly on the value of the broadening. Since the broadening parameter can be chosen freely, the results obtained in this region must be treated with great care: changing the value of the broadening could give a very different result. Sufficiently far from the excitation energy, on the other hand, the effect of the broadening is small and can be neglected. In the transparent region, we may therefore safely use the undamped linear response function.

A long list of molecular electromagnetic properties can be obtained by inserting suitable operators in the linear response function. The electric dipole polarizability is an important example. To further illustrate the procedure of obtaining molecular properties from linear response functions, we now consider the induced electric dipole moment due to higher-order multipole components of the electromagnetic field. The induced electric dipole moment due to a time-dependent uniform magnetic field is, to first order in the field, given by the electric dipole – magnetic dipole polarizability

$$G'_{ij}(-\omega; \omega) = -\text{Im}\langle\langle \mu_i; m_j \rangle\rangle_{\omega} \quad (21)$$

where the j th Cartesian component of the magnetic dipole operator is defined as

$$m_j = -\frac{e}{2m_e} \sum_s L_j^s \quad (22)$$

with the angular momentum operator of electron s given by $\mathbf{L}^s = \mathbf{r}^s \times \mathbf{p}^s$. This identification is obtained from Eq. 3 by substituting $A \rightarrow \mu_i$ and $V(\omega) \rightarrow -\mathbf{m} \cdot \mathbf{B}(\omega)$, where $\mathbf{B}(\omega)$ is the magnetic field amplitude at frequency ω . Using multipole expansions of the electric and magnetic fields, the Maxwell equations show that a time-dependent uniform magnetic field is always accompanied by a quadrupolar electric field, and the contribution to the induced dipole moment from the electric dipole – electric quadrupole polarizability

$$a_{i,jk}(-\omega; \omega) = -\langle\langle \mu_i; q_{jk} \rangle\rangle_{\omega} \quad (23)$$

is of the same order of magnitude as \mathbf{G}' . Here, $q_{jk} = -e \sum_s r_j^s r_k^s$ is the j, k component of the Cartesian electric quadrupole moment operator. The electric dipole – magnetic dipole and electric dipole – electric quadrupole polarizabilities govern the optical rotation of chiral molecules. For samples of randomly oriented molecules, the contribution from the electric dipole – electric quadrupole polarizability averages to zero, and only the trace (the sum of the diagonal elements) of the electric dipole – magnetic dipole polarizability contributes to the optical rotation. The frequency-dependence of the linear response functions gives rise to optical rotation dispersion. The residues of the electric dipole – magnetic dipole polarizability are related to the rotatory strength

$$R_{k0} = \text{Im}[\langle 0 | \mu | k \rangle \cdot \langle k | \mathbf{m} | 0 \rangle] = -\sum_i \lim_{\omega \rightarrow \omega_{k0}} \hbar(\omega - \omega_{k0}) G'_{ii}(-\omega; \omega) \quad (24)$$

which determines the differential intensity of electronic circular dichroism spectra. If the molecules are oriented, the rotatory strength is a tensor (since the intensity will depend on the orientation of the molecule relative to the propagation direction of the photon) that also includes contributions from the residues of the electric dipole – electric quadrupole polarizability.

Finally, we note that the linear response function satisfies the permutation and conjugation relations

$$\langle\langle A; B \rangle\rangle_{\omega} = \langle\langle B; A \rangle\rangle_{-\omega} \quad (25)$$

$$\langle\langle A; B \rangle\rangle_{\omega}^* = \langle\langle A^{\dagger}; B^{\dagger} \rangle\rangle_{-\omega} \quad (26)$$

where the asterisk denotes complex conjugation and A^{\dagger} is the adjoint of the operator A ($A^{\dagger} = A$ for Hermitian operators, by definition). It follows from these relations that the electric dipole polarizability is an even function of frequency,

$\alpha_{ij}(-\omega; \omega) = \alpha_{ij}(\omega; -\omega)$, and a symmetric tensor, $\alpha_{ij}(-\omega; \omega) = \alpha_{ji}(-\omega; \omega)$, which implies that only six of the nine Cartesian components of the polarizability need to be computed. The permutation and conjugation relations are satisfied by approximate linear response functions as well.

Quadratic and Higher-Order Response Functions

The quadratic response function describing the second-order induced electric dipole moment due to a uniform time-dependent electric field is related to the frequency-dependent electric dipole hyperpolarizability as

$$\beta_{ijk}(-\bar{\omega}; \omega, \omega') = -\langle\langle \mu_i; \mu_j, \mu_k \rangle\rangle_{\omega, \omega'} \quad (27)$$

where $\bar{\omega} = \omega + \omega'$. With a procedure completely analogous to that of the linear response functions, this identification is obtained by substituting $A \rightarrow \mu_i$ and $V(\omega) \rightarrow -\mu \cdot \mathbf{F}(\omega)$ in Eq. 3. A variety of optical processes are governed by the electric hyperpolarizability depending on the values of the frequencies, including Second Harmonic Generation through $\beta_{ijk}(-2\omega; \omega, \omega)$ and Optical Rectification through $\beta_{ijk}(0; \omega, -\omega)$. In an approximate calculation, the electric dipole hyperpolarizability can be obtained by computing (minus) the quadratic response function from first-order wave function parameters determined by Eq. 4 with the three Cartesian components of the electric dipole operator as perturbation operators at each set of frequencies ω , ω' , and $\bar{\omega}$. In accordance with Wigner's $2n+1$ rule, which states that the n th-order wave function parameters determine the response functions up to order $2n+1$, second-order wave function parameters are not needed. In this context, the order of the response function corresponds to the number of operators it contains such that the ground state expectation value is of order 1, the linear response function is of order 2, the quadratic response function is of order 3, and so on. The $2n+1$ rule is essential for efficient computer implementations of response theory.

To learn more about the information that can be extracted from the quadratic response function in approximate calculations, we now study the exact one. Like the exact linear response function, the exact quadratic response function can be calculated from excitation energies and transition matrix elements of the unperturbed molecule:

$$\langle\langle A; B, C \rangle\rangle_{\omega, \omega'} = \frac{1}{\hbar^2} P_{BC} \sum_{k, l \neq 0} \left\{ \begin{aligned} & \frac{\langle 0 | A | k \rangle \langle k | \bar{B} | l \rangle \langle l | C | 0 \rangle}{(\omega + \omega' - \omega_{k0}) (\omega' - \omega_{l0})} \\ & + \frac{\langle 0 | C | k \rangle \langle k | \bar{B} | l \rangle \langle l | A | 0 \rangle}{(\omega + \omega' + \omega_{l0}) (\omega' - \omega_{k0})} \\ & - \frac{\langle 0 | B | k \rangle \langle k | \bar{A} | l \rangle \langle l | C | 0 \rangle}{(\omega + \omega_{k0}) (\omega' - \omega_{l0})} \end{aligned} \right\} \quad (28)$$

The bar signals that the operator has been shifted to a zero-point defined by its ground state expectation value. For example,

$$\bar{B} = B - \langle 0 | B | 0 \rangle I \quad (29)$$

where I is the identity operator. The operator P_{BC} permutes the pairs (B, ω) and (C, ω') :

$$P_{BC} f(B, \omega, C, \omega') = f(B, \omega, C, \omega') + f(C, \omega', B, \omega) \quad (30)$$

giving a total of six distinct terms in the quadratic response function. The quadratic response function is singular at the molecular excitation energies, that is, when $\omega + \omega' = \pm \omega_{k0}$ or when $\omega = \pm \omega_{k0}$ or when $\omega' = \pm \omega_{k0}$ for some $k > 0$. These singularities can be avoided by taking into account the finite lifetime of the excited molecular states, as discussed for the linear response function in the previous section.

The quadratic response function, for exact as well as approximate states, satisfies the permutation and conjugation relations

$$\begin{aligned} \langle\langle A; B, C \rangle\rangle_{\omega, \omega'} &= \langle\langle A; C, B \rangle\rangle_{\omega', \omega} \\ &= \langle\langle B; A, C \rangle\rangle_{-\omega-\omega', \omega'} = \langle\langle B; C, A \rangle\rangle_{\omega', -\omega-\omega'} \\ &= \langle\langle C; A, B \rangle\rangle_{-\omega-\omega', \omega} = \langle\langle C; B, A \rangle\rangle_{\omega, -\omega-\omega'} \end{aligned} \quad (31)$$

$$\langle\langle A; B, C \rangle\rangle_{\omega}^* = \langle\langle A^\dagger; B^\dagger, C^\dagger \rangle\rangle_{-\omega, -\omega'} \quad (32)$$

The assumption, known as Kleinman symmetry, that the quadratic (and higher-order) response function is symmetric under interchange of any pair of operators is only true at zero frequency, as is easily verified from Eq. 31. Note that the quadratic response function is symmetric when permuting the operators B and C at $\omega = \omega'$, which is the case for, for example, Second Harmonic Generation: $\beta_{ijk}(-2\omega; \omega, \omega) = \beta_{ikj}(-2\omega; \omega, \omega)$. Kleinman symmetry is often assumed in calculations of the electric dipole hyperpolarizability at low frequencies where it is approximately valid. This reduces the number of independent tensor elements and thus the computational effort.

Owing to the more complicated dependence on field frequencies, the residues of the quadratic response function provide more information than those of the linear response function. One of the most important residues of the quadratic response function is

$$\begin{aligned} &\lim_{\omega' \rightarrow \omega_{m0}} \hbar (\omega' - \omega_{m0}) \langle\langle A; B, C \rangle\rangle_{-\omega, \omega'} \\ &= - \left[\frac{1}{\hbar} \sum_k \left(\frac{\langle 0 | A | k \rangle \langle k | B | m \rangle}{\omega_{k0} - (\omega_{m0} - \omega)} + \frac{\langle 0 | B | k \rangle \langle k | A | m \rangle}{\omega_{k0} - \omega} \right) \right] \langle m | C | 0 \rangle \end{aligned} \quad (33)$$

where the summation over k includes the ground state $|0\rangle$. The quantity in square brackets is an element of the transition matrix for two-photon excitation from the ground state to the excited state m . Another useful (double) residue is

$$\begin{aligned} & \lim_{\omega \rightarrow -\omega_{n0}} \hbar (\omega' + \omega_{m0}) \left[\lim_{\omega' \rightarrow \omega_{m0}} \hbar (\omega' - \omega_{m0}) \langle\langle A; B, C \rangle\rangle_{-\omega, \omega'} \right] \\ & = \delta_{mn} \langle 0 | A | 0 \rangle \langle 0 | B | n \rangle \langle n | C | 0 \rangle - \langle 0 | B | n \rangle \langle n | A | m \rangle \langle m | C | 0 \rangle \end{aligned} \quad (34)$$

From this expression it is possible to extract transition moments between excited states [$\langle n | A | m \rangle$]. With the specific choice $n = m$ it thus becomes possible to compute excited state properties such as dipole moments from the quadratic response of the ground state.

Higher-order response functions (cubic, quartic, quintic, etc.) can also be written in terms of molecular excitation energies and eigenstates. These functions possess singularities at the excitation energies, and the associated residues (single, double, etc. residues) can be used to identify a wide range of molecular properties. For example, residues of the cubic response function provide two-photon transition strengths (the product of two-photon transition matrix elements) and three-photon transition matrix elements, while the three-photon transition strength is obtained from the quintic response function. Linear response functions for the excited states can be extracted from the cubic response function. By calculating ground state response functions we may thus extract all molecular properties.

Static Response Functions

The interaction between a molecule and time-independent external fields is described by frequency-independent perturbation operators. Time-independent fields are nonoscillating, and interactions with molecular electrons are therefore characterized by perturbation operators at zero frequency. The limit $\omega \rightarrow 0$ is often referred to as the static limit and response functions at zero frequency are called static response functions. Although static response functions can be calculated in the same way as dynamic (frequency-dependent) ones, they are special in the sense that they are related to changes in ground state energy. For example, the electronic ground state energy of a molecule in the presence of a static uniform electric field, \mathbf{F} , can be expanded as

$$E = E_0 - \langle 0 | \mu | 0 \rangle \cdot \mathbf{F} - \mathbf{F} \cdot \alpha(0; 0) \cdot \mathbf{F} + \dots \quad (35)$$

where E_0 is the energy in the absence of the field. The expectation value of the dipole moment operator thus determines the first-order energy due to the field, whereas the static polarizability determines the second-order energy. Higher-order energy corrections are determined by the static hyperpolarizability, second

hyperpolarizability, etc. This explains why ground state expectation values are often classified as first-order properties, linear response functions as second-order properties, and so on.

The relation between static response functions and energy corrections implies that molecular properties can be identified as derivatives of the total energy at zero field strength. For example,

$$\langle 0 | \mu_i | 0 \rangle = -\frac{\partial E}{\partial F_i} \quad (36)$$

$$\alpha_{ij}(0; 0) = -\frac{\partial^2 E}{\partial F_i \partial F_j} \quad (37)$$

where it is implicitly understood that the derivatives must be evaluated at zero field strength ($\mathbf{F} = \mathbf{0}$). Hence, the calculation of static response functions does not require a wave function and may therefore be computed with electronic structure theories that only provide the energy (e.g., MP2, CASPT2, and CCSD(T)).

Even if a quantum chemical software package does not feature an implementation of a given static response property, it is often possible to evaluate the property as an energy derivative by finite difference. Most packages allow the addition of external static fields to the Hamiltonian, making it possible to compute the energy at a specified field strength, $E(\mathbf{F})$. From the definition of the partial derivative, we may thus compute the dipole moment and polarizability from finite difference formulas such as

$$\frac{\partial E}{\partial F_i} \approx \frac{E(\delta_i) - E(-\delta_i)}{2\delta_i} \quad (38)$$

$$\frac{\partial^2 E}{\partial F_i^2} \approx \frac{E(\delta_i) - 2E(0) + E(-\delta_i)}{\delta_i^2} \quad (39)$$

$$\frac{\partial^2 E}{\partial F_i \partial F_j} \approx \frac{E(\delta_i, \delta_j) + E(-\delta_i, -\delta_j) - E(-\delta_i, \delta_j) - E(\delta_i, -\delta_j)}{4\delta_i \delta_j} \quad (40)$$

where $\delta_i(\delta_j)$ is a small field strength along the i th (j th) Cartesian direction. The error in these numerical derivative formulas is quadratic in the field strength which should therefore be chosen as small as possible. Choosing too small field strengths, however, may lead to numerical instabilities when computing the difference between two large and almost identical numbers. It is therefore advisable to conduct a convergence study with respect to the field strength when computing numerical derivatives.

A number of static perturbations arise from internal interactions or fields, which are neglected in the nonrelativistic Born–Oppenheimer electronic Hamiltonian. The relativistic correction terms of the Breit–Pauli Hamiltonian are considered as perturbations in nonrelativistic quantum chemistry, including Darwin corrections, the mass-velocity correction, and spin–orbit and spin–spin interactions. Some properties, such as nuclear magnetic resonance shielding tensors and shielding

polarizabilities, are computed from perturbation operators that involve both internal and external fields.

Gauge and Origin Invariance

As described above, molecular properties are identified as response functions by selecting the observable quantity whose change is to be monitored and the external field(s) inducing the change. The operator describing the coupling between the external field and the electrons is not unique, however. The electric field $\mathbf{F}(\mathbf{r}, t)$ can be written in terms of the vector potential $\mathbf{A}(\mathbf{r}, t)$ and the scalar potential $\phi(\mathbf{r}, t)$ as

$$\mathbf{F}(\mathbf{r}, t) = -\frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} - \nabla \phi(\mathbf{r}, t) \quad (41)$$

while the magnetic field $\mathbf{B}(\mathbf{r}, t)$ is given by

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t) \quad (42)$$

Since the electric and magnetic fields are obtained by differentiation, adding a constant to the vector and scalar potentials does not alter the physical electric and magnetic fields. More generally, adding the gradient of a continuous function $\xi(\mathbf{r}, t)$ to the vector potential, that is, $\mathbf{A} \rightarrow \mathbf{A} + \nabla \xi$, while subtracting the time derivative of the same function from the scalar potential, that is, $\phi(\mathbf{r}, t) \rightarrow \phi(\mathbf{r}, t) - \partial \xi / \partial t$ will not change the electric and magnetic fields. This feature is known as gauge invariance and implies that there is a manifold of scalar and vector potentials related by gauge transformations that describe the same physical field. This leads to a manifold of different quantum mechanical operators describing the interaction between the electrons and the field. The time-dependent Schrödinger equation is invariant under such gauge transformations and the computed properties therefore also should be invariant. This is indeed the case for exact response functions.

As an example, consider again the interaction between electrons and a uniform electric field. In section “The Linear Response Function,” the electric dipole polarizability was identified using the length gauge perturbation operator

$$V(\omega) = -\boldsymbol{\mu} \cdot \mathbf{F}(\omega) \quad (43)$$

An equally valid choice would have been the velocity gauge perturbation operator

$$V(\omega) = -\frac{ie}{\omega m_e} \mathbf{p} \cdot \mathbf{A}(\omega) \quad (44)$$

where $\mathbf{A}(\omega)$ is the amplitude vector of the electromagnetic vector potential in the electric dipole approximation and $\mathbf{p} = \sum_s \mathbf{p}^s$ is the total electronic momentum

operator. This operator leads to the following expression for the electric dipole polarizability:

$$\alpha_{ij}(-\omega; \omega) = -\frac{ie}{\omega m_e} \langle\langle \mu_i; p_j \rangle\rangle_{\omega} \quad (45)$$

While Eq. 11 is called the length gauge or dipole-length gauge expression, Eq. 45 is often called the mixed gauge form since it involves both the electric dipole operator and the momentum operator. The length and mixed gauge polarizabilities are equivalent due to the equation of motion

$$\hbar\omega \langle\langle A; B \rangle\rangle_{\omega} = \langle\langle [A, H_0]; B \rangle\rangle_{\omega} + \langle 0 | [A, B] | 0 \rangle \quad (46)$$

which is satisfied by the exact linear response function, and due to the commutator relations

$$[\mu_j, H_0] = -\frac{ie\hbar}{m_e} p_j \quad (47)$$

$$[\mu_i, \mu_j] = 0 \quad (48)$$

Such relations (equation of motion and commutators) guarantee that exact linear response functions are gauge invariant. In addition, transition moments deduced from residues of the exact linear response function are also gauge invariant.

For approximate theories, however, different results are always obtained with different perturbation operators related by gauge transformations. The use of a finite basis set implies that operators are represented as finite matrices and, consequently, the commutator relations no longer hold. As the basis set quality is increased, the commutators converge to the exact values. Gauge invariance thus is recovered in the limit of a complete basis set, provided that the equation of motion Eq. 46, is fulfilled. While this is the case in theories with fully variational orbitals such as DFT, HF, and MCSCF, it is not the case in approximate theories with fixed (nonvariational) orbitals such as CI and CC.

The exact quadratic and higher-order response functions satisfy similar equations of motion, for example,

$$\hbar(\omega + \omega') \langle\langle A; B, C \rangle\rangle_{\omega, \omega'} = \langle\langle [A, H_0]; B, C \rangle\rangle_{\omega, \omega'} + \langle\langle [A, B]; C \rangle\rangle_{\omega'} + \langle\langle [A, C]; B \rangle\rangle_{\omega} \quad (49)$$

ensuring gauge invariance of higher-order molecular properties. In analogy to the case of the linear response function, the higher-order equations of motion are satisfied in approximate theories with variational determination of the orbitals such as DFT, HF, MCSCF, but not in theories such as CI and CC.

It is therefore necessary to specify which gauge (typically, length or velocity) is used when reporting computed response properties such as electric dipole (hyper-)polarizabilities and transition moments. With reasonably flexible basis sets,

each of the DFT, HF, and MCSCF methods typically provides length and velocity gauge results that are quite close to each other, whereas CC results occasionally show significant differences between length and velocity gauge. For example, the specific optical rotation of (1*S*,4*S*)-norbornenone calculated at the CC singles and doubles (CCSD) level is $-558 \text{ deg dm}^{-1} (\text{g/ml})^{-1}$ with the length gauge and $-740 \text{ deg dm}^{-1} (\text{g/ml})^{-1}$ with the velocity gauge (corrected for an unphysical static limit, see below). The fundamental problem is that there is no general physical reason to trust one gauge formulation more than another.

In some cases, however, there may be good reasons for favoring one gauge formulation over others in approximate calculations. One such example is optical rotation which is governed by the trace of the electric dipole – magnetic dipole polarizability, whose length gauge formulation is given by Eq. 21. Within the clamped-nucleus Born–Oppenheimer approximation, every electronic structure calculation is performed with fixed nuclear positions. The coordinate system used to specify the nuclear positions can be chosen arbitrarily. In particular, the results of the calculation should be independent of the choice of origin of the coordinate system.

Suppose that we translate the origin of the coordinate system along a vector \mathbf{O} . As a consequence, the positions of nuclei and electrons are shifted by the same vector, for example, $\mathbf{r}^s \rightarrow \mathbf{r}^s - \mathbf{O}$, and the electric dipole – magnetic dipole polarizability changes according to

$$\mathbf{G}'(-\omega; \omega) \rightarrow \mathbf{G}'(-\omega; \omega) + \frac{e}{2m_e} \text{Im} [\langle \langle \boldsymbol{\mu}; \mathbf{p} \rangle \rangle_{\omega} \times \mathbf{O}] \quad (50)$$

which shows that the tensor is origin dependent. The culprit is the origin-dependence of the magnetic dipole operator:

$$\mathbf{m} \rightarrow \mathbf{m} - \frac{e}{2m_e} \mathbf{p} \times \mathbf{O} \quad (51)$$

This translation of the operator can also be achieved by a particular type of gauge transformation of the magnetic vector potential, explaining why the term *gauge-origin* is often used for \mathbf{O} in the context of magnetic properties.

It can be shown that the trace of the tensor \mathbf{G}' , and hence the computed optical rotation of a sample of randomly oriented chiral molecules, is independent of the origin provided that the linear response function satisfies Eq. 46 and that the commutator of Eq. 47 is fulfilled. Consequently, approximate linear response calculations of the length gauge optical rotation depend on the chosen coordinate origin. On the other hand, the trace of the velocity gauge formulation of the electric dipole – magnetic dipole polarizability

$$\mathbf{G}'(-\omega; \omega) = \frac{e}{\omega m_e} \langle \langle \mathbf{p}; \mathbf{m} \rangle \rangle_{\omega} \quad (52)$$

is unconditionally independent of the chosen origin, also in approximate theories. (Note that Eq. 52 is only valid for real wave functions.) It may thus be argued that

the velocity gauge formulation of optical rotation should be preferred over the length gauge expression. It must be added, however, that the velocity gauge optical rotation calculated with an approximate theory suffers from a serious artifact: it predicts a nonvanishing optical rotation at zero frequency (at zero frequency, there is no light and the optical rotation must be zero). This can be traced back to the lack of gauge invariance and is rectified by the modified velocity gauge where the static limit simply is subtracted from the linear response function, that is, $\langle\langle p_i; m_i \rangle\rangle_\omega$ is replaced by $\langle\langle p_i; m_i \rangle\rangle_\omega - \langle\langle p_i; m_i \rangle\rangle_0$.

The most widely used technique for ensuring origin invariance of magnetic properties, however, is based on so-called Gauge Including Atomic Orbitals (GIAOs, also known as London Atomic Orbitals [LAOs]). Belonging to the class of perturbation-dependent basis sets, GIAOs are obtained from any conventional atomic orbital basis set by multiplying each basis function with a complex phase factor that depends explicitly on an external uniform static magnetic field. In a finite GIAO basis set, the magnetic dipole moment operator depends on the chosen coordinate origin according to

$$\mathbf{m} \rightarrow \mathbf{m} - \frac{i}{2\hbar} [\boldsymbol{\mu}, H_0] \times \mathbf{O} \quad (53)$$

which is identical to Eq. 51 in the limit of a complete basis set where the commutator of Eq. 47 is fulfilled. For fully variational approximate theories (DFT, HF, MCSCF), GIAOs thus remove the condition that the commutator must be fulfilled and the length gauge optical rotation becomes manifestly independent of origin. For approximate theories lacking variational orbital optimization (CI, CC), GIAOs are unable to remove the origin dependence of the length gauge optical rotation, since the equation of motion, Eq. 46, is not satisfied. The velocity gauge formulation with GIAOs is origin dependent in all approximate theories and, therefore, is never used.

The greatest strength of GIAOs lies in origin invariant calculation of molecular properties involving static uniform magnetic fields, including static magnetic properties such as magnetizabilities and rotational g tensors, and nuclear magnetic shieldings as well as optical properties aided by a static magnetic field such as magneto-optical activity. The use of GIAOs makes it possible to calculate the total energy in the presence of a static uniform magnetic field in an origin-independent manner for variational as well as nonvariational approximate theories, implying that the molecular properties identified as energy derivatives are independent of origin as well. In addition, owing to their being atomic eigenfunctions of the angular momentum operator correct to first order in the applied magnetic field, GIAOs provide vastly improved basis set convergence of static magnetic properties.

Effects of Nuclear Motion

Within the clamped-nucleus Born–Oppenheimer approximation, electronic response functions are computed for fixed nuclear geometries. Just like the electronic ground state energy, the excitation energies and response functions

thus become parametric functions of nuclear positions. For example, by solving the eigenvalue problem, Eq. 15, at a range of geometries, it becomes possible to compute excited state potential energy surfaces using Eq. 16 at each point. This, in turn, allows us to optimize excited state geometries using either a simple scan of the excited state potential energy surface or analytic gradients derived from Eq. 16. Since the ground state energy is required to compute the excited state energy in the response approach, calculations of excited state potential energy surfaces should only be carried out for geometries where the approximate wave function is valid. For single-reference methods such as HF and CC, one should only compute excitation energies (and thus excited state energies) for geometries not too far from the ground state equilibrium structure. As a consequence, these single-reference methods cannot be used to compute dissociative excited state surfaces, except at points close to the ground state equilibrium geometry. Multiconfigurational methods are required to describe such unbound excited states. For bound excited states, Eq. 16 also provides a viable path to excited state vibrational frequencies and wave functions. Along with those of the ground state, it then becomes possible to calculate 0–0 transition energies rather than the vertical excitation energies provided by the poles of the response functions. In addition, the vibrational structure of electronic bands becomes amenable to theoretical computations.

Electronic ground state response functions are most often calculated at the (ground state) equilibrium geometry. The geometry optimization and response function evaluation may be performed at different levels of theory and using different basis sets. For example, it may be sufficient to optimize the geometry of a covalently bonded molecule at the DFT level of theory with a relatively small polarized basis set, while the response functions are evaluated at the CC level of theory (at the DFT equilibrium geometry) with a large polarized basis set augmented with diffuse functions to ensure the correct pole structure of the response functions. In other cases, it may be advantageous to use different DFT functionals and basis sets for the geometry optimization and property calculation.

If high accuracy is required, vibrational effects must be taken into account. In a proper adiabatic Born–Oppenheimer treatment, the ground state wave function would be written as a product of an electronic and a vibrational wave function. The response of this wave function should then be computed and subsequently used to construct vibronic response functions. The sum-over-states expressions would include contributions from vibrational states in the electronic ground and excited states. Since each set of vibrational wave functions is tied to a specific electronic state within the adiabatic Born–Oppenheimer approximation, this approach is not feasible in practice. Hence, the electronic properties are considered as electronic ground state properties and therefore averaged in a vibrational state of the electronic ground state.

Computing the expectation value of the electronic property $P(\mathbf{q})$, which may be a response function or a sum of response functions depending parametrically on the dimensionless normal coordinates \mathbf{q} , in the vibrational state $|v_0\rangle$ of the electronic ground state, the vibrational correction becomes

$$\Delta P = \langle v_0 | P(\mathbf{q}) | v_0 \rangle - P_0 \quad (54)$$

where P_0 is the value of the electronic property at the equilibrium geometry. The analytic form of the dependence of P on normal coordinates normally is not known. Even if it were known, it would be too computationally demanding to evaluate the property on a grid of normal coordinates to fit the values to the analytic form. Instead, the property is expanded in a Taylor series. This implies that derivatives of the property with respect to the normal coordinates must be evaluated at the expansion point of the Taylor series. The property derivatives can themselves be expressed in terms of response functions (by considering the nuclear displacement terms of the Hamiltonian as perturbations), or they may be computed by numerical techniques analogous to the static properties of section “Static Response Functions.” The vibrational wave functions may be computed in the same manner as in the ground state theory. That is, they are computed on the basis of the harmonic approximation and often include anharmonic effects to first order in perturbation theory. In another approach, the vibrational problem is viewed as so-called mode dynamics and solved in a manner resembling electronic structure theory. This approach also makes it possible to compute the vibrational expectation value directly and only requires that an electronic structure program is available for computing energies and properties at specified nuclear geometries.

The relatively small difference between vibrational energy levels implies that a range of vibrational states are occupied at finite temperature. Including anharmonic effects to first order, the temperature-dependent vibrational correction can be expressed as

$$\begin{aligned} \Delta P = & -\frac{1}{4\hbar} \sum_n \frac{1}{\omega_n} \frac{\partial P}{\partial q_n} \sum_m k_{nmm} \coth\left(\frac{\hbar\omega_m}{2k_B T}\right) \\ & + \frac{1}{4} \sum_n \frac{\partial^2 P}{\partial q_n^2} \coth\left(\frac{\hbar\omega_n}{2k_B T}\right) \end{aligned} \quad (55)$$

where the property derivatives are to be evaluated at the equilibrium geometry, ω_n is the harmonic vibrational frequency of normal mode n , k_B is the Boltzmann constant, T is the temperature, and the cubic (anharmonic) force constants are given by

$$k_{nmm} = \frac{\partial^3 E_0}{\partial q_n \partial q_m^2} \quad (56)$$

which is also evaluated at the equilibrium geometry (E_0 is the electronic ground state energy surface, $E_0 = E_0(\mathbf{q})$). The first term in Eq. 55 accounts for the anharmonicity of the potential energy surface, whereas the second term is purely harmonic and arises from the curvature of the property surface. If anharmonic effects are neglected, only the second term is included. Anharmonic effects can be as large as the harmonic contribution and should generally be included, in particular when the molecule contains low-frequency vibrational modes (such as torsional

motions). Most quantum chemistry programs are able to compute the harmonic vibrational frequencies and normal coordinates, and some also provide cubic force constants (if not, these can be computed numerically). The property derivatives are most often computed by numerical differentiation. At zero temperature, only the vibrational ground state is populated and the zero-point vibrational correction (ZPVC) is obtained from Eq. 55 by taking the limit $T \rightarrow 0$ where the hyperbolic cotangent factors approach 1. Once the property and energy derivatives have been calculated, the temperature dependence of the vibrationally averaged property can be easily computed, including the ZPVC.

Conformationally flexible molecules are characterized by multiple low-lying minima on the potential energy surface, each minimum defining a conformer of the molecule. All conformers contribute to the experimentally observed property and must be taken into account in theoretical calculations. In the simplest approach, provided that the barriers separating the minima are sufficiently large, the vibrationally corrected property can be computed along the lines discussed above for each conformer (letting the conformer structure play the role of equilibrium geometry in each case). The total (observed) property is then obtained as a Boltzmann average of the results for each conformer:

$$\bar{P} = \sum_i^{\text{conformers}} \langle P \rangle_i X_i \quad (57)$$

Here, $\langle P \rangle_i$ and X_i are the vibrationally corrected property and mole fraction of conformer i , respectively. Note that the same formula can be applied to compute the average of the purely electronic contribution to the property (i.e., not vibrationally corrected). The mole fraction is computed according to

$$X_i = \frac{\exp(-G_i/k_B T)}{\sum_j^{\text{conformers}} \exp(-G_j/k_B T)} \quad (58)$$

where G_i is Gibbs free energy of conformer i , which, assuming the ideal gas approximation, is given by

$$G_i = E_{0,i} + \epsilon_{0,i} + RT - TS_i \quad (59)$$

where R is the universal gas constant. The electronic ground state energy of conformer i is denoted $E_{0,i}$ and $\epsilon_{0,i}$ is the zero-point vibrational energy of conformer i in the harmonic oscillator approximation. The entropy is computed from the electronic, vibrational, and rotational partition functions,

$$S_i = -k_B T \ln(Z_{\text{el},i} Z_{\text{vib},i} Z_{\text{rot},i}) \quad (60)$$

$$Z_{\text{el},i} = 1 \quad (61)$$

$$Z_{\text{vib},i} = \sum_n g_{n,i}^{\text{vib}} \exp(-\epsilon_{n,i}/k_B T) \quad (62)$$

$$Z_{\text{rot},i} = \sum_m g_{m,i}^{\text{rot}} \exp(-\epsilon_{m,i}/k_B T) \quad (63)$$

The electronic partition function is unity, as only the electronic ground state is assumed populated. The vibrational energies of conformer i , $\epsilon_{n,i}$, are computed in the harmonic oscillator approximation, which facilitates analytic summation over all vibrational levels. The rotational energy of conformer i , $\epsilon_{m,i}$, is calculated within the rigid rotor approximation. The factors $g_{n,i}^{\text{vib}}$ and $g_{m,i}^{\text{rot}}$ are the degeneracies of the vibrational and rotational energy levels, respectively. Hence, the data needed to evaluate the Boltzmann average can be obtained from ground state calculations (electronic ground state energy at the equilibrium structure, harmonic vibrational energies, and rotational constants).

The simple Boltzmann averaging relies on a range of approximations that tend to benefit from fortuitous cancellation of errors. In particular, the assumption that the conformers can be treated as independent with no couplings between their vibrational states is fragile, as barriers are often insufficient to warrant the assumption. A more correct description would be obtained by solving the vibrational problem for the entire ground state potential energy surface instead of using a first-order anharmonic approximation in the vicinity of the minima. Calculation of the entire potential energy surface is, however, out of the question for all but the smallest systems. One way out might be to compute the nuclear motion using molecular dynamics and parametrized force fields, sampling the electronic property along the trajectory. Also this approach, however, may easily turn out to be too demanding in terms of the number of electronic structure calculations required.

Further Reading

This section contains a short description of suggested references for further study of the concepts and techniques described above. The list of references is highly incomplete and should be considered as a mere starting point.

The time-independent perturbation operators describing internal molecular interactions are taken from the Breit–Pauli Hamiltonian, which includes relativistic corrections to the molecular electronic Hamiltonian of Eq. 2. The Breit–Pauli Hamiltonian is discussed in more detail in the textbooks by Bethe and Salpeter (1957) and by Moss (1973). Appropriate operators for the interaction between charged particles and external electromagnetic fields are discussed in the paper by Fiutak (1963). The textbook by Craig and Thirunamachandran (1984) contains derivations of a long range of optical properties in terms of exact molecular states, while Barron’s book (1982) is highly recommended for its thorough discussion of optical activity phenomena. The expressions derived in these books may be

translated into response functions amenable to calculation by means of approximate wave functions.

Olsen and Jørgensen (1985, 1995) have derived and discussed response functions for exact, HF, and MCSCF wave functions in great detail, while Koch and Jørgensen (1990) presented a derivation for CC wave functions. The latter was modified by Pedersen and Koch (1997) to ensure proper symmetry of the response functions. Christiansen et al. (1998) have presented a derivation of dynamic response functions for variational as well as nonvariational wave functions that resembles the way in which static response functions are deduced from energy derivatives. Linear and higher-order response functions based on DFT have been presented by Sałek et al. (2002). Damped response theory has been discussed by Norman et al. (2001) in the context of HF and MCSCF response theory. Nonperturbative calculations of static magnetic properties at the HF level have been presented by Tellgren et al. (2008, 2009).

Gauge invariance of HF and MCSCF response theory has been shown in terms of the equations of motion for the response functions by Olsen and Jørgensen (1985, 1995), whereas the lack of gauge invariance in CC response theory was demonstrated by Pedersen et al. (Pedersen and Koch 1997; Pedersen et al. 2001). The use of GIAOs to ensure origin invariance of static magnetic properties has been discussed by Helgaker et al. (Helgaker and Jørgensen 1991; Olsen et al. 1995) and for optical rotations by Bak et al. (1995).

Different approaches to the calculation of vibrational corrections to response properties can be found in the work of Sauer and Pack (2000), Ruud et al. (2000), and Kongsted and Christiansen (2006). The Boltzmann averaging procedure for conformationally flexible molecules has been critically reviewed by Crawford and Allen (2009). Mort and Autschbach (2008) have proposed an approach based on a decoupling of hindered rotations from the remaining (high-frequency) vibrational modes, which allows for a separate calculation of the hindered rotations without invoking the harmonic approximation.

Acknowledgments The author wishes to thank Profs. Michał Jaszuński, Antonio Rizzo, T. Daniel Crawford, and Trygve Helgaker for commenting on the manuscript.

Bibliography

- Bak, K. L., Hansen, A. E., Ruud, K., Helgaker, T., & Olsen, J. (1995). *Ab initio* calculation of electronic circular dichroism for *trans*-Cyclooctene using London atomic orbitals. *Theoretica Chimica Acta*, 90, 441–458.
- Barron, L. D. (1982). *Molecular light scattering and optical activity*. Cambridge, MA: Cambridge University Press.
- Bethe, H. A., & Salpeter, E. E. (1957). *Quantum mechanics of one- and two-electron atoms*. New York: Academic.
- Christiansen, O., Jørgensen, P., & Hättig, C. (1998). Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy. *International Journal of Quantum Chemistry*, 68, 1–52.

- Craig, D. P., & Thirunamachandran, T. (1984). *Molecular quantum electrodynamics*. London: Academic.
- Crawford, T. D., & Allen, W. D. (2009). Optical activity in conformationally flexible molecules: A theoretical study of large-amplitude vibrational averaging in (R)-3-chloro-1-butene. *Molecular Physics*, *107*, 1041–1057.
- Fiutak, J. (1963). The multipole expansion in quantum theory. *Canadian Journal of Physics*, *41*, 12–20.
- Helgaker, T., & Jørgensen, P. (1991). An electronic Hamiltonian for origin independent calculations of magnetic properties. *Journal of Chemical Physics*, *95*, 2595–2601.
- Koch, H., & Jørgensen, P. (1990). Coupled cluster response functions. *Journal of Chemical Physics*, *93*, 3333–3344.
- Kongsted, J., & Christiansen, O. (2006). Automatic generation of force fields and property surfaces for use in variational vibrational calculations of anharmonic vibrational energies and zero-point vibrational averaged properties. *Journal of Chemical Physics*, *125*, 124108.
- Mort, B. C., & Autschbach, J. (2008). A pragmatic approach for the treatment of hindered rotations in the vibrational averaging of molecular properties. *European Journal of Chemical Physics and Physical Chemistry*, *9*, 159–170.
- Moss, R. E. (1973). *Advanced molecular quantum mechanics*. London: Chapman and Hall.
- Norman, P., Bishop, D. M., Jensen, H. J. A., & Oddershede, J. (2001). Near-resonant absorption in the time-dependent self-consistent field and multiconfigurational self-consistent field approximations. *Journal of Chemical Physics*, *115*, 10323–10334.
- Olsen, J., & Jørgensen, P. (1985). Linear and nonlinear response functions for an exact state and for an MCSCF state. *Journal of Chemical Physics*, *82*, 3235–3264.
- Olsen, J., & Jørgensen, P. (1995). Time-dependent response theory with applications to self-consistent field and multiconfigurational self-consistent field wave functions. In D. R. Yarkony (Ed.), *Modern electronic structure theory* (Vol. 2, pp. 857–990). Singapore/River Edge: World Scientific.
- Olsen, J., Bak, K. L., Ruud, K., Helgaker, T., & Jørgensen, P. (1995). Orbital connections for perturbation-dependent basis sets. *Theoretica Chimica Acta*, *90*, 421–439.
- Pedersen, T. B., & Koch, H. (1997). Coupled cluster response functions revisited. *Journal of Chemical Physics*, *106*, 8059–8072.
- Pedersen, T. B., Fernández, B., & Koch, H. (2001). Gauge invariant coupled cluster response theory using optimized nonorthogonal orbitals. *Journal of Chemical Physics*, *114*, 6983–6993.
- Ruud, K., Astrand, P.-O., & Taylor, P. R. (2000). An efficient approach for calculating vibrational wave functions and zero-point vibrational corrections to molecular properties of polyatomic molecules. *Journal of Chemical Physics*, *112*, 2668–2683.
- Safek, P., Vahtras, O., Helgaker, T., & Agren, H. (2002). Density-functional theory of linear and nonlinear time-dependent molecular properties. *Journal of Chemical Physics*, *117*, 9630–9645.
- Sauer, S. P. A., & Packer, M. J. (2000). The Ab initio calculation of molecular properties other than the potential energy surface. In P. R. Bunker & P. Jensen (Eds.), *Computational molecular spectroscopy* (pp. 221–252). London: Wiley.
- Tellgren, E. I., Soncini, A., & Helgaker, T. (2008). Nonperturbative *ab initio* calculations in strong magnetic fields using London orbitals. *Journal of Chemical Physics*, *129*, 154114.
- Tellgren, E. I., Helgaker, T., & Soncini, A. (2009). Non-perturbative magnetic phenomena in closed-shell paramagnetic molecules. *Physical Chemistry Chemical Physics*, *11*, 5489–5498.

Alston J. Misquitta

Contents

Introduction	296
Definition of the Interaction Energy	297
The Two-Body Interaction Energy	297
Practical Methods for the Two-Body Interaction Energy	301
Supermolecular Methods	301
Perturbation Theory	303
Multipole Expansion for the Interaction Energy	310
Damping Functions	314
Many-Body Contributions to the Interaction Energy	314
Molecular Properties	316
Distributed Multipoles	316
Distributed Polarizabilities	317
Distributed Dispersion Coefficients	320
Case Studies	322
Applications	322
Polarization in Organic Crystals	324
Crystal Structure Prediction	325
Outlook	326
Programs	327
Annotated Bibliography	328
Bibliography	329

Abstract

Van der Waals interactions determine a number of phenomena in the fields of physics, chemistry and biology. As we seek to increase our understanding of

A.J. Misquitta (✉)

School of Physics, Queen Mary, University of London, London, UK

TCM Group, Cambridge, UK

e-mail: am592@cam.ac.uk

physical systems and develop detailed and more predictive theoretical models, it becomes even more important to provide an accurate description of the underlying molecular interactions. The goal of this chapter is to describe recent developments in the theory of intermolecular interactions that have revolutionised the field due to their comparatively low computational costs and high accuracies. These are the symmetry-adapted perturbation theory based on density functional theory (SAPT(DFT)) for interaction energies and the Williams–Stone–Misquitta (WSM) method for molecular properties in distributed form. These theories are applicable to systems of small organic molecules containing as many as 30 atoms each and have demonstrated accuracies comparable to the best electronic structure methods. We also discuss the numerical aspects of these theories and recent applications which demonstrate the range of problems that can now be approached with these accurate *ab initio* methods.

Introduction

Intermolecular, or van der Waals, interactions are responsible for a wide variety of phenomena in the fields of physics, chemistry, and biology. The thermodynamic properties of gases, liquids, and solids depend on these interactions. In fact, many liquids and many solids would not exist without van der Waals forces. Amongst other properties that depend strongly on the van der Waals interactions are microwave and infrared spectra of molecular complexes and bulk phases. Due to their long-range nature, van der Waals forces determine the entrance channels for chemical reactions. Some of the more exotic systems in which van der Waals interactions play an important role are Bose–Einstein condensates (BECs) and helium nanodroplets (Chang et al. 2003; Xu et al. 2003). In biological systems, van der Waals interactions are particularly important, subtle, and often hard to model. The stability of DNA and RNA arise, in part, from the stacking energy (Hobza and Sponer 2002), which is determined by this interaction. One of the important problems in which van der Waals interactions play an important role is the problem of protein folding (Lehninger 1970).

This chapter is an introduction to the field of intermolecular interaction and to the modern *ab initio* electronic structure methods – primarily those based on perturbation theory – that have been developed to study them. We will be mainly concerned with applications to small organic molecules for which accuracies of the order of a kJ mol^{-1} or less are sufficient. High-accuracy calculations on small dimers can be orders of magnitude more accurate, but these are the subject of a specialist review (see Szalewicz et al. 2005 for a review and references). Nor are we concerned with empirical methods. Our focus will be on first principles methods for the interactions of closed-shell systems in the non-relativistic limit. In the last decade, *ab initio* methods have been used to successfully model the structure of liquid water (Bukowski et al. 2007) studied the interactions between DNA base tetramers (Fiethen et al. 2008) and predicted the crystal structure of an organic

molecule (Misquitta et al. 2008b). The goal of this chapter is to describe the main theoretical developments that have been responsible for these applications.

Definition of the Interaction Energy

The interaction energy of a cluster of N interacting *rigid* molecules is defined to be

$$E_{\text{int}} = E_{ABC\dots} - E_A - E_B - E_C - \dots \quad (1)$$

Here $E_{ABC\dots}$ is the energy of the cluster and E_X , $X = A, B, C$, etc., is the energy of molecule X . Non-rigid molecules undergo a deformation in the cluster, with an associated *deformation energy cost*: $\delta E_X = E_X(x_X^*) - E_X(x_X^0)$, where x_X^0 is the geometry of monomer X in isolation and x_X^* is the geometry in the cluster. This deformation energy cost should be included as part of the interaction energy defined above, but since this chapter is mainly concerned with the interactions of rigid molecules, we will assume that the deformation energies δE_X are obtained in a suitable manner, and concern ourselves only with the interaction energy defined by Eq. 1.

E_{int} can be evaluated directly, but for computational efficiency as well as physical interpretation it is worthwhile to partition the N -body interaction energy into contributions from dimers, trimers, and so on. This leads to an exact reformulation of E_{int} that is known as the many-body expansion:

$$E_{\text{int}}(ABC\dots) = \sum_{X<Y} E_{\text{int}}(XY) + \sum_{X<Y<Z} \Delta E_{\text{int}}(XYZ) + \dots, \quad (2)$$

where $\Delta E_{\text{int}}(XYZ)$ is the three-body correction, defined as

$$\Delta E_{\text{int}}(XYZ) = E_{\text{int}}(XYZ) - E_{\text{int}}(XY) - E_{\text{int}}(XZ) - E_{\text{int}}(YZ). \quad (3)$$

In the same way, we can define four-body corrections, five-body corrections, and so on. For a cluster of N molecules, this expansion terminates at the N -body correction.

The many-body expansion would not be of much use if we had to evaluate all terms in Eq. 2. But experience has shown that the expansion converges quickly and terms beyond those involving three bodies are not so important. This is fortunate as the two-body interactions are well understood and can be evaluated for moderate-sized molecules using a variety of methods, while good approximations are available for the terms involving three and more bodies, which usually arise from the effects of polarization in the cluster.

The Two-Body Interaction Energy

The interaction energy of a pair of interacting molecules is defined as

$$E_{\text{int}} = E_{AB} - E_A - E_B. \quad (4)$$

If we calculate all three energies using a suitable electronic structure method – some of which will be described later – this definition provides a simple method for calculating the interaction energy of a pair of interacting molecules. This approach, known as the *supermolecular* method, provides us with a single number: the interaction energy E_{int} . We gain considerable physical insight into E_{int} by evaluating it not as the difference in energies suggested by the above definition, but through *perturbation theory*, which enables us to partition E_{int} into physical components: the electrostatic, induction (or polarization), dispersion, and exchange-repulsion energies. This breakup proves invaluable not only as an aid to interpretation, but also forms the basis for all analytic atom–atom potentials (Stone and Misquitta 2007).

Consider the dimer Hamiltonian partitioned as

$$H(\lambda) = H_0 + \lambda V, \quad (5)$$

where $H_0 = H_A + H_B$ is the sum of the Hamiltonians of the unperturbed monomers A and B , and V is the intermonomer interaction operator that consists of electron–electron, electron–nuclear, and nuclear–nuclear interactions between the monomers. We may now carry out a Rayleigh–Schrödinger perturbation theory – also known as the *polarization expansion* (Jeziorski et al. 1994) – using as our zeroth-order wave function the unsymmetrized product $\Phi_0 = \Phi_0^A \Phi_0^B$. The zeroth-order energy is then $E_0 = E_0^A + E_0^B$. The interaction energy appears at first and higher orders in perturbation theory (setting $\lambda = 1$):

$$E_{\text{int}} = E_{\text{pol}}^{(1)} + E_{\text{pol}}^{(2)} + \dots, \quad (6)$$

where we have used the subscript “pol” to indicate these are energies from the polarization expansion and do not include any of the effects of electron exchange between monomer. The individual terms in this series can be evaluated using the usual methods of perturbation theory (Stone 1996). At first order we have

$$E_{\text{pol}}^{(1)} = \left\langle \Phi_0^A \Phi_0^B \left| V \Phi_0^A \Phi_0^B \right. \right\rangle, \quad (7)$$

which is identified as the electrostatic energy of the interacting dimer and can be written in terms of the unperturbed total charge densities of the monomers:

$$E_{\text{pol}}^{(1)} = E_{\text{elst}}^{(1)} = \iint \rho_A^{\text{tot}}(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \rho_B^{\text{tot}}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2. \quad (8)$$

Here, the total charge density of monomer X is defined as

$$\rho_X^{\text{tot}} = \sum_{\alpha \in X} Z_\alpha \delta(\mathbf{r} - \mathbf{R}_\alpha) - \rho_X \quad (9)$$

where ρ_X is the unperturbed electronic charge density of the monomer and the term containing the Dirac delta functions represents the nuclear charge density.

At second order we get

$$E_{\text{pol}}^{(2)} = \sum'_{rs} \frac{\left| \left\langle \Phi_0^A \Phi_0^B \middle| V \Phi_r^A \Phi_s^B \right\rangle \right|^2}{E_0^A + E_0^B - E_r^A - E_s^B}, \quad (10)$$

where r and s label the excited states of monomers A and B respectively. The prime indicates that at least one of r or s refer to an excited state. The above expression can be split into two physically motivated terms: the second-order induction and second-order dispersion energies. The second-order induction energy is obtained when the summation in Eq. 10 is restricted to those dimer states including only one excited monomer. This has two natural contributions: the ground state of monomer A interacting with excitations on monomer B, and vice versa. Thus,

$$E_{\text{ind,pol}}^{(2)} = E_{\text{ind,pol}}^{(2)}(A) + E_{\text{ind,pol}}^{(2)}(B), \quad (11)$$

where,

$$E_{\text{ind,pol}}^{(2)}(A) = \sum'_r \frac{\left| \left\langle \Phi_0^A \Phi_0^B \middle| V \Phi_r^A \Phi_0^B \right\rangle \right|^2}{E_0^A - E_r^A}. \quad (12)$$

Similarly, we can define $E_{\text{ind,pol}}^{(2)}(B)$.

From Eq. 12, we see that $E_{\text{ind,pol}}^{(2)}(A)$ may be interpreted as the second-order energy correction arising when the monomer A is perturbed by the external field Ω_B of monomer B. Thus the second-order induction energy can be interpreted as a *response* to a static field, and is therefore also called the polarization energy.

The remaining part of the second-order term $E_{\text{pol}}^{(2)}$ is the dispersion energy which is defined as

$$E_{\text{disp,pol}}^{(2)} = \sum_{r \neq 0} \sum_{s \neq 0} \frac{\left| \left\langle \Phi_0^A \Phi_0^B \middle| V \Phi_r^A \Phi_s^B \right\rangle \right|^2}{E_0^A + E_0^B - E_r^A - E_s^B}. \quad (13)$$

The dispersion is purely quantum mechanical in origin and is the only term at second order in λ that describes intermonomer electron correlation. The second-order dispersion energy is long ranged, always negative, and exists between all types of molecules.

The starting wave function $\Phi_0 = \Phi_0^A \Phi_0^B$ used in the polarization expansion is not antisymmetric with respect to electron interchanges between the monomers and thus cannot be expected to be reliable when overlap effects between the monomers become important. In this region, only a fully antisymmetric theory is valid. Various procedures exist to achieve the complete antisymmetrization and these are discussed by Jeziorski and Szalewicz (2002) and Szalewicz et al. (2005). The result is that corresponding to each of the three interaction energy components discussed above we get an exchange counterpart. At first order we get the

first-order exchange-repulsion energy $E_{\text{exch}}^{(1)}$, and at second order we get the second-order exchange-induction and exchange-dispersion energies, $E_{\text{ind,exch}}^{(2)}$ and $E_{\text{disp,exch}}^{(2)}$.

We now define the *total* induction and *total* dispersion energies as the sum of their polarization and exchange terms:

$$E_{\text{ind,tot}}^{(2)} = E_{\text{ind,pol}}^{(2)} + E_{\text{ind,exch}}^{(2)} \quad (14)$$

$$E_{\text{disp,tot}}^{(2)} = E_{\text{disp,pol}}^{(2)} + E_{\text{disp,exch}}^{(2)}. \quad (15)$$

These sums are the physically relevant energies; the individual components are useful from the theoretical standpoint, but are somewhat arbitrary and depend on the particular perturbation theory used.

Therefore, the total interaction energy can be defined as

$$E_{\text{int}} = E_{\text{elst}}^{(1)} + E_{\text{exch}}^{(1)} + E_{\text{ind,tot}}^{(2)} + E_{\text{disp,tot}}^{(2)} + \Delta E_{\text{int}}^{(3-\infty)}, \quad (16)$$

where the last term collects all contributions of third order and higher in the interaction operator V . We will describe the theoretical methods for calculating the interaction energy components in section “[Perturbation Theory](#).” Now let us see the kind of interpretational power perturbation theory gives us by taking a look at a few examples.

Examples

Based on the relative contributions of the interaction energy components, we can classify intermolecular bonds. For example, hydrogen-bonded complexes tend to be bound primarily by the electrostatic and induction (polarization) components of the interaction energy, while van der Waals complexes are bound primarily by the dispersion energy. Of course, most complexes will fall in between, so it is probably better not to regard such classifications as being rigorous, but they are useful as they provide us with a physical picture of the mechanisms of the intermolecular bond.

Interaction energies and components for three representative complexes are presented in Table 1. We have used both approaches: perturbation theory and the supermolecular method.

- *Hydrogen bonding*: At its minimum energy geometry, the water dimer shows a bonding pattern that is typical of hydrogen-bonded complexes (Buckingham et al. 2008). Binding is primarily due to the electrostatic and induction energies. In fact, this is the only one of the three complexes which is bound at first order, that is, the sum of the first-order energies is negative. Even so, the dispersion energy is certainly non-negligible and contributes almost 40 % of the binding of this complex. The terms of third and higher orders contribute significantly to the binding of these complexes.
- *van der Waals bonding*: The benzene dimer in the sandwich geometry is a typical example of a van der Waals bound complex. The dimer is unbound at first order and the electrostatic and induction energies make minor contributions to the interaction energy. Almost all the binding arises at second order from the

Table 1 Interaction energy components for the water dimer, benzene dimer, and the $\text{Cl}_2 \cdots \text{OH}_2$ complex. The water dimer is in its equilibrium geometry, the benzene dimer is in the parallel stacked geometry with center-of-mass separation 3.7 Å, and the $\text{Cl}_2 \cdots \text{OH}_2$ complex is in a symmetrical complex with all atoms in a plane and collinear with $r(\text{Cl} \cdots \text{O}) = 2.755$ Å. All energies are reported in kJ mol^{-1}

Energy	$(\text{H}_2\text{O})_2$	$\text{Cl}_2 \cdots \text{OH}_2$	$(\text{C}_6\text{H}_6)_2$
Perturbation theory [SAPT(DFT)]:			
$E_{\text{elst}}^{(1)}$	-28.90	-19.24	-1.63
$E_{\text{exch}}^{(1)}$	23.97	24.96	20.99
$E_{\text{ind,pol}}^{(2)}$	-10.96	-16.09	-7.65
$E_{\text{ind,exch}}^{(2)}$	5.79	12.21	6.70
$E_{\text{ind,tot}}^{(2)}$	-5.16	-3.88	-0.95
$E_{\text{disp,pol}}^{(2)}$	-10.02	-12.49	-29.03
$E_{\text{disp,exch}}^{(2)}$	1.68	1.76	2.89
$E_{\text{disp,tot}}^{(2)}$	-8.34	-10.73	-26.14
$\Delta E_{\text{int}}^{(3-\infty)}$	-2.82	-3.24	+0.34
E_{int}	-21.26	-12.13	-7.39
Supermolecular:			
$E_{\text{int}}[\text{MP}(2)]$	-20.34	-12.43	-15.23 ^a
$E_{\text{int}}[\text{CCSD}(T)]$	-20.45	-11.21	-7.57 ^a

^aFrom Sinnokrot et al. (2002). These energies are close to the basis set limit

dispersion energy. Third- and higher-order terms are generally small for van der Waals complexes, and in this particular case, they are positive.

- *Intermediate case:* As an example of an intermediate case consider the so-called halogen-bonded complex $\text{Cl}_2 \cdots \text{OH}_2$. The exchange-repulsion energy of this complex is similar to that of the water dimer, but the electrostatic and induction energies are not as strong. The dispersion contributes a substantial fraction of the interaction energy, but less than for the benzene dimer. Third- and higher-order terms are quite large for this particular system.

By comparison, the supermolecular methods provide just the interaction energy: a single number. But perturbation theory does not provide only physical interpretation; the breakup of the interaction energy proves invaluable when constructing model interaction potentials (Stone and Misquitta 2007) and semi-empirical methods, some of which we shall see later.

Practical Methods for the Two-Body Interaction Energy

Supermolecular Methods

The two-body interaction energy can be calculated directly through Eq. 4 as the difference in energies of the dimer and the monomers. This approach, known as the *supermolecular method*, has the advantage of allowing E_{int} to be calculated

using a variety of electronic structure methods. We will only briefly describe some aspects of this method below; for a more complete description see Chalasinski and Szczesniak (2000).

Two of the most commonly used electronic structure methods to evaluate E_{int} are MP2 (second-order Moller–Plesset perturbation theory) and CCSD(T) (coupled-cluster with singles and doubles and non-iterated triples). Interaction energies calculated using these methods are shown in Table 1. Interaction energies calculated using MP2 are usually close to those calculated using the more sophisticated and computationally expensive CCSD(T), but the exception is the benzene dimer system for which MP2 overestimates the binding by almost a factor of 2. This deserves some explanation.

Moller–Plesset perturbation theory starts from a Hartree–Fock description of the system and builds in electron correlation perturbatively. In MP2, electron correlation is described using second-order perturbation theory. For a pair of interacting molecules the two orders of perturbation theory are used to describe the dispersion interaction, and therefore cannot be used to simultaneously describe electron correlation *within* each of the interacting molecules. This seems to cause severe problems for systems with π -bonding for which MP2 is completely inappropriate. Recently, this problem has been corrected using a hybrid scheme in which the dispersion component of the MP2 interaction energy is replaced with the dispersion energy calculated using SAPT(DFT) (see below) (Pitonak and Hesselmann 2010).

The CCSD(T) method takes a very different approach to electron correlation and is today regarded as the reference method for many kinds of problems. However, the computational cost of CCSD(T) scales as the seventh power of the system size, and therefore this method cannot be applied to large systems.

One of the sources of error in a supermolecular calculation of the interaction energy is the basis set superposition error (BSSE). This is the artificially enhanced binding that occurs because of the use of finite basis sets. This problem has been discussed at length in the literature, so we will not get into details here. A simple way of correcting for this error was suggested by Boys and Bernardi (1970) who suggested that the BSSE could be reduced by using the dimer basis to calculate the monomer energies. This leads to the *counterpoise* (CP) corrected interaction energy defined as

$$E_{\text{int}}^{\text{CP}} = E_{AB}(AB) - E_A(AB) - E_B(AB). \quad (17)$$

Here $E_A(AB)$ is the energy of A calculated in the dimer basis AB, and likewise for $E_B(AB)$. The CP form of the interaction energy should always be used, except possibly when the basis set is too poor. However, in this case, neither the CP corrected nor the uncorrected energies may be meaningful.

There is another aspect of the BSSE worth keeping in mind: An energy minimum on the CP corrected surface will generally occur at a different geometry from the corresponding minimum on the uncorrected surface. Since geometry optimizations are typically performed on the latter surface, the resulting structure will need to

be re-optimized on a CP corrected surface, but this is often rather difficult and computationally expensive.

Density Functional Theory

Density functional theory (DFT) is not presently suitable for intermolecular interactions (Tsuzuki and Lüthi 2001; van Mourik and Gdanitz 2002). The main reason for this failure of DFT is the highly non-local nature of the dispersion interaction which is present even when charge densities do not overlap. Since most density functionals are local or semi-local in the density, that is, they depend on the electron density and gradients of the density, they are unable to account for non-local correlations, and therefore cannot describe the dispersion energy.

Recently, there have been many attempts made to include the dispersion in DFT, either *explicitly* through a dispersion correction (Grimme 2004; Neumann and Perrin 2005) or *implicitly* through a non-local functional (Dion et al. 2004). The former method is necessarily empirical and makes a priori assumptions about the nature of the dispersion interaction. Though the latter methods are promising, it is still unclear if they are accurate enough in practice. Also recently, Becke and Johnson (2005) have proposed that the dispersion energy could be included in DFT through correlations in the exchange-hole dipole moments evaluated using the occupied orbitals. This method is promising and has been given a theoretical justification by Angyan (2007) and Hesselmann (2009).

In a recent development, Pernal et al. (2009) have used SAPT(DFT) (see below) interaction energies for a number of dimers to develop an exchange-correlation functional that does not include the dispersion energy. They have subsequently added the dispersion energy calculated from SAPT(DFT). This is a significant development as it is far better founded on theoretical grounds than other dispersion-corrected DFT methods. In practice a more efficient, but less accurate method could be to use dispersion models such as the WSM models described below.

Perturbation Theory

Symmetry-Adapted Perturbation Theory

Until recently, and possibly still so for certain applications, one of the most successful perturbation theories for intermolecular interactions was the symmetry-adapted perturbation theory (SAPT) (Jeziorski and Szalewicz 2002; Jeziorski et al. 1994). SAPT is a triple perturbation theory based on a Hartree–Fock description of the interacting monomers. Intramolecular correlation is built up using two of the perturbations, and intermolecular correlation by the third. SAPT has been applied to a large number of systems with very good success. See for example, applications to the water dimer (Mas et al. 1997), the carbon dioxide dimer (Bukowski et al. 1999), the He–HCCCN dimer (Akin-Ojo et al. 2003), Ar–CO₂ dimer (Misquitta et al. 2000), and the Ne–HCN dimer (Murdachaew et al. 2001). However, SAPT, like the CCSD(T) method, is computationally expensive with scaling of $\mathcal{O}N^7$, where N is the size of the system (which can be taken to be the number of electrons), and

quickly become impossible to apply as this size increases, and cannot be applied to systems of more than about ten atoms without a significant reduction to the level of correlation.

Symmetry-Adapted Perturbation Theory Based on DFT

Fairly recently, Szalewicz, Misquitta, and Jeziorski (Misquitta and Szalewicz 2002, 2005; Misquitta et al. 2003, 2005b), and independently, Jansen and Hesselmann (Hesselmann and Jansen 2002a, b, 2003a, b) developed a variant of SAPT based on DFT that has largely superseded SAPT both in terms of accuracy as well as computational efficiency. In this method, termed SAPT(DFT) or DFT-SAPT, the interaction energy components are formulated to make their dependence on molecular properties like the charge density and density response functions explicit. These molecular properties are then calculated using DFT and linear-response time-dependent DFT. In some sense, DFT is used to describe intramolecular correlation effects, thereby significantly simplifying the perturbation theory as the intramolecular perturbation operators are no longer required.

The first-order electrostatic energy $E_{\text{elst}}^{(1)}$ has already been shown to arise from the Coulomb interaction of the unperturbed total charge densities of the interacting monomers (Eq. 8). In SAPT(DFT), these densities are obtained from DFT. To see how the second-order induction and dispersion energies can be expressed in terms of molecular properties we need to define the frequency-dependent density susceptibility (FDDS):

$$\alpha_X(\mathbf{r}, \mathbf{r}' | \omega) = 2 \sum_{r \neq 0} \frac{E_r^X - E_0^X}{(E_r^X - E_0^X)^2 - \omega^2} \langle \Phi_0^X | \hat{\rho}_X(\mathbf{r}) | \Phi_r^X \rangle \langle \Phi_r^X | \hat{\rho}_X(\mathbf{r}') | \Phi_0^X \rangle, \quad (18)$$

where ω is a frequency and $\hat{\rho}_X(\mathbf{r}) = -\sum_{i \in X} \delta(\mathbf{r} - \mathbf{r}_i)$ is the electron density operator. The FDDS describes the linear response of the electron density to a frequency-dependent perturbation.

$E_{\text{ind,pol}}^{(2)}(A)$ can be written in terms of the FDDS evaluated at zero frequency (Magnasco and McWeeny 1991):

$$E_{\text{ind,pol}}^{(2)}(A) = -\frac{1}{2} \iint V_B(\mathbf{r}) \alpha_A(\mathbf{r}, \mathbf{r}' | 0) V_B(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \quad (19)$$

where $V_B(\mathbf{r}) = \int \rho_B^{\text{tot}}(\mathbf{r}') / |\mathbf{r} - \mathbf{r}'| d\mathbf{r}'$ is the electrostatic potential of monomer B, and similarly for $E_{\text{ind,pol}}^{(2)}(B)$. We therefore see that the induction energy is a *response* energy of the molecule to the potential arising from the interacting partner.

The dispersion energy involves the FDDSs of the two molecules evaluated at imaginary frequency (Longuet-Higgins 1965; Zaremba and Kohn 1976):

$$E_{\text{disp,tot}}^{(2)} = -\frac{1}{2\pi} \int_0^\infty d\omega \int d\mathbf{r}_1 d\mathbf{r}'_1 d\mathbf{r}_2 d\mathbf{r}'_2 \frac{\alpha_A(\mathbf{r}_1, \mathbf{r}'_1; i\omega) \alpha_B(\mathbf{r}_2, \mathbf{r}'_2; i\omega)}{|\mathbf{r}_1 - \mathbf{r}_2| |\mathbf{r}'_1 - \mathbf{r}'_2|}. \quad (20)$$

There is no easy way to interpret this expression which is a coupling of the response functions of the two molecules, but at imaginary frequency. However, see the discussion in Angyan (2007) and Hesselmann (2009) where an approximate form of this expression is related to correlations in fluctuating exchange-holes of the monomers.

Linear-response time-dependent DFT, also known as coupled Kohn–Sham (CKS), provides us with a convenient framework for evaluating the FDDS. CKS theory has been used for some time now to obtain excitation energies of small systems (Grabo et al. 2000; Onida et al. 2002) and molecular properties like polarizabilities and hyperpolarizabilities (Adamo et al. 1999; Tozer and Handy 1998) (see Koch 2000 for a comprehensive summary), so it is not surprising that it can also be used to evaluate the second-order dispersion and induction energies through the FDDS (Hesselmann and Jansen 2003b; Misquitta et al. 2003). Within CKS theory (and also coupled Hartree–Fock (CHF) theory), the FDDS takes the form (Casida 1995; Colwell et al. 1995; Petersilka et al. 1996):

$$\alpha(r, r' | \omega) = \sum_{i v, i' v'} C_{i v, i' v'}(\omega) \phi_i(\mathbf{r}) \phi_v(\mathbf{r}) \phi_{i'}(\mathbf{r}') \phi_{v'}(\mathbf{r}'), \quad (21)$$

where the subscripts i and i' (v and v') denote occupied (virtual) molecular orbitals and ϕ_i is a molecular orbital obtained from Kohn–Sham DFT. The coupling between the molecular orbitals is contained in the coefficients $C_{i v, i' v'}(\omega)$ which are defined in Colwell et al. (1995), Ioannou et al. (1997), Misquitta et al. (2005b), and Podeszwa et al. (2006b). These coefficients involve Coulomb integrals of the molecular orbitals as well as integrals containing the functional derivative of the exchange–correlation potential. The method used to evaluate the coefficients $C_{i v, i' v'}(\omega)$ is quite important as the dispersion energy calculated through Eq. 20 varies appreciably depending on the type of FDDS (Misquitta et al. 2005b).

The exchange energies depend explicitly on molecular wavefunctions through the interaction density matrix (Jeziorski et al. 1994; Moszyński et al. 1994). In SAPT(DFT), these energies are evaluated using Kohn–Sham molecular orbitals and, for the second-order exchange energies, the effect of orbital relaxation is estimated by scaling:

$$E_{\text{ind,exch}}^{(2)} = E_{\text{ind,exch}}^{(2)}(\text{KS}) \times \frac{E_{\text{ind,pol}}^{(2)}}{E_{\text{ind,pol}}^{(2)}(\text{KS})} \quad (22)$$

$$E_{\text{disp,exch}}^{(2)} = E_{\text{disp,exch}}^{(2)}(\text{KS}) \times \frac{E_{\text{disp,pol}}^{(2)}}{E_{\text{disp,pol}}^{(2)}(\text{KS})} \quad (23)$$

where the energies calculated without orbital relaxation are indicated by “KS,” that is, they use un-relaxed Kohn–Sham orbitals. The validity of using Kohn–Sham orbitals for the first-order exchange-repulsion energy has been justified by an asymptotic analysis of the exact and Kohn–Sham interaction density matrices (Misquitta and Szalewicz 2002) as well as extensive numerical tests (Misquitta and Szalewicz 2005). In the DFT-SAPT formulation, the second-order exchange energies are defined via coupled-response theory (Hesselmann et al. 2005).

Exchange-correlation functional: Which exchange-correlation functional should be used in SAPT(DFT) calculations? The short answer is that while SAPT(DFT) interaction energies are not very sensitive to the exchange-correlation functional used, it has been concluded from extensive numerical experiments that the asymptotically corrected (Tozer 2000; Tozer and Handy 1998) PBE0 (Adamo and Barone 1999) exchange-correlation functional results in the most accurate interaction energies for a variety of systems. For details see Hesselmann and Jansen (2002a), Misquitta and Szalewicz (2005), and Misquitta et al. (2005b). The PBE0 functional (also called PBE1PBE) is a hybrid functional with 75 % of its exchange energy determined from PBE and 25 % from the so-called exact, or Hartree–Fock exchange. This means that the FDDS must also be constructed as the hybrid of the FDDSs from CKS and CHF theories. This is quite important as significant errors in the dispersion energy are introduced if the FDDS is constructed using CKS theory alone (Misquitta et al. 2005b). For large molecules, however, the terms in the coefficients $C_{i_v, i'v'}(\omega)$ that depend on the functional derivative of the exchange-correlation potential are computationally demanding to evaluate using the PBE functional. Instead, a more practical approach is to calculate these coefficients using the exchange-only LDA functional. Therefore, while the Kohn–Sham molecular orbitals and eigenvalues are obtained using the PBE0 functional, the FDDS is best constructed using the less accurate LDA + CHF kernel. This approximation results in a small (less than 1 %) loss in accuracy which is more than compensated by an order of magnitude reduction in computational expense (Misquitta et al. 2005b).

Asymptotic correction: The asymptotic correction is needed to correct the tails of the exchange-correlation potential, and consequently, the density tails that are crucial for intermolecular interactions. In order to implement the asymptotic correction, accurate vertical ionization potentials (IPs) are needed for the monomers. When they are not available experimentally, good estimates may be obtained from the difference between the energies of the N and $N - 1$ electron systems. The PBE0 functional is best suited for this calculation too as tests on atoms, diatoms, and small organic molecules have shown that it gives IPs with mean errors centered about 0.0 a.u. with a standard deviation of only 0.007 a.u. (Ernzerhof and Scuseria 1999). This correction is definitely needed if the individual interaction energy components are required to be accurate. However, it is cumbersome to apply for large systems, when a single IP may be questionable and local IPs are hard to define. In such cases, the asymptotic correction is best neglected, particularly if total interactions are all that is desired.

Density-Fitting

The SAPT(DFT) interaction energy components are not evaluated directly through the expressions provided in the previous section, but are formulated to make use of a procedure known as density-fitting (DF) or the resolution of the identity (RI) (Dunlap 2000; Dunlap et al. 1979). The idea here is to expand pairs of molecular orbitals that appear in the energy expressions in an auxiliary basis. For example, we may perform the DF expansion of one such pair that appears in the FDDS defined through Eq. 21 as follows:

$$\rho_{iv}(\mathbf{r}) = \phi_i(\mathbf{r})\phi_v(\mathbf{r}) \approx \tilde{\rho}_{iv}(\mathbf{r}) = \sum_k D_{iv,k} \chi_k(\mathbf{r}), \quad (24)$$

where χ_k are the auxiliary basis functions and $D_{iv,k}$ are the expansion coefficients. In the standard density-fitting procedure (Dunlap et al. 1979), the fitted density $\tilde{\rho}$ is found by minimizing the integrals

$$\Delta_{iv} = \int \int (\tilde{\rho}_{iv}(\mathbf{r}) - \rho_{iv}(\mathbf{r})) \frac{1}{|\mathbf{r} - \mathbf{r}'|} (\tilde{\rho}_{iv}(\mathbf{r}') - \rho_{iv}(\mathbf{r}')) d\mathbf{r} d\mathbf{r}'. \quad (25)$$

With these expansions, the FDDS can be rewritten as (Misquitta et al. 2003)

$$\alpha(\mathbf{r}, \mathbf{r}' | \omega) = \sum_{k,l} \tilde{C}_{kl}(\omega) \chi_p(\mathbf{r}) \chi_q(\mathbf{r}'), \quad (26)$$

where $\tilde{C}_{kl}(\omega)$ are the transformed coefficients given by $\tilde{C}_{kl}(\omega) = \sum_{i_v, i'v'} D_{iv,k} C_{i_v, i'v'}(\omega) D_{i'v', l}$. Now the FDDS defined through Eq. 21 involves a quadruple sum over molecular orbitals, but in Eq. 26 the sum is over only two indices. This represents a huge reduction in computational expense provided that the size of the density-fitting basis – commonly called the *auxiliary basis* – is of the same order as the size of the main basis (the basis used to obtain the molecular orbitals), which is indeed the case with optimized auxiliary basis sets.

With the density-fitted formulation of SAPT(DFT), commonly called DF-SAPT(DFT) or DF-DFT-SAPT, the computational cost of the second-order interaction energy scales as only $\mathcal{O}(N^5)$ if hybrid functionals are used (Bukowski et al. 2005; Podeszwa et al. 2006b). In the DFT-SAPT implementation the FDDS is calculated without a contribution from CHF theory. This reduces the computational scaling for evaluating the FDDS to $\mathcal{O}(N^4)$, but the overall scaling is still $\mathcal{O}(N^5)$ because of the 4-index two-electron integrals (Hesselmann et al. 2005). This is the same computational scaling as MP2, though DF-SAPT(DFT) calculations take a few times longer due to the additional complexity of the method. Compare this to the $\mathcal{O}(N^7)$ scaling of SAPT and CCSD(T).

Higher-Order Contributions

Contributions of third and higher order to the two-body interaction energy are often large and cannot be neglected. In particular, for systems of polarizable molecules with large permanent multipole moments terms of third and higher order can contribute as much as 30 % of the binding energy at equilibrium geometries. Such systems tend to have a large $E_{\text{ind,pol}}^{(2)}$. Examples are the water dimer and the $\text{Cl}_2 \cdots \text{OH}_2$ systems in Table 1. For dispersion-bound systems the higher-order contributions are smaller and constitute only as much as 5 % of the total interaction energy (see the benzene dimer in Table 1) and can therefore be neglected if high accuracies are not required.

The most common method for approximating the terms of third and higher-order, $E_{\text{int}}^{(3-\infty)}$, is the $\delta_{\text{int,resp}}^{\text{HF}}$ correction (Jeziorska et al. 1987; Moszynski et al. 1996):

$$E_{\text{int}}^{(3-\infty)} \approx \delta_{\text{int,resp}}^{\text{HF}} = E_{\text{int}}^{\text{HF}} - \left(E_{\text{elst}}^{(1)}(\text{HF}) + E_{\text{exch}}^{(1)}(\text{HF}) + E_{\text{ind,tot}}^{(2)}(\text{CHF}) \right), \quad (27)$$

where the $E_{\text{int}}^{\text{HF}}$ is the counterpoise-corrected supermolecular interaction energy calculated using Hartree–Fock theory, and the interaction energies are calculated using Hartree–Fock orbitals (denoted by “HF”) or, for the second-order induction energy, using CHF theory. This approximation was used to evaluate $E_{\text{int}}^{(3-\infty)}$ for the water dimer and $\text{Cl}_2 \cdots \text{OH}_2$ systems (Table 1) for which this term contributes 13 % and 27 % of the total interaction energy.

There is evidence (Patkowski et al. 2006) that the $\delta_{\text{int,resp}}^{\text{HF}}$ estimate is not appropriate for dispersion-bound systems, for which a better estimate of $E_{\text{int}}^{(3-\infty)}$ is (Misquitta and Stone 2008a)

$$E_{\text{int}}^{(3-\infty)} \approx E_{\text{ind,tot}}^{(3)}(\text{KS}) \quad (28)$$

where $E_{\text{ind,tot}}^{(3)}(\text{KS})$ is the total third-order induction energy evaluated without relaxation, that is, using uncoupled Kohn–Sham theory. This estimate has been used for the benzene dimer in Table 1.

For large systems it is quite conceivable that no clear distinction may be possible that would justify the use of either one of these approximations. In such cases, it would probably be better to use the $\delta_{\text{int,resp}}^{\text{HF}}$ approximation rather than ignore the higher-order energies altogether.

Basis Sets: Charge-Transfer

In general, the basis sets used to calculate intermolecular interaction energies need to be large and augmented with diffuse functions. In part, this is because of the second-order energies which are response energies. Consequently, the basis set has to be good enough to describe not only the charge density accurately (particularly in the region of the density tail), but also the *response* of the density to perturbations. These requirements are adequately met only by basis sets of the augmented triple- ζ

kind (and larger), such as aug-cc-pVTZ (Kendall et al. 1992; Woon and Dunning 1994) and Sadlej-pVTZ (Sadlej 1992; Sadlej and Urban 1991).

However, when the overlap of the charge densities of the interacting monomers becomes significant (typically at around the equilibrium separation) the basis set must be flexible enough to describe the intermolecular electron–electron cusp as well as the intermolecular charge-transfer (CT). The former effect is manifested as part of $E_{\text{disp,tot}}^{(2)}$ and the latter as part of $E_{\text{ind,tot}}^{(2)}$.

It is often quite difficult to converge $E_{\text{disp,tot}}^{(2)}$ with basis set. This is because it is hard to describe the intermolecular electron–electron cusp in the bonding region using basis functions located on atomic sites only (Burcl et al. 1995). In this case, fairly high-angular-momenta functions are needed in order to do so. However, this leads to very large basis sets and a consequent increase in computational requirements. An alternative is to use basis sets augmented with the so-called *mid-bond* functions: a small set of basis functions located in the bonding region (see Burcl et al. 1995; Williams et al. 1995 and references therein). The dispersion energy is not so sensitive to the exact composition of the mid-bond basis, which is usually chosen to consist of a set of 3 *s*, 2 *p*, and 1 *d* diffuse functions (Mas et al. 1997). A convenient choice for the location of the mid-bond set has been given in Akin-Ojo et al. (2003).

An alternative to mid-bond functions is to use the basis set extrapolation scheme of Helgaker et al. (1997). For intermolecular interactions, this involves calculating $E_{\text{disp,tot}}^{(2)}$ using two correlation-consistent basis sets, say the aug-cc-pV *n*Z and the aug-cc-pV(*n* + 1)Z Dunning bases. This energy can be fitted to the form $a + b/X^3$ where $X = n$ and $n + 1$ and the constants determined. The complete basis set (CBS) estimate of $E_{\text{disp,tot}}^{(2)}$ is obtained by extrapolating $X \rightarrow \infty$, that is, it is the constant *a*. This scheme has the advantage that it applies equally well to small as well as large molecules, for which the mid-bond scheme is potentially ambiguous. However, two electronic structure calculations need to be performed, so there is an increase in computational cost.

Charge transfer poses a very different problem. This subject is discussed more completely in Stone and Misquitta (2009) and Stone (1996). The CT energy is the part of the short-range induction energy that involves excitations from the occupied orbitals on one molecule into the virtual orbitals of another. Consequently, in Stone (1993) it was suggested that the CT energy could be calculated as the difference of $E_{\text{ind,tot}}^{(2)}$ calculated in the dimer and monomer basis. The understanding here is that the monomer basis is localized, so excitations into the virtual space of the interacting partner should be suppressed. In any case, if CT excitations are to be correctly included, the induction energy should be calculated using the dimer basis. In practice, this is rather excessive, and it has been shown (Williams et al. 1995) that only a subset of the basis of the interacting partner need be included. This is typically taken to be just the functions of *s* and *p* symmetry. Thus, when calculating the induction energy of A in the presence of B, that is, $E_{\text{ind,tot}}^{(2)}(A)$, we use in addition to the basis set of A, the *s* and *p* symmetry functions of B located where the atomic sites of B would be. These ghost functions are known as the *far-bond* functions.

The combined basis consisting of functions on the molecule and the mid-bond and far-bond functions is referred to as the *monomer-centered plus* or “MC+” basis type. The “+” indicating the presence of the additional basis functions. A typical SAPT(DFT) calculation of the interaction energy would involve two such bases; one for each of the two interacting molecules.

The additional functions are required for the second-order energies. If all that is needed are the first-order energies, the MC basis can be used, that is, basis functions need be included only on the monomer. First-order electrostatic and exchange-repulsion energies calculated in such a basis are generally very close to those calculated in the larger MC+ type of basis if a triple- ζ -quality basis is used.

Finally, if each molecule is described using the dimer basis we obtain the DC or DC+ basis types; the latter additionally including the mid-bond set. These basis types must be used for supermolecular calculations of the interaction energy.

Auxiliary Basis Sets

One downside of density-fitting is that the auxiliary basis used needs to be paired with the main basis. Nevertheless, optimized auxiliary basis sets are available (Weigend et al. 1998, 2002) for most of the commonly used main basis sets. These tend to be between two and three times larger than the main basis.

A few points should be noted when using density-fitting for intermolecular forces:

- The auxiliary basis used needs to be of the DC or DC+ type even if the main basis is of the MC or MC+ type. It is not clear why this is the case, but a failure to do so sometimes results in unacceptable errors in $E_{\text{elst}}^{(1)}$ and $E_{\text{ind,tot}}^{(2)}$.
- The Sadlej-pVTZ basis does not have an optimized auxiliary basis associated with it, but experience suggests that the aug-cc-pVTZ auxiliary basis may be used.
- When using the Dunning aug-cc-pV nZ basis sets, higher accuracies may be obtained by using auxiliary basis sets optimized for the aug-cc-pV($n+1$)Z bases.

Multipole Expansion for the Interaction Energy

One of the major strengths of perturbation theory is that each of the interaction energy components possesses a multipole expansion that allows us to evaluate the interaction energy analytically in terms of molecular properties alone, at least when the molecular charge densities do not overlap appreciably.

The intermolecular interaction operator V defined through Eq. 5 can be written as

$$V = \sum_{\alpha \in A} \sum_{\beta \in B} \frac{q_{\alpha} q_{\beta}}{r_{\alpha\beta}}, \quad (29)$$

where $\alpha(\beta)$ label all the particles (electrons and nuclei) of monomers A (B), q_{α} is the charge (in atomic units) of particle α , and $r_{\alpha\beta}$ is the inter-particle distance. If

A and B are well separated so that their charges are distinct and do not overlap (strictly this is never exactly possible), we can expand the interaction operator in the multipole expansion to give (in the compact notation of Stone 1996):

$$V = \sum_{t,u} \widehat{Q}_t^A T_{tu}^{AB} \widehat{Q}_u^B. \quad (30)$$

There is a lot compressed in this expression so it is well that we spend some time unravelling it. First of all, the subscripts t and u label the angular momenta of the real spherical harmonics and take the lm values 00, 10, 11 *c*, 11 *s*, 20, 21 *c*, 21 *s*, 22 *c*, 22 *s*, ... (the labels “*c*” and “*s*” stand for cosine and sine respectively). \widehat{Q}_t^A is the real form of the multipole moment operator of rank t centered on A and expressed in the local-axis system of A, while T_{tu}^{AB} is the so-called T -tensor that carries the distance and angular dependence. The T -function of ranks l_1 and l_2 has a distance dependence $R^{-l_1-l_2-1}$ where R is the separation between the centers of A and B.

Using this form of the interaction energy operator, the first-order electrostatic energy can be written as

$$E_{\text{elst}}^{(1)}(\text{MP}) = \sum_{t,u} Q_t^A T_{tu}^{AB} Q_u^B, \quad (31)$$

where Q_t^A is now the rank t multipole moment of the total charge density of A. This expression, which follows directly from Eqs. 8 and 30, is convenient as multipole moments are relatively easy to calculate. Therefore, Eq. 31 provides us an easy method for calculating the first-order electrostatic energy. But consider the following points:

- Contributions to the electrostatic energy that decay exponentially with separation R do not contribute to the multipole expansion. Such terms arise when the charge densities overlap and are therefore referred to as the *penetration* contributions to the interaction energy.
- The non-expanded electrostatic energy (Eq. 8) exhibits a R^{-1} divergence when the nuclei of the interacting molecules approach each other, but the multipole expansion introduces a higher-order divergence because it includes terms of the form R^{-n} with $n > 1$. These terms need to be *damped* out at short range. This is usually done using a *damping function* $f_n(R)$ that has the property that $f_n(R) \rightarrow R^n$ for small R and $f_n(R) \rightarrow 1$ at large R . We will discuss specific forms of these functions later.

Therefore, we can write the electrostatic energy as

$$E_{\text{elst}}^{(1)} = Ae^{-\alpha R} + \sum_{t,u} Q_t^A f_{(tu)}(R) T_{tu}^{AB} Q_u^B, \quad (32)$$

where the first term is the exponential penetration contribution (that is usually negative) and the second is the damped multipole expansion of the electrostatic energy. In general, the coefficients A and α in the penetration term will depend on orientation and possibly on distance. The exact form is not critical, what is important is that this term exists at short separations and that it can be large in magnitude.

This analysis also applies to the multipole expansions for the second-order induction and dispersion energies. But before getting to these, let us ask an important question: When is the single-center multipole expansion valid? To answer this question fully would take us beyond the scope of this chapter. Instead we will introduce the main ideas here without proof, and refer the reader to the detailed discussion in Chap. 7 of Stone (1996).

A key concept is the idea of the sphere of divergence. This is defined as the sphere centered on the expansion center (typically the center-of-mass) and just enclosing all charges. For a molecule this poses a problem as the electronic density formally extends to infinity. But it has been shown (Stone and Alderton 1985) that for charge densities expressed as the sum of Gaussian functions, it is sufficient that the sphere enclose all nuclear sites. The multipole expansion is valid only if the spheres of divergence of all interacting molecules *do not overlap*.

This poses a problem for all but the smallest of molecules. For example, consider the benzene dimer in the sandwich configuration used in Table 1. The equilibrium separation is about 3.7 Å, but the radius of a benzene molecule is about 2.4 Å, so the molecules would have to be at least 4.8 Å apart before the single-center multipole expansion could be used. And even then, the expansion would be only slowly convergent.

A solution to this problem is to use a multi-centered multipole expansion, more commonly called a *distributed multipole expansion*. In principle the centers could be arbitrarily chosen, but it is convenient to use the atomic nuclei as centers. The distributed multipole expansion of the intermolecular interaction operator V is

$$V = \sum_{a \in A} \sum_{b \in B} \sum_{tu} \widehat{Q}_t^a T_{tu}^{ab} \widehat{Q}_u^b, \quad (33)$$

which is similar to Eq. 30 except that now \widehat{Q}_t^a is the rank t multipole moment operator for site a . We will define the precise form of these multipole moment operators later.

We are now in a position to define the distributed multipole expansions for $E_{\text{elst}}^{(1)}$, $E_{\text{ind,tot}}^{(2)}$, and $E_{\text{disp,tot}}^{(2)}$. Recall that the exchange energies decay exponentially with molecular separation and have no expansion in $1/R$. Consequently, the multipole expansion for $E_{\text{ind,tot}}^{(2)}$ is the same as that for $E_{\text{ind,pol}}^{(2)}$ and will be denoted simply by $E_{\text{ind}}^{(2)}(\text{MP})$. Likewise for $E_{\text{disp}}^{(2)}(\text{MP})$. The damping functions have been omitted from the expansions given below. To insert them is straightforward: for every T_{tu}^{ab} include a damping function of the form $f_{(tu)}(r_{ab})$.

- Electrostatic energy: The generalization of the single-center expansion is straightforward and we get

$$E_{\text{elst}}^{(1)}(\text{MP}) = \sum_{a \in A} \sum_{b \in B} \sum_{tu} Q_t^a T_{tu}^{ab} Q_u^b. \quad (34)$$

- Second-order induction: The second-order induction energy for molecule A is given by

$$E_{\text{ind}}^{(2)}(\text{MP}) = (A) = -\frac{1}{2} \sum_{a, a' \in A} \sum_{b, b' \in B} \sum_{tu t' u'} Q_u^b T_{tu}^{ab} \alpha_{tt'}^{aa'} T_{t'u'}^{a'b'} Q_{u'}^{b'}, \quad (35)$$

where $\alpha_{tt'}^{aa'}$ is the static polarizability of rank t by t' for sites a and a' and is defined by

$$\alpha_{tt'}^{aa'} = \iint \widehat{Q}_t^a(\mathbf{r}) \alpha(\mathbf{r}, \mathbf{r}' | 0) \widehat{Q}_{t'}^{a'}(\mathbf{r}') d\mathbf{r} d\mathbf{r}'. \quad (36)$$

Likewise, we define $E_{\text{ind}}^{(2)}(\text{MP}) (B)$.

From Eq. 35 we see that the second-order induction energy of A is a *linear response* to permanent multipoles of B.

- Second-order dispersion: The second-order dispersion energy has the distributed multipole expansion

$$E_{\text{disp}}^{(2)}(\text{MP}) = -\frac{1}{2\pi} \sum_{a, a' \in A} \sum_{b, b' \in B} \sum_{tu t' u'} T_{tu}^{ab} T_{t'u'}^{a'b'} \int_0^\infty \alpha_{tt'}^{aa'}(i\omega) \alpha_{uu'}^{bb'}(i\omega) d\omega, \quad (37)$$

where $\alpha_{tt'}^{aa'}(i\omega)$ is the frequency-dependent distributed polarizability defined analogously to Eq. 36 and evaluated at imaginary frequency. This form of the asymptotic dispersion energy involves a quadrupole sum over sites and is too cumbersome for most applications. Instead, very often a simplification is made by *localizing* the distributed polarizabilities to obtain a polarizability description that contains terms that contain one site index only. With this simplification, the dispersion expansion involves a double sum over sites and can be cast into the more familiar site–site form (Stone and Tough 1984):

$$\begin{aligned} E_{\text{disp}}^{(2)}(\text{MP}) &= -\frac{1}{2\pi} \sum_{a \in A} \sum_{b \in B} \sum_{tu t' u'} T_{tu}^{ab} T_{t'u'}^{ab} \int_0^\infty \alpha_{tt'}^a(i\omega) \alpha_{uu'}^b(i\omega) d\omega \\ &= -\sum_{a \in A} \sum_{b \in B} \left(\frac{C_6^{ab}}{R_{ab}^6} + \frac{C_7^{ab}}{R_{ab}^7} + \frac{C_8^{ab}}{R_{ab}^8} + \dots \right). \end{aligned} \quad (38)$$

In general, the dispersion coefficients are angle dependent (Stone and Tough 1984), but for the special case of spherical sites (an idealization), they are independent of angle and only terms of even order survive.

Damping Functions

As explained above, the divergence of the multipole expansions of the interaction energies must be damped out using *damping functions*. The exact form of the damping functions are not known, particularly for the induction energy, but experience from high-accuracy calculations on small dimers suggests that the damping functions from Tang and Toennies (1992), which are incomplete gamma functions, are probably the most suitable. The order- n Tang–Toennies damping function takes the form

$$f_n(\beta R_{ab}) = 1 - \exp(-\beta R_{ab}) \sum_{k=0}^n \frac{(\beta R_{ab})^k}{k!} \quad (39)$$

where β is the damping constant which may be angle dependent, but is often assumed to be a simple constant dependent on the pair of sites (a , b). Often, even the site dependence is dropped and β is assumed to depend on the molecular types alone.

In general, the damping constant β will need to be determined by comparison with the non-expanded energies [from SAPT(DFT)]. For the induction energy, since the multipole expansion does not include the charge transfer energies, the comparison should be made with $E_{\text{ind,tot}}^{(2)}$ calculated in the MC basis. This differs from the recommendation in Misquitta and Stone (2008a) where it was suggested that the damping constant be derived from the ionization energies of the interacting molecules as

$$\beta = \sqrt{2I_A} + \sqrt{2I_B}, \quad (40)$$

where I_A and I_B are the molecular vertical ionization energies in atomic units. There is now evidence (Sebetci and Beran 2010) to suggest that this results in too small a damping when used to evaluate the many-body induction energy through the polarization expansion, potentially resulting in a divergence of the expansion. However, there is extensive evidence that this value of the damping constant is appropriate for the dispersion energy (Misquitta and Stone 2008b) and results in accurate dispersion energies even at molecular charge-density overlap.

Many-Body Contributions to the Interaction Energy

In the condensed phase and in clusters, the two-body energy is often a poor approximation to the total interaction energy which can contain a substantial contribution from the many-body energies (Eqs. 2 and 3). These terms contribute most strongly in systems of polarizable molecules with strong permanent multipole moments such

as those that exhibit strong hydrogen-bonding. For example, Milet et al. (1999) have shown that in clusters of water molecules, the three-body contribution to the interaction energy can be as much as 28 % at the equilibrium geometries and even more – as much as 50 % – at the barriers. Likewise, the many-body non-additive energies have been shown to be responsible for the structural properties of liquid water (Bukowski et al. 2006). Many-body effects are also present in dispersion-bound systems like the argon liquid (Bukowski and Szalewicz 2001) small trimers containing inert gases (Ernesti and Hutson 1997) and the benzene crystal (Podeszwa et al. 2008). In all cases the non-additive dispersion makes a small but non-negligible contribution to the total interaction energy.

The three-body correction (Eq. 3) is the dominant source of non-additivity (Bukowski et al. 2006; Mas et al. 2003a, b). This correction may be calculated using the supermolecular approach, but the calculation is tedious and the BSSE needs to be corrected for. Nevertheless, this approach is possible for small clusters (Hodges et al. 1997). The most appropriate method for calculating the three-body correction is through a three-body version of SAPT(DFT) (Podeszwa and Szalewicz 2007). Using representative trimers of benzene, Podeszwa et al. (2008) have estimated the three-body SAPT(DFT) contribution to the lattice energy of the benzene crystal and have found the result to be in good agreement with experiment. However this cannot be done in general as the number of trimers in the condensed phase scales as the cube of the number of molecules. One solution is to construct an analytic representation of the three-body correction. Though this has been successfully done for water (Mas et al. 2003a, b), it is a formidable task, and is likely unreasonable to perform for most systems.

There are, however, alternative, computationally feasible procedures that are based on the (distributed) multipole forms of the induction and dispersion that are the dominant source of non-additivity in clusters (Bukowski and Szalewicz 2001; Mas et al. 2003b). The exchange non-additivity is significant in small clusters (Hodges et al. 1997; Milet et al. 1999), but it is relatively less important in large clusters and the condensed phase (Mas et al. 2003b).

From a detailed examination of the three-body SAPT energies for the argon liquid, Bukowski et al. (Bukowski and Szalewicz 2001) have concluded that the bulk of the three-body dispersion energy can be estimated using the Axilrod–Teller–Muto triple dipole term. For isotropic local distributed polarizabilities this takes the form (Stone 1996)

$$\Delta E_{\text{disp}}^{3\mu}(ABC) = \sum_{a \in A} \sum_{b \in B} \sum_{c \in C} C_9 \frac{(1 + 3 \cos \hat{a} \cos \hat{b} \cos \hat{c})}{R_{ab}^3 R_{bc}^3 R_{ac}^3}, \quad (41)$$

where R_{ab} , etc., are the lengths of the sides and \hat{a} , \hat{b} , and \hat{c} are the angles of the triangle formed by the three atoms, and

$$C_9 = \frac{3}{\pi} \int_0^{\infty} \alpha^a(i\omega) \alpha^b(i\omega) \alpha^c(i\omega) d\omega. \quad (42)$$

One of the most efficient and convenient ways of approximating the many-body induction energy is through a *polarizability model* based on distributed polarizabilities. The many-body induction energy of a cluster of molecules can be calculated using a generalization of Eq. 35 (Stone 1996):

$$E_{\text{ind}} (\text{Many} - \text{body}) = \frac{1}{2} \sum_A \sum_{a \in A} \sum_{B \neq A} \sum_{b \in B} \sum_{tu} \Delta Q_t^a T_{tu}^{ab} Q_u^b. \quad (43)$$

Here ΔQ_t^a is the *change* in multipole moment t at a due to the self-consistent polarization of site a in the field of all sites on other molecules and is given by

$$\Delta Q_t^a = - \sum_{a' \in A} \sum_{B \neq A} \sum_{b \in B} \sum_{t'v} \alpha_{tt'}^{aa'} T_{t'v}^{a'b} (Q_v^b + \Delta Q_v^b), \quad (44)$$

Notice that Eq. 44 must be solved iteratively for all molecules in the system, as the ΔQ occur on both sides of the equation. The cumulative effect of the iterations is small for a dimer, but can be substantial for a cluster. As before, each of the T -functions should be associated with a suitable damping function. Mas et al. (2003b) have validated this model for the many-body energies in liquid water.

Molecular Properties

In order to evaluate the distributed multipole forms of the interaction energy components we need to calculate distributed multipoles and distributed frequency-dependent polarizabilities of the interacting molecules.

There is no unique method for distributed molecular properties. Rather than get into a philosophical discussion of the concept of an *atom-in-a-molecule* (Matta and Bader 2006; Parr et al. 2005) (see also the discussion in Lillestolen and Wheatley 2008) we will adopt the more pragmatic approach and require that whatever the method used, the distributed properties be the most accurate possible within the constraints imposed by the model. Additionally, we also need the methods used to be numerically robust, applicable to small organic molecules, relatively easy to use, and theoretically consistent with the electronic structure method used to calculate the non-expanded energies, which, in our case, will be SAPT(DFT).

Distributed Multipoles

The distribution of the multipole moments has been the subject of many decades of research (see Stone 1996 and Stone and Misquitta 2007 for reviews). The Distributed Multipole Analysis (DMA) of Stone and Alderton (1985) is widely

used, and a recent modification (Stone 2005) overcomes a shortcoming of the earlier method that arose when diffuse functions were present in the basis.

A special case of distributed multipoles are the point charge models. Such models may be obtained in a variety of ways. In one of the most commonly used methods the charge model is required to reproduce (in a least-squares sense) the molecular electrostatic potential at a set of points around the molecule (Singh and Kollman 1984). These points are generally chosen to lie sufficiently far as to exclude or minimize the effects of penetration. Another possibility is to require that the charges best reproduce some of the molecular multipole moments.

The simplicity of the point charge models comes at a price. While such models may be adequate for systems with an insignificant contribution from the electrostatic energy, they are clearly inadequate for others, in particular, when hydrogen bonding is present. This is because hydrogen bonds are strongly directional (Buckingham and Fowler 1983; Buckingham et al. 2008), whereas the electrostatic interaction between two point charges is completely isotropic. Besides the energetic importance of the hydrogen bond, the strong directionality significantly reduces the available configuration space. This can have important consequences in simulations and crystal energy landscape searches (Day et al. 2005).

Basis set: The higher-ranking multipoles require large and more diffuse basis sets to be well described. Experience has shown that an augmented triple- ζ basis is often adequate.

Distributed Polarizabilities

The distributed polarizability is defined through Eq. 36. This definition suggests at least two methods for performing the distribution: one possibility is to require that the site multipole moment operator \widehat{Q}_t^a is non-zero only around site a , and likewise for $\widehat{Q}_{t'}^{a'}$. This *real-space* partitioning scheme has been implemented using integration grids (Le Sueur and Stone 1994) and Bader's theory of atoms in a molecule (Angyan et al. 1994). However, the shapes of the atomic domains in these schemes tend to be irregular, which could lead to artifacts in the higher-rank polarizabilities. Additionally, these methods tend to result in unphysically large charge-flow terms that are hard to localize. These terms are rank zero contributions to the polarizability (where one or both of the subscripts t and t' are 00 in Eq. 36) that describe the flow of charge along the molecule (Stone 1996). The effects of the charge-flow terms can often be described using higher-ranking polarizabilities, but in order to do so, the charge-flow terms must be small.

Yet another possibility is to partition the FDDS that appears in Eq. 36 into contributions from pairs of atoms. These are known as the *basis-space* partitioning methods. Early attempts to do so failed as the resulting polarizabilities were completely unphysical for large basis sets (Le Sueur and Stone 1993).

More recent methods (Gagliardi et al. 2004; Lillestolen and Wheatley 2007; Wheatley and Lillestolen 2008) have been more successful, but are either applicable

only to the static polarizability or are too cumbersome to use routinely. Also very recently, distribution schemes were proposed by Williams and Stone (2003) and Misquitta and Stone (2006). These two methods have been combined into the Williams–Stone–Misquitta (WSM) method which has proved to be one of the most successful methods for obtaining distributed polarizabilities. This is what we will describe next.

Williams–Stone–Misquitta (WSM) Distribution

The WSM method involves three stages. In the first, the constrained density-fitting scheme (Misquitta and Stone 2006) is used to calculate distributed, non-local polarizabilities. These are then *localized* in stage two. And subsequently, in the final stage, *refined* using a method based on Ref. Williams and Stone (2003).

Stage 1: From Eq. 26 we see that the density-fitted FDDS can be partitioned into atom–atom contributions by partitioning the auxiliary basis set into contributions from individual sites, that is, $\{\chi\} = \{\chi^{(1)}, \chi^{(2)}, \dots\}$, and then the FDDS can be written as

$$\alpha(\mathbf{r}, \mathbf{r}' | \omega) \approx \sum_{aa'} \alpha^{a,a'}(\mathbf{r}, \mathbf{r}' | \omega), \quad (45)$$

where the contribution from sites a and a' is given by

$$\alpha^{a,a'}(\mathbf{r}, \mathbf{r}' | \omega) = \sum_{p \in a, p' \in a'} \tilde{C}_{pp'}(\omega) \chi_p(\mathbf{r}) \chi_{p'}(\mathbf{r}'). \quad (46)$$

We can now insert $\alpha^{a,a'}(\mathbf{r}, \mathbf{r}' | \omega)$ in Eq. 36 to define the distributed polarizability for sites (a, a') :

$$\alpha_{tt'}^{aa'}(\omega) = \sum_{p \in a, p' \in a'} \tilde{C}_{pp'}(\omega) \mathcal{N}_t^p \mathcal{N}_{t'}^{p'}, \quad (47)$$

where $\mathcal{N}_t^p = \int \hat{Q}_t(\mathbf{r} - \mathbf{a}) \chi_p(\mathbf{r}) d^3\mathbf{r}$, where \mathbf{a} is a suitable reference origin for site a that will typically be taken to be the nucleus. Here $\hat{Q}_t(\mathbf{r} - \mathbf{a}) \equiv \hat{Q}_t^a$ is the multipole moment operator of rank t , centered on site a .

This simple idea fails when the standard density-fitting procedure is used. The resulting distributed polarizabilities are very sensitive to the auxiliary basis used (Misquitta and Stone 2006) and become unphysical as the basis grows larger and more complete. This is because in a large diffuse basis, diffuse functions on a site can be used to describe properties on another, adjacent site. This means that the basis functions located on a site are not representative of the site properties.

This problem can be overcome by modifying the density-fitting algorithm (section “Density-Fitting”) to include two additional constraints: one to impose orthogonality and the other to remove the non-physical terms (Misquitta and Stone

2006). With this modification, the density-fitted form of the FDDS leads to much more sensible values (Misquitta and Stone 2006) and are only weakly dependent on the basis set used.

Step 2: The distributed polarizabilities obtained using this constrained density-fitting procedure contain non-local terms, that is, terms involving pairs of distinct sites. In contrast to other distribution methods, the non-local terms describing flow of charge from site to site are very small (around 10^{-2} and 10^{-3} a.u. in magnitude) for all systems, irrespective of the type of bonding involved. Nevertheless, non-local terms are best avoided as they complicate the description unnecessarily. The localization methods of Le Sueur and Stone (1994) and Lillestolen and Wheatley (2007) can be used to transform the non-local terms into local polarizabilities and remove the charge-flow terms altogether. The localization by this procedure causes a deterioration of the convergence properties of the model, because multipole expansions are used to move the polarizabilities around. In principle, the Le Sueur and Stone procedure can increase the radius of divergence of the description to be equal to the size of the molecule, thereby causing significant losses in accuracy for large molecules for which the Lillestolen and Wheatley localization method may be more appropriate. Thus, while good results have been obtained for molecules like formamide and urea, there is already an appreciable loss in accuracy for *N*-methyl propanamide (Misquitta and Stone 2006).

Step 3: The quality of the local polarizability model from the previous step can be dramatically improved using the method of Williams and Stone (2003). In this step, the local polarizabilities are *refined* by requiring them to reproduce a set of point-to-point polarizabilities which describe the response of the electrostatic potential at a point to the frequency-dependent potential produced by a unit oscillating point charge at another. Details of the refinement process are given in Misquitta and Stone (2008a) and Misquitta et al. (2008a).

Analyzing the Models

The WSM method can be used to construct a variety of polarization models. These can include terms to rank 3 and can even be made isotropic. The refinement step ensures that whatever the choice of model, it will be the most accurate possible. A convenient way of assessing the accuracy of the models is by mapping induction energies onto a suitable surface around the molecule in question (Misquitta et al. 2008a). Such a mapping can be performed using a spherical energy probe which, for the induction energy, is best taken to be a point charge with charge qe . The induction energy of a molecule in the field of a point charge depends quadratically on the magnitude of the charge, and an appropriate value of q needs to be used in interpreting the energy scales in these maps. Setting $q = 1$ gives the response to a unit charge, but this is larger than typical local charges in a molecule, which are not expected to exceed $0.5 e$. The surface around the molecule is constructed using the algorithm described in Misquitta et al. (2008a).

In Fig. 1 we show difference maps of the induction energy of the formamide molecule in the field of a unit point charge, computed using the non-local models, local models, and WSM local models, respectively. The large errors in the

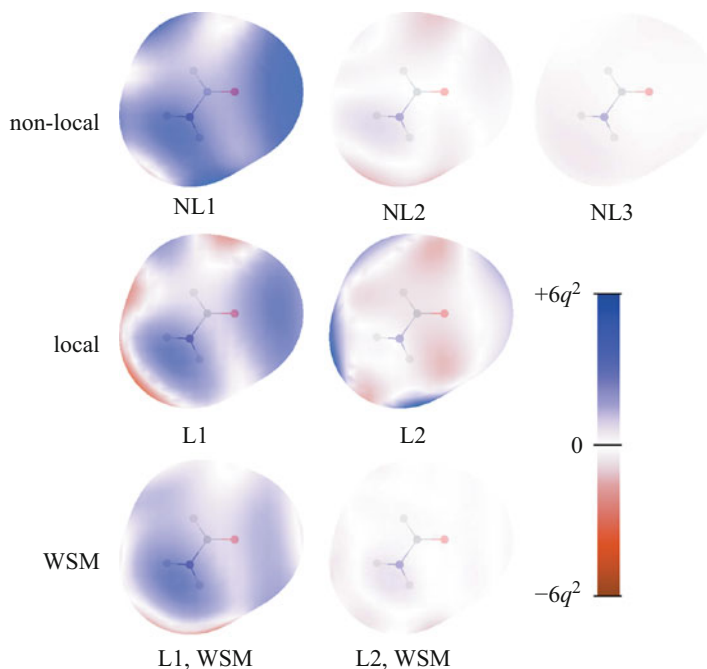


Fig. 1 Difference maps of the induction energy (kJ mol^{-1}) arising from a charge q atomic units on the $\text{vdW} \times 2$ surface of formamide using distributed non-local description of ranks 1, 2, and 3, distributed local description of ranks 1 and 2 obtained from the non-local models using the Le Sueur and Stone localization technique (Le Sueur and Stone 1994), and WSM distributed local descriptions of ranks 1 and 2. The differences are taken against SAPT(DFT) second-order induction energies obtained using a molecular description with the Sadlej/MC basis set

non-local rank 1 model, particularly near the oxygen and the polar hydrogen atoms, are quite clearly displayed. These errors are reduced in the rank 1 local model and are still smaller in the WSM rank 1 local model. The largest residual errors always seem to occur in the same regions, perhaps indicating the need for higher-ranking polarizabilities on some sites than on others. At rank 2, the non-local model is in almost perfect agreement with SAPT(DFT), but the local model obtained using the Le Sueur and Stone localization method exhibits rather large deficiencies near the polar hydrogens. These are removed in the WSM rank 2 local model which is comparable in accuracy to the rank 2 non-local model.

Distributed Dispersion Coefficients

Having obtained the localized frequency-dependent polarizabilities, we can now evaluate the dispersion coefficients that appear in Eq. 38. The dispersion coefficients depend on integrals of the form

$$\int_0^{\infty} \alpha_{ll'}^a(\nu) \alpha_{ll'}^b(\nu) d\nu, \quad (48)$$

where rank- l_1 – rank- l'_1 and rank- l_2 – rank- l'_2 polarizability tensors contribute to a C_n coefficient with $n = l_1 + l'_1 + l_2 + l'_2 + 2$. At present, the WSM polarizabilities can be calculated to rank 3, so while we can calculate C_{10} and C_{12} terms, they will lack the contributions from the hexadecapole and higher-rank polarizabilities. However, this has been shown not to be a serious limitation (Misquitta and Stone 2008b).

Basis sets: For a good dispersion model the basis set needs to be large and diffuse. To be consistent with SAPT(DFT) interaction energies calculated using the MC+ (or DC+) basis types, the dispersion coefficients should be calculated using a more diffuse basis than was used for the monomer part of the SAPT(DFT) basis. This is to compensate for the effects of the mid-bond functions that cannot be used when calculating dispersion coefficients. For example, if an aug-cc-pVTZ MC+ basis is used to calculate SAPT(DFT) interaction energies, the dispersion model should be constructed from WSM polarizabilities calculated using a d-aug-cc-pVTZ basis (Misquitta and Stone 2008b).

Models: Dispersion models need to be rather elaborate before they are accurate enough to describe the dispersion energy at small separations. Figure 2 shows dispersion energy maps for benzene and N-methyl propanamide using neon as a probe atom as described in Misquitta and Stone (2008b). From the error maps we see

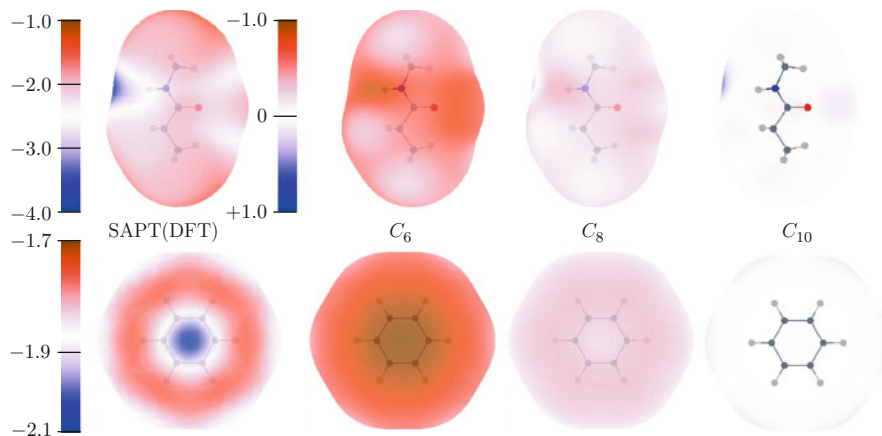


Fig. 2 Dispersion energy maps of benzene and N-methyl propanamide with neon as a probe. The SAPT(DFT) dispersion energy $E_{\text{disp,pol}}^{(2)}$ is displayed using an absolute scale in kJ mol^{-1} . The model dispersion energies are displayed as *differences* taken against $E_{\text{disp,tot}}^{(2)}$ from SAPT(DFT). The models are not damped. The d-aug-cc-pVTZ basis has been used for N-methyl propanamide and the aug-cc-pVTZ basis for benzene

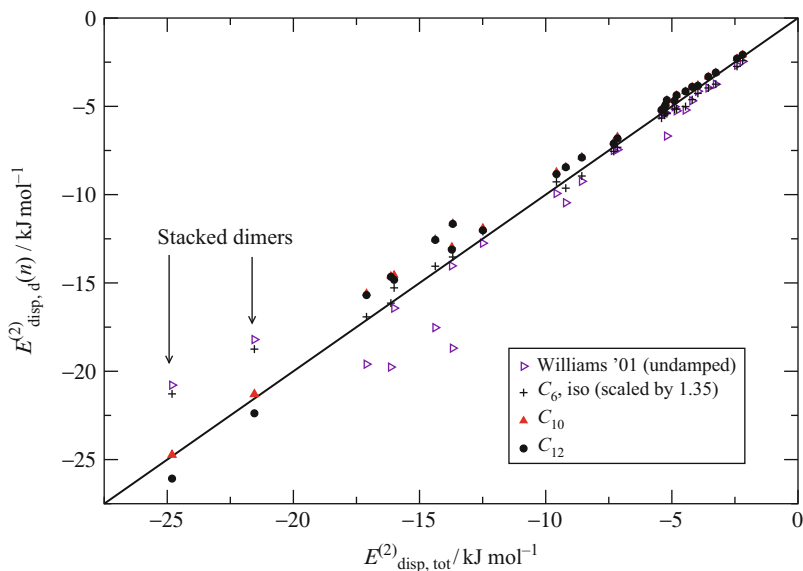


Fig. 3 Dispersion energies for the benzene dimer. Scatter plot of dispersion energies calculated using the damped anisotropic C_{10} and C_{12} models and the isotropic scaled C_6 model plotted against $E_{\text{disp,tot}}^{(2)}$ calculated using SAPT(DFT). Also shown are undamped dispersion energies calculated using the C_6 parameters from the Williams' 01 potential (Williams 2001a, b)

that the C_6 dispersion models significantly underestimate the dispersion interaction, and it is only with the C_{10} model that we get a good representation of the dispersion. More examples are provided in Misquitta and Stone (2008b).

However, the C_{10} model is rather elaborate and unwieldy. For many applications simpler models may suffice. Fortunately, the WSM method allows the construction of far simpler *isotropic* dispersion models. Even more simplifications can be made by scaling the isotropic C_6 WSM model to best reproduce the SAPT(DFT) dispersion energies (Misquitta and Stone 2008b), and the results are quite acceptable. From Fig. 3 we see that the scaled isotropic C_6 dispersion model (damped) is quite acceptable for the benzene dimer and is a significant improvement over the empirical dispersion model from a popular and well parameterized empirical potential. The significant failure occurs at the stacked configurations where anisotropy is essential to describe the dispersion interaction correctly.

Case Studies

Applications

Over the last few years, SAPT(DFT) has been used to study the interactions between a number of small and medium-sized molecules. Initial applications tended to

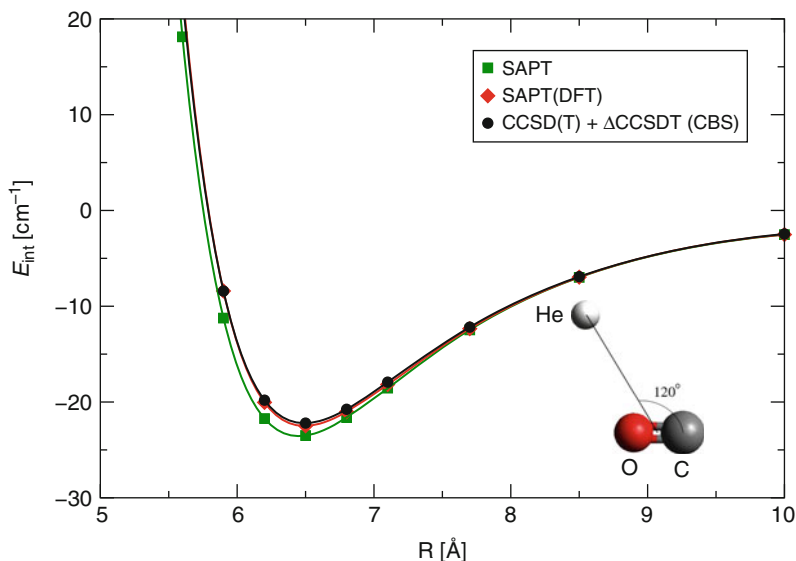


Fig. 4 Interaction energy scan for the He \cdots CO system. Comparison of SAPT, SAPT(DFT), and CCSDT (From Peterson and McBane 2005)

focus on small systems such as dimers of rare gas atoms (Misquitta et al. 2005a; Podeszwa and Szalewicz 2005) where detailed comparisons were possible with the more established CCSD(T) method, from which the accuracy of SAPT(DFT) was validated. The agreement was generally excellent. Consider for example the He \cdots CO complex (Fig. 4). The agreement of SAPT(DFT) energies with the CCSDT energies extrapolated to the complete basis set limit (Peterson and McBane 2005) is remarkable. By contrast, SAPT yields a potential energy curve that is somewhat too negative.

Subsequently, applications were made to larger systems such as the ethyne and benzene dimers (Hesselmann et al. 2005; Podeszwa and Szalewicz 2005) where, once again, the accuracy of the method was demonstrated. More recently, the DFT-SAPT implementation has been used to study the interaction of water on graphene (Jenness and Jordan 2009), DNA base pairs (Hesselmann et al. 2006), and DNA base pair tetramers (Fiethen et al. 2008). Calculations of such accuracy are simply not possible using conventional electronic structure methods in the supermolecular approach, particularly as the only correlated ab initio method with a reasonable computational scaling is MP2, and MP2 is not suitable for the interactions of systems with π -bonding (Sinnokrot et al. 2002).

In a truly remarkable application on water, Bukowski et al. (2006) derived a SAPT(DFT) two-body potential, together with a three-body potential from SAPT (Mas et al. 2003a, b) and a polarization model for the many-body effects, and were able to describe not only the vibrational-rotational tunneling spectrum of the water dimer and trimer, but also the structure of liquid water.

Complete potential energy surfaces of a few complexes have been calculated using SAPT(DFT). These include the CO dimer, for which ro-vibrational transitions were calculated and found to be in good agreement with experiment (Vissers et al. 2005), and the benzene dimer (Podeszwa et al. 2006a). Recently, an anisotropic, transferable atom–atom potential for polyaromatic hydrocarbons has been parametrized from SAPT(DFT) calculations on the benzene, naphthalene, anthracene, and pyrene dimers (Podeszwa and Szalewicz 2008; Totton et al. 2010). Such potentials could bring a new level of accuracy to the simulations of condensed matter.

Polarization in Organic Crystals

One of the successes of the WSM polarizability models has been in describing the role of polarization in organic crystals (Welch et al. 2008). Amongst the many crystals of organic molecules studied by Welch et al. let us consider one here: oxalyl dihydrazide (Fig. 5). This molecule has been found to crystallize in five polymorphic forms. In the α form, the molecular conformation is quite different from the other four, with two fewer internal hydrogen bonds. While the experimental lattice energies of these five forms are not known, it is generally expected that they should lie close together in energy. However, lattice energies calculated using methods that only modeled the exchange-repulsion, dispersion, electrostatic, and conformational energies predict the lattice energy of the α form to be between 20 and 28 kJ mol⁻¹ higher than the others. Such a high energy is generally considered to be out of the range of possible polymorphism.

Using a WSM polarizability model, Welch et al. showed that this difference could be explained as arising from polarization (induction) effects that were missing in the empirical force fields (see Fig. 5). When polarization is taken into account explicitly and added to the lattice energy (without subsequent lattice relaxation) all five polymorphs of oxalyl dihydrazide are seen to have a similar lattice energy. It is not hard to see why an explicit model was found to be important here. Empirical force fields (Williams 2001a, b) will always contain some of the effects of polarization in an average manner. This is why they work in many cases. But when there is a difference in the hydrogen-bonding network, such as happens in polymorphs of oxalyl dihydrazide where we see a difference in the number of intra and intermolecular hydrogen bonds, these potentials are no longer able to describe the relative energies of the systems, and we need an explicit treatment of the polarization. See Karamertzanis et al. (2008) for a detailed investigation of this instance of conformational polymorphism.

These effects will undoubtedly be important in other systems which involve different hydrogen-bonding environments such as biological molecules. It is very likely that even a simple treatment of polarization may significantly improve the predictions of the properties of such systems.

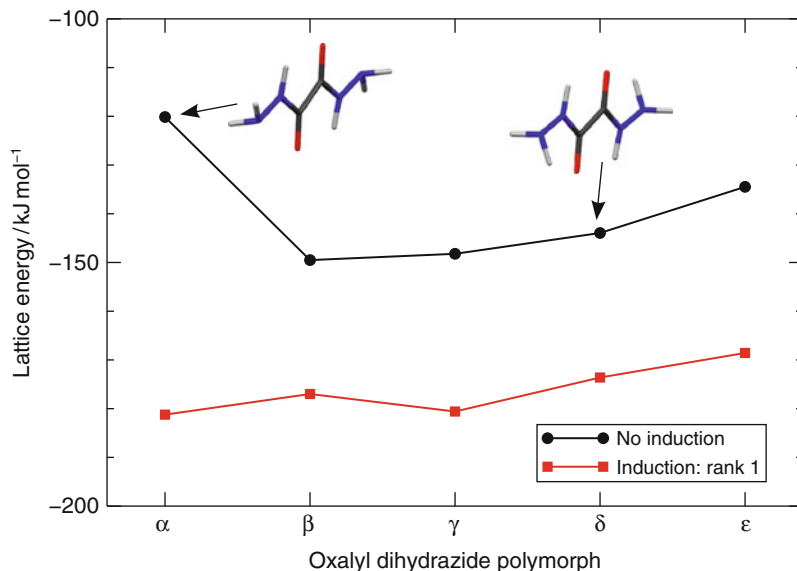


Fig. 5 Lattice energy of the five polymorphs of oxalyl dihydrazide calculated with and without an explicit polarization contribution. Without polarization, the α form is too high in energy compared with the others, but on calculating the polarization contribution to the lattice energy self-consistently, the five forms are seen to be similar in energy, which is what would be expected for polymorphs (Welch et al. 2008)

Crystal Structure Prediction

One of the significant successes of the electronic structure methods described in this chapter is the high level of predictive power they have brought to the field of organic crystals. Consider that just about 20 years ago the inability of computational scientists to predict crystal structures of organic molecules was considered a “scandal” (Maddox 1988). While we still cannot claim to have solved the problem, we do now have a number of noteworthy successes in this field.

One of the target molecules in the 2007 blind test of organic crystal structure prediction organized by the Cambridge Crystallographic Data Center (CCDC) was 1,3-dibromo-2-chloro-5-fluorobenzene ($C_6Br_2ClFH_2$). This molecule contains halogen atoms which are not well represented by most empirical potentials, primarily because they are strongly anisotropic (Day and Price 2003; Stone 1996) and empirical atom–atom potentials assume spherical atomic shapes. Using an anisotropic atom–atom potential calculated using first-order SAPT(DFT) energies and a dispersion model from WSM polarizabilities, Misquitta et al. (2008b) were able to predict a crystal structure for this system that was later found to correspond exactly with the experimental one (Fig. 6) (Day et al. 2009). This was the first completely ab initio prediction of a crystal structure of an organic molecule.

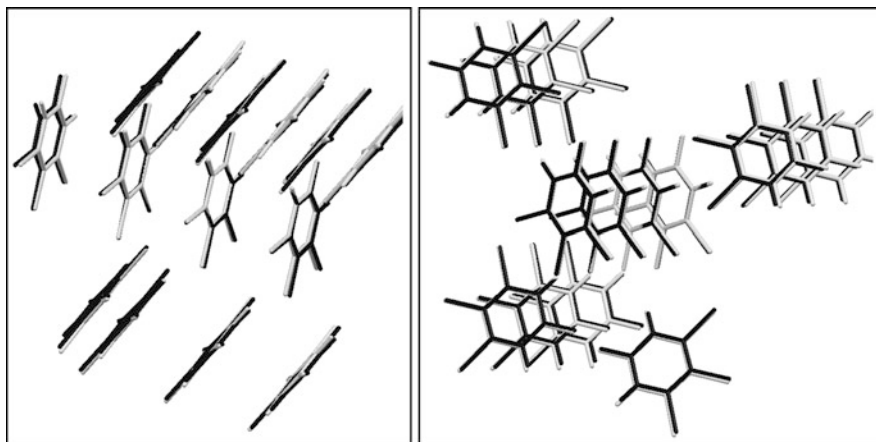


Fig. 6 Two views of the overlay of the experimental (*black*) and predicted (*gray*) crystal structures of $C_6Br_2ClFH_2$. The two structures are hard to separate as the overlay is nearly perfect

Also recently, Podeszwa et al. (2008) have used potentials derived from SAPT(DFT) to model the crystal structures of cyclotrimethylene trinitramine (RDX) and benzene, the latter being arguably the more difficult of the two systems. This once again demonstrates the accuracy of force fields derived from SAPT(DFT) energies.

Outlook

Applications such as these would have been almost unthinkable as little as a decade ago. It is developments in the theory of intermolecular forces together with improvements in our computational resources that have made these calculations possible. We can only speculate where future developments will take us in the next few years. For many applications computational requirement will limit us to use atom–atom intermolecular potentials. We have already seen examples of the success of such potentials when derived using accurate *ab initio* methods, but this is only the beginning of what is possible. Accurate *ab initio* methods like SAPT(DFT) and the WSM method can be the basis of a new generation of atom–atom potentials that will surpass any empirical potential in accuracy.

While we have good reason for optimism, it is worthwhile bearing in mind that at least some of the successes described here are fortuitous. The crystal structure predictions listed above were successful because the observed crystal form corresponded to the global minimum in the free energy. However, this is not always true, particularly for industrially important polymorphic molecules for which kinetic effects can be important. If we are to fully understand the interplay between kinetic and thermodynamic effects, we first need to remove any uncertainty in our ability to model molecular interactions – probably of both the intra and intermolecular types.

This raises the question of intramolecular flexibility which has not yet been adequately addressed. The potential energy surface involving intermolecular and intramolecular degrees of freedom is a formidable object and so far has been constructed for only a small number of small dimers. Methods for the inclusion of internal degrees of freedom have been suggested (Jankowski 2004, 2008; Jankowski and Ziolkowski 2006; Murdachaew et al. 2002), but so far have been applied to dimers of small systems only. This remains one of the outstanding problems with this approach.

Yet another outstanding problem is the correct treatment of the interactions of small gap materials such as fullerenes. It is quite likely that second-order perturbation will not be adequate for such systems. Furthermore, the strong electron delocalization in these semi-conductor-like materials means that the standard atom–atom models of interaction fail due to their inherent assumption of locality (Misquitta et al. 2010). This is possibly the next hurdle to be faced by the theory of intermolecular interactions.

Programs

There are a large number of programs that can be used to study intermolecular interactions. If the supermolecular method is used, practically any electronic structure program capable of using correlated methods like MP2 and CCSD(T) can be used. However, the following programs are likely more suited than most for this field.

- **SAPT2008:** (free, source available, parallel) The SAPT2008 program (Bukowski et al. 2002) contains implementations of both SAPT and DF-SAPT(DFT) and allows the calculation of three-body energies through the three-body SAPT and SAPT(DFT) theories. Besides the usual interaction energy corrections, this program also contains a variety of state-of-the-art energy components. The SAPT2008 program uses molecular orbitals and eigenvalues from a variety of SCF programs, but the DF-SAPT(DFT) module works only with the DALTON 2.0 program.
- **CAMCASP:** (free, source available, serial) The CAMCASP (Misquitta and Stone 2007) contains implementations of DF-SAPT(DFT), the WSM method for molecular properties, and a variety of modules and programs to analyze and interpret mechanisms of intermolecular interactions in small organic molecules (through interfaces with the ORIENT program) and construct analytic, polarizable atom–atom potentials for small organic molecules. Additionally, distributed multipoles can be calculated using CAMCASP through the included GDMA 2.2 module. This program has been used for almost all the results presented in this chapter. CAMCASP uses molecular orbitals and eigenvalues calculated using the DALTON 2.0 program (Helgaker et al. 2005).
- **ORIENT:** (free, source available, serial) ORIENT (Stone et al. 2006) is not an electronic structure program, but is a very versatile platform for calculating

interaction energies of clusters, performing geometry optimizations, simulations and displaying energies using the Open-GL module. This program is used quite closely by CAMCASP through a variety of interfaces.

- MOLPRO 2008: (commercial, source available, parallel) The MOLPRO 2008 program (Werner et al. 2008) contains a number of state-of-the-art electronic structure methods including DF-DFT-SAPT. This program is probably the most computationally efficient implementation of SAPT(DFT), though the FDDS cannot include any fraction of CHF (see section “Symmetry-Adapted Perturbation Theory Based on DFT”).
- DALTON 2.0: (free, source available, parallel) The DALTON 2.0 program (Helgaker et al. 2005) is used by both SAPT2008 and CAMCASP as the front-end to calculate the Kohn–Sham molecular orbitals and eigenvalues and some of the integrals needed to evaluate the FDDS.

Annotated Bibliography

The papers highlighted in this section are meant to be the most recent or comprehensive in the subject. This list is by no means complete. They should be a good starting point for researchers entering the field.

- Intermolecular forces: The book by Stone (1996) is probably the most comprehensive introduction to the theory from the atomistic point of view. Israelachvili (2007) takes a very different approach and is more concerned with the manifestations of these forces on surfaces, bulk media, and complex systems. Yet another approach is taken by Parsegian (2005) who tackles the Casimir forces between extended media in various geometries. The two volumes – (Wales 2005a, b) – offer articles on the current status of the field.
- Types of intermolecular bonds: An excellent overview of hydrogen bonds can be found in the article by Buckingham, Del Bene, and McDowell (Buckingham et al. 2008). In the last few years interest has developed in halogen bonded systems. See Bernal-Uruchurtu et al. (2009) for an introduction and an analysis based on a SAPT decomposition of the interaction energy.
- SAPT: The review articles by Jeziorski, Szalewicz, and others (Jeziorski and Szalewicz 1998, 2002; Jeziorski et al. 1994; Szalewicz 2002; Szalewicz et al. 2005) are probably the most comprehensive articles on symmetry-adapted perturbation theory. See also the discussion of these theories in Stone (1996). For a detailed examination of symmetry-forcing techniques and the convergence of intermolecular perturbation theories see Szalewicz et al. (2005).
- SAPT(DFT): The theory of the SAPT(DFT) method is described in Misquitta et al. (2005b) and DF-SAPT(DFT) in Podeszwa et al. (2006b). The DF-DFT-SAPT method is described in Hesselmann et al. (2005).
- Potentials: Obtaining an analytic expression for the intermolecular interaction energy in atom–atom form is often difficult. See Misquitta et al. (2008b) and Stone and Misquitta (2007) for a discussion of how this could be done.

- Crystal structure prediction: The field of organic crystal structure prediction remains one of the best testing grounds for intermolecular potentials. Accuracies need not be as high as that needed for spectroscopic calculations, but the effects of molecular flexibility and many-body non-additivity need to be accounted for. See Price (2008, 2009) for recent reviews of this subject. For a description of dispersion-corrected DFT methods specially parametrized for organic crystals see Neumann and Perrin (2005). For a comprehensive examination of the role of detailed distributed multipole models in this field see Day et al. (2005).

Acknowledgments The theoretical methods described in this chapter have been strongly influenced by Anthony Stone and Krzysztof Szalewicz. I owe particular thanks to Anthony who has had a strong influence on this chapter through the many discussions we have had and papers we have written together.

Bibliography

- Adamo, C., & Barone, V. (1999). Toward reliable density functional methods without adjustable parameters: The pbe0 model. *Journal of Chemical Physics*, *110*, 6158–6170.
- Adamo, C., Cossi, M., Scalmani, G., & Barone, V. (1999). Accurate static polarizabilities by density functional theory: Assessment of the PBE0 model. *Chemical Physics Letters*, *307*, 265–271.
- Akin-Ojo, O., Bukowski, R., & Szalewicz, K. (2003). Ab initio studies of He-HCCCN interaction. *Journal of Chemical Physics*, *119*, 8379–8396.
- Angyan, J. G. (2007). On the exchange-hole model of London dispersion forces. *Journal of Chemical Physics*, *127*, 024108.
- Angyan, J. G., Jansen, G., Loos, M., Hattig, C., & Hess, B. A. (1994). Distributed polarizabilities using the topological theory of atoms in molecules. *Chemical Physics Letters*, *219*, 267–273.
- Becke, A. D., & Johnson, E. R. (2005). Exchange-hole dipole moment and the dispersion interaction. *Journal of Chemical Physics*, *122*, 154104.
- Bernal-Uruchurtu, M. I., Hernandez-Lamonedá, R., & Janda, K. C. (2009). On the unusual properties of halogen bonds: A detailed ab initio study of x_2 -(h₂o)₁₋₅ clusters (x = cl and br). *Journal of Physical Chemistry A*, *113*, 5496–5505.
- Boys, S. F., & Bernardi, F. (1970). The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Molecular Physics*, *19*, 553.
- Buckingham, A. D., & Fowler, P. W. (1983). Do electrostatic interactions predict structures of van der Waals molecules? *Journal of Chemical Physics*, *79*, 6426–6428.
- Buckingham, A. D., Bene, J. E. D., & McDowell, S. A. C. (2008). The hydrogen bond. *Chemical Physics Letters*, *463*, 1–10.
- Bukowski, R., & Szalewicz, K. (2001). Complete ab initio three-body nonadditive potential in monte carlo simulations of vapor-liquid equilibria and pure phases of argon. *Journal of Chemical Physics*, *114*, 9518.
- Bukowski, R., Sadlej, J., Jeziorski, B., Jankowski, P., Szalewicz, K., Kucharski, S. A., Williams, H. L., & Rice, B. M. (1999). Intermolecular potential of carbon dioxide dimer from symmetry-adapted perturbation theory. *Journal of Chemical Physics*, *110*, 3785–3803.
- Bukowski, R., Cencek, W., Jankowski, P., Jeziorski, B., Jeziorska, M., Lotrich, V., Kucharski, S., Misquitta, A. J., Moszynski, R., Patkowski, K., Podeszwa, R., Rybak, S., Szalewicz, K.,

- Williams, H., Wheatley, R. J., Wormer, P. E. S., & Zuchowski, P. S. (2002). *SAPT2008: An ab initio program for many-body symmetry-adapted perturbation theory calculations of intermolecular interaction energies*. University of Delaware and University of Warsaw. <http://www.physics.udel.edu/~szalewic/>. Accessed 18 July 2011.
- Bukowski, R., Podeszwa, R., & Szalewicz, K. (2005). Efficient generation of the coupled Kohn–Sham dynamic susceptibility functions and dispersion energy with density fitting. *Chemical Physics Letters*, *414*, 111–116.
- Bukowski, R., Szalewicz, K., Groenenboom, G., & van der Avoird, A. (2006). Interaction potential for water dimer from symmetry-adapted perturbation theory based on density functional description of monomers. *Journal of Chemical Physics*, *125*, 044301.
- Bukowski, R., Szalewicz, K., Groenenboom, G. C., & van der Avoird, A. (2007). Predictions of the properties of water from first principles. *Science*, *315*, 1249–1252.
- Burcl, R., Chalasinski, G., Bukowski, R., & Szczesniak, M. M. (1995). On the role of bond functions in interaction energy calculations: $\text{Ar} \cdots \text{HCl}$, $\text{Ar} \cdots \text{H}_2\text{O}$, $(\text{HF})_2$. *Journal of Chemical Physics*, *103*, 1498–1507.
- Casida, M. E. (1995). Time-dependent density-functional response theory for molecules. In D. P. Chong (Ed.), *Recent advances in density-functional theory* (p. 155). Singapore: World Scientific.
- Chalasinski, G., & Szczesniak, M. M. (2000). State of the art and challenges of the ab initio theory of intermolecular interactions. *Chemical Reviews*, *100*, 4227–4252.
- Chang, B., Akin-Ojo, O., Bukowski, R., & Szalewicz, K. (2003). Potential energy surface and rovibrational spectrum of He–N₂O dimer. *Journal of Chemical Physics*, *119*, 11654.
- Colwell, S. M., Handy, N. C., & Lee, A. M. (1995). Determination of frequency-dependent polarizabilities using current density-functional theory. *Physical Review A*, *53*, 1316–1322.
- Day, G. M., & Price, S. L. (2003). A nonempirical anisotropic atom-atom model potential for chlorobenzene crystals. *Journal of the American Chemical Society*, *125*, 16434–16443.
- Day, G. M., Motherwell, W. D. S., & Jones, W. (2005). Beyond the isotropic atom model in crystal structure prediction of rigid molecules: Atomic multipoles versus point charges. *Crystal Growth and Design*, *5*, 1023–1033.
- Day, G. M., Cooper, T. G., Cabeza, A. J. C., Hejczyk, K., Ammon, H. L., Boerrigter, S. X. M., Tan, J., Valle, R. G. D., Venuti, E., Jose, J., Gadre, S. R., Desiraju, G. R., Thakur, T. S., van Eijck, B. P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Gavezzotti, A., Hofmann, D. W. M., Neumann, M., Leusen, F. J. J., Price, J. K. S. L., Misquitta, A. J., Karamertzanis, P. G., Welch, G., Scheraga, H. A., Arnautova, Y. A., Schmidt, M. U., van de Streek, J., Wolf, A., & Schweizer, B. (2009). Significant progress in predicting the crystal structures of small organic molecules – A report on the fourth blind test. *Acta Crystallographica. Section B*, *65*, 107–125.
- Dion, M., Rydberg, H., Schroder, E., Langreth, D. C., & Lundqvist, B. I. (2004). Van der Waals density functional for general geometries. *Physical Review Letters*, *92*, 246401–246404.
- Dunlap, B. I. (2000). Robust and variational fitting. *Physical Chemistry Chemical Physics*, *2*, 2113–2116.
- Dunlap, B. I., Connolly, J. W. D., & Sabin, J. R. (1979). On first-row diatomic molecules and local density models. *Journal of Chemical Physics*, *71*, 4993–4999.
- Ernesti, A., & Hutson, J. M. (1997). Non-additive intermolecular forces from the spectroscopy of van der Waals trimers: A comparison of ar₂-hf and ar₂-hcl, including h/d isotope effects. *Journal of Chemical Physics*, *106*, 6288.
- Ernzerhof, M., & Scuseria, G. E. (1999). Assessment of the perdew–burke–ernzerhof exchange–correlation functional. *Journal of Chemical Physics*, *110*, 5029–5036.
- Fiethen, A., Jansen, G., Hesselmann, A., & Schutz, M. (2008). Stacking energies for average b-dna structures from the combined density functional theory and symmetry-adapted perturbation theory approach. *Journal of the American Chemical Society*, *130*, 1802–1803.
- Gagliardi, L., Lindh, R., & Karlstrom, G. (2004). Local properties of quantum systems: The loprop approach. *Journal of Chemical Physics*, *121*, 4494.
- Grabo, T., Petersilka, M., & Gross, E. K. U. (2000). Molecular excitation energies from time-dependent density functional theory. *Journal of Molecular Structure (Theochem)*, *501*, 353.

- Grimme, S. (2004). Accurate description of van der Waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry*, 25, 1463–1473.
- Helgaker, T., Klopper, W., Koch, H., & Noga, J. (1997). Basis set convergence of correlated calculations on water. *Journal of Chemical Physics*, 106, 9639–9646.
- Helgaker, T., Jensen, H. J. A., Joergensen, P., Olsen, J., Ruud, K., Aagren, H., Auer, A., Bak, K., Bakken, V., Christiansen, O., Coriani, S., Dahle, P., Dalskov, E. K., Enevoldsen, T., Fernandez, B., Haettig, C., Hald, K., Halkier, A., Heiberg, H., Hettema, H., Jonsson, D., Kirpekar, S., Kobayashi, R., Koch, H., Mikkelsen, K. V., Norman, P., Packer, M. J., Pedersen, T. B., Ruden, T. A., Sanchez, A., Saue, T., Sauer, S. P. A., Schimmelpfennig, B., Sylvester-Hvid, K. O., Taylor, P. R., & Vahtras, O. (2005). *Dalton, a molecular electronic structure program, release 2.0*. See <http://www.kjemi.uio.no/software/dalton/dalton.html>. Accessed 18 July 2011.
- Hesselmann, A. (2009). Derivation of the dispersion energy as an explicit density- and exchange-hole functional. *Journal of Chemical Physics*, 130, 084104–084105.
- Hesselmann, A., & Jansen, G. (2002a). First-order intermolecular interaction energies from Kohn–Sham orbitals. *Chemical Physics Letters*, 357, 464–470.
- Hesselmann, A., & Jansen, G. (2002b). Intermolecular induction and exchange-induction energies from coupled-perturbed Kohn–Sham density functional theory. *Chemical Physics Letters*, 362, 319–325.
- Hesselmann, A., & Jansen, G. (2003a). The helium dimer potential from a combined density functional theory and symmetry-adapted perturbation theory approach using an exact exchange-correlation potential. *Physical Chemistry Chemical Physics*, 5, 5010.
- Hesselmann, A., & Jansen, G. (2003b). Intermolecular dispersion energies from time-dependent density functional theory. *Chemical Physics Letters*, 367, 778–784.
- Hesselmann, A., Jansen, G., & Schutz, M. (2005). Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *Journal of Chemical Physics*, 122, 014103.
- Hesselmann, A., Jansen, G., & Schutz, M. (2006). Interaction energy contributions of h-bonded and stacked structures of the at and gc dna base pairs from the combined density functional theory and intermolecular perturbation theory approach. *Journal of the American Chemical Society*, 128, 11730–11731.
- Hobza, P., & Spöner, J. (2002). Toward true dna base-stacking energies: Mp2, ccSD(T), and complete basis set calculations. *Journal of the American Chemical Society*, 124, 11802.
- Hodges, M. P., Stone, A. J., & Xantheas, S. S. (1997). Contribution of many-body terms to the energy for small water clusters: A comparison of ab initio calculations and accurate model potentials. *Journal of Physical Chemistry A*, 101, 9163–9168.
- Ioannou, A. G., Colwell, S. M., & Amos, R. D. (1997). The calculation of frequency-dependent polarizabilities using current density functional theory. *Chemical Physics Letters*, 278, 278–284.
- Israelachvili, J. N. (2007). *Intermolecular and surface forces* (2nd ed.). Amsterdam: Academic.
- Jankowski, P. (2004). Approximate generation of full-dimensional ab initio van der Waals surfaces for high-resolution spectroscopy. *Journal of Chemical Physics*, 121, 1655–1662.
- Jankowski, P. (2008). Exploring the new three-dimensional ab initio interaction energy surface of the Ar-HF complex: Rovibrational calculations for Ar-HF and Ar-DF with vibrationally excited diatoms. *Journal of Chemical Physics*, 128, 154311.
- Jankowski, P., & Ziolkowski, M. (2006). Fitting the derivative surfaces for full-dimensional interaction potentials. *Molecular Physics*, 104, 2293–2302.
- Jenness, G. R., & Jordan, K. D. (2009). Df-dft-sapt investigation of the interaction of a water molecule to coronene and dodecabenzocoronene: Implications for the water-graphite interaction. *The Journal of Physical Chemistry*, 113, 10242–10248.
- Jeziorska, M., Jeziorski, B., & Cizek, J. (1987). Direct calculation of the hartree-fock interaction energy via exchange-perturbation expansion. The He ... He interaction. *International Journal of Quantum Chemistry*, 32, 149–164.
- Jeziorski, B., & Szalewicz, K. (1998). *Encyclopedia of computational chemistry* (Vol. 2, p. 1376). Chichester: Wiley.

- Jeziorski, B., & Szalewicz, K. (2002). Symmetry-adapted perturbation theory. In S. Wilson (Ed.), *Handbook of molecular physics and quantum chemistry* (Vol. 8, pp. 37–83). Chichester: Wiley. Chap. 8.
- Jeziorski, B., Moszynski, R., & Szalewicz, K. (1994). Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chemical Reviews*, *94*, 1887–1930.
- Karamertzanis, P. G., Day, G. M., Welch, G. W. A., Kendrick, J., Leusen, F. J. J., Neumann, M. A., & Price, S. L. (2008). Modeling the interplay of inter- and intramolecular hydrogen bonding in conformational polymorphs. *Journal of Chemical Physics*, *128*, 244708–244717.
- Kendall, R. A., Dunning, T. H., & Harrison, R. J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *Journal of Chemical Physics*, *96*, 6796–6806.
- Koch, W. (2000). *A chemist's guide to density functional theory*. Weinheim/Chichester: Wiley.
- Le Sueur, C. R., & Stone, A. J. (1993). Practical schemes for distributed polarizabilities. *Molecular Physics*, *78*, 1267–1291.
- Le Sueur, C. R., & Stone, A. J. (1994). Localization methods for distributed polarizabilities. *Molecular Physics*, *83*, 293–308.
- Lehninger, A. L. (1970). *Biochemistry*. New York: Worth Publishers.
- Lillestolen, T. C., & Wheatley, R. J. (2007). First-principles calculation of local atomic polarizabilities. *Journal of Physical Chemistry A*, *111*, 11141–11146.
- Lillestolen, T. C., & Wheatley, R. J. (2008). Redefining the atom: Atomic charge densities produced by an iterative stockholder approach. *Chemical Communications*, 5909–5911.
- Longuet-Higgins, H. C. (1965). Intermolecular forces. *Discussions of the Faraday Society*, *40*, 7.
- Maddox, J. (1988). Crystals from first principles. *Nature*, *335*, 201.
- Magnasco, V., & McWeeny, R. (1991). Weak interaction between molecules and their physical interpretations. In Z. B. Maksić (Ed.), *Theoretical models of chemical bonding* (Vol. 4, pp. 133–169). New York: Springer.
- Mas, E. M., Szalewicz, K., Bukowski, R., & Jeziorski, B. (1997). Pair potential for water from symmetry-adapted perturbation theory. *Journal of Chemical Physics*, *107*, 4207–4218.
- Mas, E. M., Bukowski, R., & Szalewicz, K. (2003a). Ab initio three-body interactions for water. I. Potential and structure of water trimer. *Journal of Chemical Physics*, *118*, 4386–4403.
- Mas, E. M., Bukowski, R., & Szalewicz, K. (2003b). Ab initio three-body interactions for water. II. Effects on structure and energetics of liquid. *Journal of Chemical Physics*, *118*, 4404–4413.
- Matta, C. F., & Bader, R. F. W. (2006). An experimentalist's reply to "what is an atom in a molecule?". *Journal of Physical Chemistry A*, *110*, 6365–6371.
- Milet, A., Moszynski, R., Wormer, P. E. S., & van der Avoird, A. (1999). Hydrogen bonding in water clusters: Pair and many-body interactions from symmetry-adapted perturbation theory. *Journal of Physical Chemistry A*, *103*, 6811–6819.
- Misquitta, A. J., & Stone, A. J. (2006). Distributed polarizabilities obtained using a constrained density-fitting algorithm. *Journal of Chemical Physics*, *124*, 024111.
- Misquitta, A. J., & Stone, A. J. (2007). *CAMCASP: A program for studying intermolecular interactions and for the calculation of molecular properties in distributed form*. University of Cambridge. <http://www-stone.ch.cam.ac.uk/programs.html#CamCASP>. Accessed 18 July 2011.
- Misquitta, A. J., & Stone, A. J. (2008a). Accurate induction energies for small organic molecules: I. Theory. *Journal of Chemical Theory and Computation*, *4*, 7–18.
- Misquitta, A. J., & Stone, A. J. (2008b). Dispersion energies for small organic molecules: First row atoms. *Molecular Physics*, *106*, 1631–1643.
- Misquitta, A. J., & Szalewicz, K. (2002). Intermolecular forces from asymptotically corrected density functional description of monomers. *Chemical Physics Letters*, *357*, 301–306.
- Misquitta, A. J., & Szalewicz, K. (2005). Symmetry-adapted perturbation-theory calculations of intermolecular forces employing density-functional description of monomers. *Journal of Chemical Physics*, *122*, 214109.
- Misquitta, A. J., Bukowski, R., & Szalewicz, K. (2000). Spectra of Ar-CO₂ from ab initio potential energy surfaces. *Journal of Chemical Physics*, *112*, 5308–5319.

- Misquitta, A. J., Jeziorski, B., & Szalewicz, K. (2003). Dispersion energy from density-functional theory description of monomers. *Physical Review Letters*, *91*, 33201.
- Misquitta, A., Podeszwa, R., Jeziorski, B., & Szalewicz, K. (2005a). Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional calculations. *Journal of Chemical Physics*, *123*, 214103.
- Misquitta, A. J., Podeszwa, R., Jeziorski, B., & Szalewicz, K. (2005b). Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional theory. *Journal of Chemical Physics*, *123*, 214103.
- Misquitta, A. J., Stone, A. J., & Price, S. L. (2008a). Accurate induction energies for small organic molecules: II. Models and numerical details. *Journal of Chemical Theory and Computation*, *4*, 19–32.
- Misquitta, A. J., Welch, G. W. A., Stone, A. J., & Price, S. L. (2008b). A first principles prediction of the crystal structure of $C_6 Br_2 ClFH_2$. *Chemical Physics Letters*, *456*, 105–109.
- Misquitta, A. J., Spencer, J., Stone, A. J., & Alavi, A. (2010). Dispersion interactions between semiconducting wires. *Physical Review B*, *82*, 075312–075317.
- Moszyński, R., Jeziorski, B., Rybak, S., Szalewicz, K., & Williams, H. L. (1994). Many-body theory of exchange effects in intermolecular interactions. Density matrix approach and applications to $He-F^+$, $He-HF$, H_2-HF , and $Ar-H_2$ dimers. *Journal of Chemical Physics*, *100*, 5080–5092.
- Moszynski, R., Heijmen, T. G. A., & Jeziorski, B. (1996). Symmetry-adapted perturbation theory for the calculation of hartree-fock interaction energies. *Molecular Physics*, *88*, 741–758.
- Murdachaew, G., Misquitta, A. J., Bukowski, R., & Szalewicz, K. (2001). Intermolecular potential energy surfaces and spectra of Ne-HCN complex from ab initio calculations. *Journal of Chemical Physics*, *114*, 764.
- Murdachaew, G., Szalewicz, K., & Bukowski, R. (2002). Efficient generation of flexible-monomer intermolecular potential energy surfaces. *Physical Review Letters*, *88*, 12320.
- Neumann, M. A., & Perrin, M.-A. (2005). Energy ranking of molecular crystals using density functional theory calculations and an empirical van der Waals correction. *Journal of Physical Chemistry B*, *109*, 15531–15541.
- Onida, G., Reinig, L., & Rubio, A. (2002). Electronic excitations: Density-functional versus many-body greens-function approaches. *Reviews of Modern Physics*, *74*, 601.
- Parr, R. G., Ayers, P. W., & Nalewajski, R. F. (2005). What is an atom in a molecule? *Journal of Physical Chemistry A*, *109*, 3957–3959.
- Parsegian, V. A. (2005). *Van der Waals forces: A handbook for biologists, chemists, engineers, and physicists*. Cambridge: Cambridge University Press.
- Patkowski, K., Szalewicz, K., & Jeziorski, B. (2006). Third-order interactions in symmetry-adapted perturbation theory. *Journal of Chemical Physics*, *125*, 154107.
- Pernal, K., Podeszwa, R., Patkowski, K., & Szalewicz, K. (2009). Dispersionless density functional theory. *Physical Review Letters*, *103*, 263201(4).
- Petersilka, M., Gossmann, U. J., & Gross, E. K. U. (1996). Excitation energies from time-dependent density-functional theory. *Physical Review Letters*, *76*, 1212–1215.
- Peterson, K. A., & McBane, G. C. (2005). A hierarchical family of three-dimensional potential energy surfaces for He-CO. *Journal of Chemical Physics*, *123*, 084314–084315.
- Pitonak, M., & Hesselmann, A. (2010). Accurate intermolecular interaction energies from a combination of mp2 and tddft response theory. *Journal of Chemical Theory and Computation*, *6*, 168–178.
- Podeszwa, R., & Szalewicz, K. (2005). Accurate interaction energies for argon, krypton, and benzene dimers from perturbation theory based on the Kohn–Sham model. *Chemical Physics Letters*, *412*, 488.
- Podeszwa, R., & Szalewicz, K. (2007). Three-body symmetry-adapted perturbation theory based on Kohn–Sham description of the monomers. *Journal of Chemical Physics*, *126*, 194101.
- Podeszwa, R., & Szalewicz, K. (2008). Physical origins of interactions in dimers of polycyclic aromatic hydrocarbons. *Physical Chemistry Chemical Physics*, *10*, 2735–2746.

- Podeszwa, P., Bukowski, R., & Szalewicz, K. (2006a). Potential energy surface for the benzene dimer and perturbational analysis of $\pi - \pi$ interactions. *Journal of Physical Chemistry A*, *110*, 10345–10354.
- Podeszwa, R., Bukowski, R., & Szalewicz, K. (2006b). Density-fitting method in symmetry-adapted perturbation theory based on Kohn–Sham description of monomers. *Journal of Chemical Theory and Computation*, *2*, 400–412.
- Podeszwa, R., Rice, B. M., & Szalewicz, K. (2008). Predicting structure of molecular crystals from first principles. *Physical Review Letters*, *101*, 115503.
- Price, S. L. (2008). Computational prediction of organic crystal structures and polymorphism. *International Reviews in Physical Chemistry*, *27*, 541–568.
- Price, S. L. (2009). Computed crystal energy landscapes for understanding and predicting organic crystal structures and polymorphism. *Accounts of Chemical Research*, *42*, 117–126.
- Sadlej, A. J. (1992). Medium-sized polarized basis sets for high-level correlated calculations of molecular electric properties. V. Fourth-row atoms Sn–I. *Theoretical Chemistry Accounts*, *81*, 339–354.
- Sadlej, A. J., & Urban, M. (1991). Medium-sized polarized basis sets for high-level correlated calculations of molecular electric properties. III. Alkali atoms (Li, Na, K, Rb) and alkaline-earth atoms (Be, Mg, Ca, Sr). *Theochem – Journal of Molecular Structure*, *80*, 147–171.
- Sebetci, A., & Beran, G. J. O. (2010). Spatially homogeneous qm/mm for systems of interacting molecules with on-the-fly ab initio force-field parametrization. *Journal of Chemical Theory and Computation*, *6*, 155–167.
- Singh, U. C., & Kollman, P. A. (1984). An approach to computing electrostatic charges for molecules. *Journal of Computational Chemistry*, *5*, 129–145.
- Sinnokrot, M. O., Valeev, E. F., & Sherrill, C. D. (2002). Estimates of the ab initio limit for pi-pi interactions: The benzene dimer. *Journal of the American Chemical Society*, *124*, 10887–10893.
- Stone, A. J. (1993). Computation of charge-transfer energies by perturbation theory. *Chemical Physics Letters*, *211*, 101–109.
- Stone, A. J. (1996). *The theory of intermolecular forces*. Oxford: Clarendon.
- Stone, A. J. (2005). Distributed multipole analysis: Stability for large basis sets. *Journal of Chemical Theory and Computation*, *1*, 1128–1132.
- Stone, A. J., & Alderton, M. (1985). Distributed multipole analysis – methods and applications. *Molecular Physics*, *56*, 1047–1064.
- Stone, A. J., & Misquitta, A. J. (2007). Atom–atom potentials from ab initio calculations. *International Reviews in Physical Chemistry*, *26*, 193–222.
- Stone, A. J., & Misquitta, A. J. (2009). Charge-transfer in symmetry-adapted perturbation theory. *Chemical Physics Letters*, *473*, 201–205.
- Stone, A. J., & Tough, R. J. A. (1984). Spherical tensor theory of long-range intermolecular forces. *Chemical Physics Letters*, *110*, 123–129.
- Stone, A. J., Dullweber, A., Engkvist, O., Fraschini, E., Hodges, M. P., Meredith, A. W., Nutt, D. R., Popelier, P. L. A., & Wales, D. J. (2006). *Orient: A program for studying interactions between molecules, version 4.6*. University of Cambridge. <http://www-stone.ch.cam.ac.uk/programs.html#Orient>. Accessed 18 July 2011.
- Szalewicz, K. (2002). Hydrogen bond. In R. A. Meyers et al. (Eds.), *Encyclopedia of physical science and technology* (3rd ed., Vol. 7, pp. 505–538). San Diego: Academic.
- Szalewicz, K., Patkowski, K., & Jeziorski, B. (2005). Intermolecular interactions via perturbation theory: From diatoms to biomolecules. In D. J. Wales (Ed.), *Intermolecular forces and clusters II* (Structure and bonding, Vol. 116, pp. 43–117). Berlin: Springer.
- Tang, K. T., & Toennies, J. P. (1992). The damping function of the van der Waals attraction in the potential between rare gas atoms and metal surfaces. *Surface Science Letters*, *279*, 203–206.
- Totton, T., Misquitta, A. J., & Kraft, M. (2010). A first principles development of a general anisotropic potential for polycyclic aromatic hydrocarbons. *Journal of Chemical Theory and Computation*, *6*, 683–695.
- Tozer, D. J. (2000). The asymptotic exchange potential in Kohn–Sham theory. *Journal of Chemical Physics*, *112*, 3507–3515.

- Tozer, D. J., & Handy, N. C. (1998). Improving virtual Kohn–Sham orbitals and eigenvalues: Application to excitation energies and static polarizabilities. *Journal of Chemical Physics*, *109*, 10180–10189.
- Tsuzuki, S., & Lüthi, H. P. (2001). Interaction energies of van der Waals and hydrogen bonded systems calculated using density functional theory: Assessing the pw91 model. *Journal of Chemical Physics*, *114*, 3949–3957.
- van Mourik, T., & Gdanitz, R. J. (2002). A critical note on density functional theory studies on rare-gas dimers. *Journal of Chemical Physics*, *116*, 9620–9623.
- Visser, G. W. M., Hesselmann, A., Jansen, G., Wormer, P. E. S., & van der Avoird, A. (2005). New CO–CO interaction potential tested by rovibrational calculations. *Journal of Chemical Physics*, *122*(5), 054306. doi:10.1063/1.1835262.
- Wales, D. J. (Ed.). (2005a). *Intermolecular forces and clusters I* (Structure and bonding). Berlin: Springer.
- Wales, D. J. (Ed.). (2005b). *Intermolecular forces and clusters II* (Structure and bonding). Berlin: Springer.
- Weigend, F., Haser, M., Patzelt, H., & Ahlrichs, R. (1998). RI-MP2: Optimized auxiliary basis sets and demonstration of efficiency. *Chemical Physics Letters*, *294*, 143–152.
- Weigend, F., Kohn, A., & Hattig, C. (2002). Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *Journal of Chemical Physics*, *116*, 3175–3183.
- Welch, G. W. A., Karamertzanis, P. G., Misquitta, A. J., Stone, A. J., & Price, S. L. (2008). Is the induction energy important for modeling organic crystals? *Journal of Chemical Theory and Computation*, *4*, 522–532.
- Werner, H.-J., Knowles, P. J., Lindh, R., Manby, F. R., Schütz, M., Celani, P., Korona, T., Mitrushenkov, A., Rauhut, G., Adler, T. B., Amos, R. D., Bernhardsson, A., Berning, A., Cooper, D. L., Deegan, M. J. O., Dobbyn, A. J., Eckert, F., Goll, E., Hampel, C., Hetzer, G., Hrenar, T., Knizia, G., Köppl, C., Liu, Y., Lloyd, A. W., Mata, R. A., May, A. J., McNicholas, S. J., Meyer, W., Mura, M. E., Nicklass, A., Palmieri, P., Pflüger, K., Pitzer, R., Reiher, M., Schumann, U., Stoll, H., Stone, A. J., Tarroni, R., Thorsteinsson, T., Wang, M., & Wolf, A. (2008). *Molpro, version 2008.3, a package of ab initio programs*. See <http://www.molpro.net>. Accessed 18 July 2011.
- Wheatley, R. J., & Lillestolen, T. C. (2008). Local polarizabilities and dispersion energy coefficients. *Molecular Physics*, *106*, 1545–1556.
- Williams, D. E. (2001a). Improved intermolecular force field for crystalline oxohydrocarbons including O–H...O hydrogen bonds. *Journal of Computational Chemistry*, *22*, 1–20.
- Williams, D. E. (2001b). Improved intermolecular force field for molecules containing H, C, N, and O atoms, with applications to nucleoside and peptide crystals. *Journal of Computational Chemistry*, *22*, 1154–1166.
- Williams, G. J., & Stone, A. J. (2003). Distributed dispersion: A new approach. *Journal of Chemical Physics*, *119*, 4620–4628.
- Williams, H. L., Mas, E. M., Szalewicz, K., & Jeziorski, B. (1995). On the effectiveness of monomer-, dimer-, and bond-centered basis functions in calculations of intermolecular interaction energies. *Journal of Chemical Physics*, *103*, 7374–7391.
- Woon, D. E., & Dunning, T. H. (1994). Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties. *Journal of Chemical Physics*, *100*, 2975–2989.
- Xu, Y., Jager, W., Tang, J., & McKellar, A. R. W. (2003). *Physical Review Letters*, *91*, 163401(4).
- Zaremba, E., & Kohn, W. (1976). Van der Waals interaction between an atom and a solid surface. *Physical Review B*, *13*, 2270–2285.

Molecular Dynamics Simulation: From “Ab Initio” to “Coarse Grained”

9

Chris Lorenz and Nikos L. Doltsinis

Contents

Introduction	338
Choosing the Right Method	338
Theoretical Background	340
Ab Initio Molecular Dynamics	343
Classical Molecular Dynamics	344
Hybrid Quantum/Classical (QM/MM) Molecular Dynamics	347
Coarse-Grained Molecular Dynamics	350
Interaction Potentials/Force Fields	355
Classical Force Fields	355
First-Principles Electronic Structure Methods	362
Building the System/Collecting the Ingredients	365
Setting Up an AIMD Simulation	365
Setting Up a Classical MD Simulation	368
Preparing an Input File	370
Optimization Algorithms	370
Controlling Temperature: Thermostats	373
Controlling Pressure: Barostats	378
Setting the Timestep	380
Postprocessing	384
Data Analysis	384
Visualization	388
Bibliography	388

C. Lorenz (✉)

Department of Physics, King’s College London, Strand, London, UK

e-mail: chris.lorenz@kcl.ac.uk

N.L. Doltsinis

Institute for Solid State Theory and Center for Multiscale Theory and Computation, University of Münster, Münster, Germany

e-mail: nikos.doltsinis@wwu.de

Abstract

This chapter provides an overview of different hierarchical levels of molecular dynamics (MD) spanning a wide range of time and length scales – from first-principles approaches via classical atomistic methods to coarse-graining techniques. The theoretical background of the most widely used methods and algorithms is briefly reviewed, and practical instructions are given on the choice of input parameters for an actual computer simulation. In addition, important postprocessing procedures such as data analysis and visualization are discussed.

Introduction

Molecular dynamics (MD) simulations in their different flavors are widely used in a large variety of research areas of computational physics and chemistry. They represent a powerful tool to study the motion of atoms in molecules, liquids, and solids. The term MD typically refers to the propagation of point particles – atomic nuclei or effective particles combining several nuclei – according to the laws of classical mechanics. In particular, the forces acting on the particles are calculated “on the fly” only at discrete points along the trajectory. Following this definition, we discuss in this chapter *ab initio* MD (AIMD), i.e., the atomic forces are calculated from first principles, classical atomistic MD using analytical empirical interaction potentials (force fields), which sometimes is referred to as force-field molecular dynamics; and coarse-grained MD using analytical empirical potentials between effective particles representing groups of atoms. We exclude methods which go beyond classical nuclei, such as path integral MD (Tuckerman et al. 1993; Tuckerman and Hughes 1998; Tuckerman 2002) and wave packet dynamics (Balint-Kurti 2008; Worth et al. 2008), or beyond the Born–Oppenheimer approximation (Doltsinis and Marx 2002a, b). This overview, furthermore, leaves out the vast area of semi-empirical methods (see, e.g., Bredow and Jug (2005) for a recent review) including self-consistent charge density functional tight-binding (SCC-DFTB) (Elstner et al. 1998) and empirical valence-bond (EVB) theory (Warshel 1991; Aqvist and Warshel 1993; Shurki and Warshel 2003; Warshel 2003).

The aim of this chapter is to offer practical guidance on how to choose the appropriate technique for a particular physical problem, how to set up a simulation, and how to analyze and visualize the output. In addition it should provide the theoretical background required to become a competent user of the available simulation software packages.

Choosing the Right Method

When choosing which type of molecular dynamics simulations to perform, it is important to understand the capabilities of each technique. The differences in the various methods are basically dependent on the detail with which each one models a physical system.

The most detailed molecular dynamics simulation technique is the ab initio (quantum) molecular dynamics simulation approach that explicitly models the electrons within the system by solving the time-independent Schrödinger equation in some approximate fashion whereas force-field atomistic molecular dynamics simulations model the interactions between the atoms within the system by analytical functions that include the effects of the electrons only implicitly. Then the method that incorporates the least amount of detail is that of coarse-grained molecular dynamics models where multiple atoms are grouped together before being represented by a single interaction "bead."

Therefore, ab initio molecular dynamics simulations will generate the most accurate and versatile modeling of interatomic interactions as electrons are the basis of all such interactions. Quantum simulations allow for certain phenomena like electronic excitation or charge transport within a system to be modeled, which cannot be modeled in force-field or coarse-grained molecular dynamics simulations because they do not treat electrons explicitly. Also, in order to model chemical reactions, quantum simulations are the most accurate approach (note: there have been force-field and coarse-grained molecular dynamics simulations that have modeled the formation and breaking of bonds, but some a priori knowledge must then be included in the model to allow for the reaction to take place). The major limitation of quantum simulations is that they are very computationally intensive, which results in the capability to model only small system sizes (typically 10^2 – 10^3 atoms) and short times (10^{-12} – 10^{-10} s). Thus, the systems that can be modeled are limited to relatively small- to medium-sized molecules or portions of larger molecules (i.e., specific amino acids within a protein).

Atomistic force-field molecular dynamics simulations offer the ability to model molecules at the particle level. Often, information from quantum simulations is used to develop the empirical interaction potentials (force field) that are used to describe the interactions between particles. Because force-field molecular dynamics simulations use less detail than the quantum simulations, they are able to model systems that are significantly larger in size ($\sim 10^6$ particles) for a longer period of time ($\sim 10^{-6}$ s). Therefore, measuring the structural, mechanical, and/or transport properties of medium- to large-sized systems (i.e., proteins, functionalized nanoparticles, . . .) is possible.

Finally, coarse-grained molecular dynamics simulations reduce the number of degrees of freedom within the simulated system even further by grouping several atoms into one interaction bead. Therefore, even larger system sizes ($\sim \mu\text{m}$) and times (on the order of seconds) are accessible via these simulations. Several of the same properties measured via atomistic force-field molecular dynamics simulations can be measured with coarse-grained molecular dynamics simulations (i.e., structural, mechanical, and transport properties). However, due to the reduced detail in the models of the molecules, it is not possible to investigate specific chemical interactions within a system, such as hydrogen bonding.

Once you have chosen the appropriate method for the particular system and property to be investigated, the next choice is what simulation package to use. For classical MD simulations, there are several free molecular dynamics packages

that can be found on the Web including DL_POLY (Smith et al. 2002; http://www.cse.scitech.ac.uk/ccg/software/DL_POLY/), GROMACS (Spoel et al. 2005; <http://www.gromacs.org/>), HOOMD (Anderson et al. 2008; <http://codeblue.umich.edu/hoomd-blue/about.html/>), LAMMPS (Plimpton 1995; <http://lammps.sandia.gov/>), MOLLY (Refson 2000; <http://www.ccp5.ac.uk/moldy/moldy.html/>), TINKER (<http://dasher.wustl.edu/tinker/>), and NAMD (Phillips et al. 2005; <http://www.ks.uiuc.edu/Research/namd/>), and there are also commercial packages including AMBER (Case et al. 2005; <http://ambermd.org/>), CHARMM (Brooks et al. 2009; <http://www.charmm.org/>), and GROMOS (Scott et al. 1999; <http://www.igc.ethz.ch/GROMOS/index/>). Generally, these codes can be divided into those that are mostly used for simulations of biological systems (AMBER, CHARMM, GROMACS, GROMOS, NAMD) and those that are more general simulation packages (HOOMD, LAMMPS, MOLLY). When choosing between these options, an important criterion is to choose a code that you feel comfortable using. Outside of comfort, another aspect to take into consideration is that packages will differ in the features they offer and the additional tools to perform analysis (usually lists of analysis tools can be found in the packages' documentation).

For AIMD simulations the user may choose from a large number of codes, for instance, ABINIT (www.abinit.org; Aulbur et al. 2000), CASTEP (www.tcm.phy.cam.ac.uk/castep/; Segall et al. 2002; Clark et al. 2005), CONQUEST (www.conquest.ucl.ac.uk/; Bowler et al. 2006), CP2K (Hutter et al.; VandeVondele et al. 2005, 2006), CPMD (Hutter et al.; Marx and Hutter 2000, 2009), CP-PAW (<https://orion.pt.tu-clausthal.de/paw/>; Blochl 1994; Blochl et al. 2003), DACAPO (<http://dcwww.camp.dtu.dk/campos/Dacapo/>), FHI98md (www.fhi-berlin.mpg.de/th/fhimd/; Bockstedte et al. 1997), NWChem (www.emsl.pnl.gov/docs/nwchem; Kendall et al. 2000), ONETEP (www.tcm.phy.cam.ac.uk/onetep/; Skylaris et al. 2005), PINY (<http://homepages.nyu.edu/~mt33/PINY.MD/PINY.html>), PWscf (www.pwscf.org/; Giannozzi et al. 2009), QuantumEspresso (www.quantum-espresso.org; Giannozzi et al. 2009), SIESTA (www.uam.es/siesta/; Soler et al. 2002; Artacho et al. 2008), S/PHI/nX (www.sphinxlib.de; Boeck 2009), or VASP (<http://cms.mpi.univie.ac.at/vasp/>; Kresse and Furthmüller 1996).

To perform hybrid quantum/classical (QM/MM) simulations, one needs to interface a force-field MD program with an ab initio electronic structure code. This can be done, for example, with the following codes: ChemShell (www.chemshell.org/), CPMD (Hutter et al.), CP2K (Hutter et al.), LAMMPS (<http://lammps.sandia.gov/>), AMBER (<http://ambermd.org/>), and GROMACS (<http://www.gromacs.org/>).

Theoretical Background

Born–Oppenheimer Approximation

Let us begin by introducing our nomenclature and by reviewing some well-known basic relations within the Schrödinger formulation of quantum mechanics. A complete, nonrelativistic description of a dynamic system of N atoms hav-

ing the positions $\mathbf{R} = \{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_I, \dots, \mathbf{R}_N\}$ with n electrons located at $\mathbf{r} = \{\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_i, \dots, \mathbf{r}_n\}$ would involve solving the time-dependent Schrödinger equation

$$\mathcal{H}\Phi(\mathbf{r}, \mathbf{R}; t) = i\hbar \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}; t), \quad (1)$$

with the total Hamiltonian

$$\mathcal{H}(\mathbf{r}, \mathbf{R}) = \mathcal{T}(\mathbf{R}) + \mathcal{T}(\mathbf{r}) + \mathcal{V}_{\text{nn}}(\mathbf{R}) + \mathcal{V}_{\text{ne}}(\mathbf{r}, \mathbf{R}) + \mathcal{V}_{\text{ee}}(\mathbf{r}), \quad (2)$$

being the sum of kinetic energy of the atomic nuclei,

$$\mathcal{T}(\mathbf{R}) = -\frac{\hbar^2}{2} \sum_{I=1}^N \frac{\nabla_I^2}{M_I}, \quad (3)$$

kinetic energy of the electrons,

$$\mathcal{T}(\mathbf{r}) = -\frac{\hbar^2}{2m_e} \sum_{i=1}^n \nabla_i^2, \quad (4)$$

internuclear repulsion,

$$\mathcal{V}_{\text{nn}}(\mathbf{R}) = \frac{e^2}{4\pi\epsilon_0} \sum_{I=1}^{N-1} \sum_{J>I}^N \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \quad (5)$$

electronic–nuclear attraction,

$$\mathcal{V}_{\text{ne}}(\mathbf{r}, \mathbf{R}) = -\frac{e^2}{4\pi\epsilon_0} \sum_{I=1}^N \sum_{i=1}^n \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}, \quad (6)$$

and interelectronic repulsion,

$$\mathcal{V}_{\text{ee}}(\mathbf{r}) = \frac{e^2}{4\pi\epsilon_0} \sum_{i=1}^{n-1} \sum_{j>i}^n \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (7)$$

Here, M_I and Z_I denote the mass and atomic number of nucleus I ; m_e and e are the electronic mass and elementary charge, and ϵ_0 is the permittivity of vacuum. The nabla operators ∇_I and ∇_i act on the coordinates of nucleus I and electron i , respectively. The total wave function $\Phi(\mathbf{r}, \mathbf{R}; t)$ simultaneously describes the motion of both electrons and nuclei.

The Born–Oppenheimer approximation (Kołos 1970; Kutzelnigg 1997; Doltsinis and Marx 2002b) separates nuclear and electronic motion based on the assumption that the much faster electrons adjust their positions instantaneously to the comparatively slow changes in nuclear positions. The electronic problem is then reduced to the time-independent (electronic) Schrödinger equation for clamped nuclei,

$$\mathcal{H}_{\text{el}}(\mathbf{r}; \mathbf{R})\Psi_k(\mathbf{r}; \mathbf{R}) = E_k(\mathbf{R})\Psi_k(\mathbf{r}; \mathbf{R}), \quad (8)$$

where $\mathcal{H}_{\text{el}}(\mathbf{r}; \mathbf{R})$ is the electronic Hamiltonian,

$$\mathcal{H}_{\text{el}}(\mathbf{r}, \mathbf{R}) = \mathcal{T}(\mathbf{r}) + \mathcal{V}_{\text{nn}}(\mathbf{R}) + \mathcal{V}_{\text{ne}}(\mathbf{r}, \mathbf{R}) + \mathcal{V}_{\text{ee}}(\mathbf{r}), \quad (9)$$

and $\Psi_k(\mathbf{r}; \mathbf{R})$ is the electronic wave function of state k . Meanwhile, nuclear motion is described by

$$[\mathcal{T}(\mathbf{R}) + E_k(\mathbf{R})]\chi_k = i\hbar \frac{\partial}{\partial t} \chi_k \quad (10)$$

with the nuclear wave function $\chi_k(\mathbf{R}, t)$ evolving on the potential energy surface $E_k(\mathbf{R})$ of the electronic state k . The total wave function is then the direct product of the electronic and the nuclear wave function,

$$\Phi(\mathbf{r}, \mathbf{R}; t) = \Psi_k(\mathbf{r}, \mathbf{R})\chi_k(\mathbf{R}, t) \quad (11)$$

In the classical limit (Doltsinis and Marx 2002b), the nuclear wave equation (10) is replaced by Newton’s equation of motion

$$M_I \ddot{\mathbf{R}}_I = -\nabla_I E_k \quad (12)$$

For a great number of physical situations the Born–Oppenheimer approximation can be safely applied. On the other hand, there are many important chemical phenomena such as charge transfer and photoisomerization reactions, whose very existence is due to the inseparability of electronic and nuclear motion. Inclusion of nonadiabatic effects is beyond the scope of this chapter and the reader is referred to the literature (e.g., Refs. Doltsinis and Marx 2002b; Doltsinis 2006) for more details.

The above approximations form the basis of conventional molecular dynamics, Eqs.(22) together with (8) being the working equations. Thus, in principle, a classical trajectory calculation merely amounts to integrating Newton’s equations of motion (22). In practice, however, this deceptively simple task is complicated by the fact that the stationary Schrödinger equation (8) cannot be solved exactly for any many-electron system. The potential energy surface therefore has to be approximated using ab initio electronic structure methods or empirical interaction potentials (the so-called force-field molecular dynamics (Sutmann 2002; Allen and Tildesley 1987)). The former approach, usually referred to as ab initio molecular dynamics

(AIMD), will be the subject of section “Ab Initio Molecular Dynamics,” while the latter – force-field molecular dynamics – will be discussed in section “Classical Molecular Dynamics.”

Ab Initio Molecular Dynamics

In the following, we shall focus on first-principles molecular dynamics methods. Due to the high computational cost associated with ab initio electronic structure calculations of large molecules, computation of the entire potential energy surface prior to the molecular dynamics simulation is best avoided. A more efficient alternative is the evaluation of electronic energy and nuclear forces “on the fly” at each step along the trajectory.

Born–Oppenheimer Molecular Dynamics

In the so-called Born–Oppenheimer implementation of such a scheme (Marx and Hutter 2000), the nuclei are propagated by integration of Eq. (22), where the exact energy E_k is replaced with the eigenvalue, \tilde{E}_k , of some approximate electronic Hamiltonian, $\tilde{\mathcal{H}}_{\text{el}}$, which is calculated at each time step. For the electronic ground state, i.e., $k = 0$, the use of Kohn–Sham (KS) density functional theory (Parr and Yang 1989; Dreizler and Gross 1990) has become the method of choice.

Car–Parrinello Molecular Dynamics

In order to further increase computational efficiency, Car and Parrinello have introduced a technique to bypass the need for wave function optimization at each molecular dynamics step (Car and Parrinello 1985; Marx and Hutter 2000). Instead, the molecular wave function is dynamically propagated along with the atomic nuclei according to the equations of motion

$$M_I \ddot{\mathbf{R}}_I = -\nabla_I \langle \Psi_k | \tilde{\mathcal{H}}_{\text{el}} | \Psi_k \rangle \quad (13)$$

$$\mu_i \ddot{\psi}_i = -\frac{\delta}{\delta \psi_i^*} \langle \Psi_k | \tilde{\mathcal{H}}_{\text{el}} | \Psi_k \rangle + \sum_j \lambda_{ij} \psi_j, \quad (14)$$

where the KS one-electron orbitals ψ_i are kept orthonormal by the Lagrange multipliers λ_{ij} . These are the Euler–Lagrange equations

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{\partial \mathcal{L}}{\partial q}, \quad (q = \mathbf{R}_I, \psi_i^*) \quad (15)$$

for the Car–Parrinello Lagrangian (Car and Parrinello 1985)

$$\mathcal{L} = \sum_I \frac{1}{2} M_I \dot{\mathbf{R}}_I^2 + \sum_i \frac{1}{2} \mu_i \langle \dot{\psi}_i | \dot{\psi}_i \rangle - \langle \Psi_k | \tilde{\mathcal{H}}_{\text{el}} | \Psi_k \rangle + \sum_{ij} \lambda_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij}) \quad (16)$$

that is formulated here for an arbitrary electronic state Ψ_k , an arbitrary electronic Hamiltonian $\hat{\mathcal{H}}_{\text{el}}$, and an arbitrary basis (i.e., without invoking the Hellmann–Feynman theorem).

Classical Molecular Dynamics

While first-principles molecular dynamics simulations deal with the electrons in a system, this results in a large number of particles and inevitably requires the use of quantum mechanics. Therefore, the calculations become significantly time-consuming. Classical molecular dynamics ignore electronic motions and calculate the energy of a system as a function of the nuclear positions only and therefore are used to simulate larger, less detailed systems for larger time scales. The successive configurations of the system are generated by solving the differential equations that constitute Newton's second law (Eq. 22):

$$\frac{d^2 X_I}{dt^2} = \frac{F_{X_I}}{M_I} \quad (17)$$

This equation describes the motion of particle I of mass M_I along one dimension (X_I), where F_{X_I} is the force on the particle in that dimension. The solution of these differential equations results in a trajectory that specifies how the positions and velocities of the particles in the system vary with time.

In realistic models of intermolecular interactions, the force on particle I changes whenever particle I changes its position or whenever another atom with which particle I interacts changes its position. Therefore, the motions of all the particles are coupled together, which results in a many-body problem that cannot be solved analytically. Therefore, finite difference methods are used to integrate the equations of motion.

Generally, the integration of Eq. (17) is broken into consecutive steps that are conducted at different times t that are separated by increments of δt , which is generally referred to as the time step. First, the total force on each particle in the system at time t is calculated as the vector sum of its interactions with other particles.

Then, assuming the force is constant over the course of the time step, the accelerations of the particles are calculated, which are then combined with positions and velocities of the particles at time t to determine the positions and velocities at time $t + \delta t$. Finally, the forces on the particles in their new positions are determined, and then new accelerations, positions, and velocities are determined at $t + 2\delta t$ and so on.

A common approach in the various finite difference methods used to integrate the equations of motions for classical molecular dynamics simulations is that it is assumed that the positions, velocities, and accelerations (as well as all other dynamic properties) can be approximated using Taylor series expansions:

$$\mathbf{R}(t + \delta t) = \mathbf{R}(t) + \delta t \mathbf{V}(t) + \frac{1}{2} \delta t^2 \mathbf{A}(t) + \frac{1}{6} \delta t^3 \mathbf{B}(t) + \frac{1}{24} \delta t^4 \mathbf{C}(t) + \dots \quad (18)$$

$$\mathbf{V}(t + \delta t) = \mathbf{V}(t) + \delta t \mathbf{A}(t) + \frac{1}{2} \delta t^2 \mathbf{B}(t) + \frac{1}{6} \delta t^3 \mathbf{C}(t) + \dots \quad (19)$$

$$\mathbf{A}(t + \delta t) = \mathbf{A}(t) + \delta t \mathbf{B}(t) + \frac{1}{2} \delta t^2 \mathbf{C}(t) + \dots \quad (20)$$

where \mathbf{R} is the position, \mathbf{V} ($= \dot{\mathbf{R}}$) is the velocity, \mathbf{A} ($= \ddot{\mathbf{R}}$) is the acceleration, and \mathbf{B} and \mathbf{C} are the third and fourth derivatives of the positions with respect to time, respectively.

Verlet Algorithm

One of the most widely used finite difference methods in classical molecular dynamics simulations is the Verlet algorithm (Verlet 1967). In the Verlet algorithm, the positions and accelerations at time t and the positions from the previous time step $\mathbf{R}(t - \delta t)$ are used to calculate the updated positions $\mathbf{R}(t + \delta t)$ using the equation:

$$\mathbf{R}(t + \delta t) = 2\mathbf{R}(t) - \mathbf{R}(t - \delta t) + \delta t^2 \mathbf{A}(t). \quad (21)$$

The accelerations can be conveniently calculated using Newton's equation (17), where the forces are obtained by differentiating the energy with respect to the nuclear positions,

$$\mathbf{A} = \frac{\mathbf{F}}{M} = -\frac{\nabla E}{M} \quad (22)$$

While the velocities do not explicitly appear in Eq. (21), they can be calculated from the difference in position over the entire time step:

$$\mathbf{V}(t) = |\mathbf{R}(t + \delta t) - \mathbf{R}(t - \delta t)| / 2\delta t \quad (23)$$

or the difference in position over a half time step ($t + \frac{1}{2}\delta t$):

$$\mathbf{V}(t + \frac{1}{2}\delta t) = |\mathbf{R}(t + \delta t) - \mathbf{R}(t)| / \delta t \quad (24)$$

The fact that the velocities are not explicitly represented in the Verlet algorithm is one of the drawbacks to this method in that no velocities are available until the positions have been determined at the next time step. Also, in order to calculate the position of particles at $t = \delta t$, it is necessary to determine the positions at $t = -\delta t$ since the algorithm requires the position at time $t - \delta t$ to calculate the position at time $t + \delta t$. Often, this drawback is overcome by using the Taylor series to calculate $\mathbf{R}(-\delta t) = \mathbf{R}(0) - \delta t \mathbf{V}(0) + \frac{1}{2} \delta t^2 \mathbf{A}(0) + \dots$. A final drawback of the Verlet algorithm is that there may be a loss of precision in the resulting trajectories that result from the fact that the positions are calculated by adding a small term ($\delta t^2 \mathbf{A}(t)$) to the difference of two larger terms ($2\mathbf{R}(t)$ and $\mathbf{R}(t - \delta t)$) in equation (21).

“Leap-Frog” Algorithm

In an attempt to improve upon the original Verlet algorithm, several variations have been developed. The leap-frog algorithm (Hockney 1970) is one of the variations that use the following equations to update the positions:

$$\mathbf{R}(t + \delta t) = \mathbf{R}(t) + \delta t \mathbf{V}(t + \frac{1}{2} \delta t), \quad (25)$$

and the velocities:

$$\mathbf{V}(t + \frac{1}{2} \delta t) = \mathbf{V}(t - \frac{1}{2} \delta t) + \delta t \mathbf{A}(t). \quad (26)$$

In the leap-frog algorithm, the velocities $\mathbf{V}(t + \frac{1}{2} \delta t)$ are first calculated from the velocities at time $t - \frac{1}{2} \delta t$ and the accelerations at time t using Eq. (25). Then the positions $\mathbf{R}(t + \delta t)$ are calculated from the velocities $\mathbf{V}(t + \frac{1}{2} \delta t)$ and the positions $\mathbf{R}(t)$ using Eq. (26). The algorithm gets its name from the fact that the velocities are calculated in a manner such that they “leap-frog” over the positions to give their values $t - \frac{1}{2} \delta t$. Then the positions are calculated such that they “leap-frog” over the velocities, and then the algorithm continues.

The “leap-frog” algorithm improves upon the standard Verlet algorithm in that the velocity is explicitly included in the calculations and also the “leap-frog” algorithm does not require the calculation of the differences of large numbers so the precision of the calculation should be improved. However, the fact that the calculated velocities and positions are not synchronized in time results in the fact that the kinetic energy contribution to the total energy cannot be calculated for the time at which the positions are defined. In response to this shortcoming in the “leap-frog” algorithm, a formalism to calculate the velocities at time t has been developed that follows

$$\mathbf{V}(t) = [\mathbf{V}(t + \frac{\delta t}{2}) + \mathbf{V}(t - \frac{\delta t}{2})]/2 \quad (27)$$

Velocity Verlet Algorithm

The velocity Verlet method (Swope et al. 1982), which is a variation of the standard Verlet method, calculates the positions, velocities, and accelerations at the same time by using the following equations:

$$\mathbf{R}(t + \delta t) = \mathbf{R}(t) + \delta t \mathbf{V}(t) + \frac{1}{2} \delta t^2 \mathbf{A}(t) \quad (28)$$

$$\mathbf{V}(t + \delta t) = \mathbf{V}(t) + \frac{1}{2} \delta t [\mathbf{A}(t) + \mathbf{A}(t + \delta t)]. \quad (29)$$

The velocity Verlet method is a three-stage algorithm because the calculation of the new velocities (Eq. 29) requires both the acceleration at time t and at time $t + \delta t$.

Therefore, first, the positions at $t + \delta t$ are calculated using Eq. (28) and the velocities and accelerations at time t . The velocities at time $t + \frac{1}{2}\delta t$ are then calculated using

$$\mathbf{V}(t + \frac{1}{2}\delta t) = \mathbf{V}(t) + \frac{1}{2}\delta t \mathbf{A}(t). \quad (30)$$

Then the forces are computed from the current positions, which results in being able to calculate $\mathbf{A}(t + \delta t)$. Then the final step consists of calculating the velocities at time $t + \delta t$ using:

$$\mathbf{V}(t + \delta t) = \mathbf{V}(t + \frac{1}{2}\delta t) + \frac{1}{2}\delta t \mathbf{A}(t + \delta t). \quad (31)$$

Therefore, the velocity Verlet allows for the velocities and positions to be calculated in a time-synchronized manner and thus allows for the kinetic energy contribution of the total energy. Also, the precision of the results will be improved upon those from the standard Verlet algorithm as there are no differences of large numbers within the formalism of the method.

The selection of the best time integration method for a given problem and the size of the time step to use will be discussed in section “Setting the Timestep.”

Hybrid Quantum/Classical (QM/MM) Molecular Dynamics

The *ab initio* and classical simulation techniques discussed in the previous sections can be viewed as complementary. While AIMD is capable of dealing with electronic processes such as chemical reactions, charge transfer, and electronic excitations, its applicability is limited to systems of modest size, precluding its use in complex, large-scale biochemical simulations. Classical MD, on the other hand, can describe much larger systems on longer time scales but misses any of the abovementioned electronic effects, e.g., bond breaking and formation. The basic idea of the QM/MM approach is to combine the strengths of the two methods treating a chemically active region at the quantum level and the environment using molecular mechanics (i.e., a force field). There are several excellent review articles on the QM/MM method in the literature (Senn and Thiel 2009; Thiel 2009; Groenhof 2013).

Partitioning Schemes

The entire system, **S**, is partitioned into a chemically active inner region, **I**, and a chemically inert outer region, **O**. If the border between these regions cuts through chemical bonds, so-called link atoms, **L**, are usually introduced to cap the inner region (see section “Bonds Across the QM/MM Boundary”).

Subtractive Scheme

In a subtractive scheme, the total energy, $E_{\text{QM/MM}}^{\text{S}}$, of the entire system,

$$E_{\text{QM/MM}}^{\text{S}} = E_{\text{MM}}^{\text{S}} + E_{\text{QM}}^{\text{I,L}} - E_{\text{MM}}^{\text{I,L}} \quad (32)$$

is calculated from three separate energy contributions: (i) the MM energy of the entire system, E_{MM}^{S} , (ii) the QM energy of the active region (including any link atoms), $E_{\text{QM}}^{\text{IL}}$, and (iii) the MM energy of the active region $E_{\text{MM}}^{\text{IL}}$.

The role of the third term in Eq. (32) is to avoid double counting and to correct for any artifacts caused by the link atoms. For the latter to be effective, the force field has to reproduce the quantum-mechanical forces reasonably well in the link region.

Additive Scheme

In an additive scheme, the total energy of the system is given by

$$E_{\text{QM/MM}}^{\text{S}} = E_{\text{MM}}^{\text{O}} + E_{\text{QM}}^{\text{IL}} + E_{\text{QM-MM}}^{\text{IO}} \quad (33)$$

The difference to the subtractive scheme is that here a pure MM calculation is performed for only the outer region and the interaction between QM and MM regions is achieved by an explicit coupling term,

$$E_{\text{QM-MM}}^{\text{IO}} = E_{\text{QM-MM}}^{\text{bond}} + E_{\text{QM-MM}}^{\text{vdW}} + E_{\text{QM-MM}}^{\text{el}} \quad (34)$$

where $E_{\text{QM-MM}}^{\text{bond}}$, $E_{\text{QM-MM}}^{\text{vdW}}$, and $E_{\text{QM-MM}}^{\text{el}}$ are bonded, van der Waals, and electrostatic interaction energies, respectively.

The simplest way to treat electrostatic interactions between the **I** and **O** subsystems is to assign fixed electric charges to all **I** atoms (*mechanical embedding*). In this case the QM problem is solved for the isolated subsystem **I** without taking into account the effects of the surrounding atomic charges in **O**. The majority of implementations use an *electrostatic embedding* scheme in which the MM point charges of region **O** are incorporated in the QM Hamiltonian through a QM-MM coupling term,

$$\hat{H}_{\text{QM-MM}}^{\text{el}} = - \sum_i \sum_{\alpha \in \text{O}} \frac{q_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} + \sum_{I \in \text{I+L}} \sum_{\alpha \in \text{O}} \frac{q_\alpha Z_I}{|\mathbf{R}_I - \mathbf{R}_\alpha|} \quad (35)$$

where q_α are the MM point charges at positions \mathbf{R}_α (all other symbols as defined in section “Born–Oppenheimer Approximation”). In this way, the electronic structure of the QM region adjusts to the moving MM charge distribution. A problem that arises when an MM point charge is in close proximity to the QM electron cloud is overpolarization of the latter, sometimes referred to as a “spill-out” effect. This can be avoided by modifying the Coulomb potential in the first term of Eq. (35) at short range (see, e.g., Laio et al. 2002).

At present, in all commonly used partitioning schemes, the partitions remain *fixed* over time, i.e., an MM atom cannot turn into a QM atom and vice versa. This can present a serious limitation, for instance, in the case of solvent diffusion through the chemically active region. A number of *adaptive* partitioning methods have been proposed to remedy this problem (Kerdcharoen et al. 1996; Hofer et al. 2005; Kerdcharoen and Morokuma 2002; Heyden et al. 2007; Buló et al. 2009; Nielsen et al. 2010; Pezeshki and Lin 2011; Pezeshki et al. 2014; Waller et al. 2014);

however, the computational overhead is significant and physical conservation laws are not strictly satisfied.

Bonds Across the QM/MM Boundary

Partitioning the total system into QM and MM regions in such a way that cuts chemical bonds is best avoided. However, in many cases this is inevitable. Then one has to make sure that any atoms participating in chemical reactions are at least three bonds away from boundary. Furthermore, it is preferable to cut a bond that is unpolar and not part of a conjugated chain.

Link Atoms

Cutting a single covalent bond will create a dangling bond which must be capped by a so-called link atom, and in most applications a hydrogen atom is chosen. In the QM calculation, the atoms of region **I** together with the link atoms **L** are treated as an isolated molecule in the presence of the point charges of the environment **O**. The original QM–MM bond, cut by the partitioning, is only treated at the MM level.

Boundary Atoms

Boundary atom schemes have been developed to avoid the artifacts introduced by a link atom. The boundary atom appears as a normal MM atom in the MM calculation while carrying QM features to saturate the QM–MM bond and to mimic the electronic properties of the MM side. The QM interactions are achieved by placing a pseudopotential at the position of the boundary atom, parametrized to reproduce electronic properties of a certain chemical end group, e.g., a methyl group in the case of a cut C–C bond. Among the various flavors that have been proposed, the pseudobond method for first principles QM calculations (Zhang et al. 1999; Zhang 2005, 2006) and the pseudopotential approach for plane-wave DFT (Laio et al. 2002) are the most relevant in the present context.

Frozen Localized Orbitals

The basic idea behind the various frozen orbital methods (Warshel and Levitt 1976; Théry et al. 1994; Monard et al. 1996; Assfeld and Rivail 1996; Assfeld et al. 1998; Ferré et al. 2002; Fornili et al. 2003, 2006a, b; Sironi et al. 2007; Loos and Assfeld 2007; Philipp and Friesner 1999; Murphy et al. 2000; Gao et al. 1998; Amara et al. 2000; Garcia-Viloca and Gao 2004; Pu et al. 2004a, b, 2005; Jung et al. 2007; Jensen et al. 1994; Day et al. 1996; Kairys and Jensen 2000; Gordon et al. 2001; Nemukhin et al. 2002, 2003; Grigorenko et al. 2002) is to saturate the cut QM–MM bond by placing on either the MM or the QM atom at the boundary localized orbitals that have been determined in a prior quantum-mechanical SCF calculation on a model molecule containing the bond under consideration. To preserve the properties of the bond, the localized orbitals are then kept fixed in the subsequent QM/MM calculation. Different flavors are the local SCF (LSCF) method (Théry et al. 1994; Monard et al. 1996; Assfeld and Rivail 1996; Assfeld et al. 1998; Ferré et al. 2002), extremely localized molecular orbitals (ELMOs) (Fornili et al. 2003, 2006a; Sironi et al. 2007), frozen core orbitals (Fornili et al. 2006b), optimized

LSCF (Loos and Assfeld 2007), frozen orbitals (Philipp and Friesner 1999; Murphy et al. 2000), generalized hybrid orbitals (Gao et al. 1998; Amara et al. 2000; Garcia-Viloca and Gao 2004; Pu et al. 2004a, b, 2005; Jung et al. 2007), and effective fragment potentials (EFP) (Jensen et al. 1994; Day et al. 1996; Kairys and Jensen 2000; Gordon et al. 2001; Nemukhin et al. 2002, 2003; Grigorenko et al. 2002).

Of the three types of boundary treatment, the link atom method is the simplest both conceptually and in practice and is hence the most widely used. The boundary atom and in particular the frozen orbital methods can potentially achieve higher accuracy but require careful a priori parametrization and bear limitations on transferability (Senn and Thiel 2009).

Coarse-Grained Molecular Dynamics

A large number of important problems in fields that are often studied using molecular dynamics simulations (i.e., soft condensed matter physics, structural biology, chemistry, and materials science) take place over a time span of microseconds to seconds and distances of few hundred nanometers to a few microns. However, these time and length scales are still unattainable via quantum or atomistic force-field molecular dynamics methods despite significant computational hardware advances (Shirts and Pande 2000; Mervis 2001; Reed 2003) and the development of increasingly powerful software (MacKerell et al. 1998; Lindahl and van der Spoel 2001; Wang et al. 2004a; Phillips et al. 2005). Therefore, one approach that has been utilized in order to be able to study these complex problems is to reduce the computational demand of the simulation by reducing the number of atoms represented and therefore the degrees of freedom of the simulated system. This procedure of reducing the number of atoms represented in a system is done by grouping atoms together and representing them as a single interaction site and is generally referred to as “coarse graining” of the system. Figure 1 shows a comparison of the atomistic, united-atom, and coarse-grained representation.

The “bead-spring” coarse-grained model of polymer chains that was created by Kremer and Grest in 1990 has served as the foundation for many of the coarse-grained models that have been developed for a wide range of phenomena (at the current date this paper has been cited over 1440 times) including various studies of polymers and biomolecules including DNA solutions. Many of the more recent coarse-grained models have been developed for biological macromolecules since there are many examples of interesting biophysical phenomena that occur at large length and time scales. The most widely used coarse-grained models for biological systems include the generic model of Lipowsky et al. (Goetz et al. 1999; Shillcock and Lipowsky 2002), the solvent-free model of Deserno et al. (Cooke et al. 2005), and the specific models of the Klein group (Shelley et al. 2001), the Voth group (which is called the multi-scale coarse-grained model) (Izvekov and Voth 2005, 2006), the Essex group (called the ELBA force-field) (Orsi et al. 2011), and the Marrink group (called the MARTINI force-field) (Marrink et al. 2007). The above coarse-grained models have generally been developed for lipid

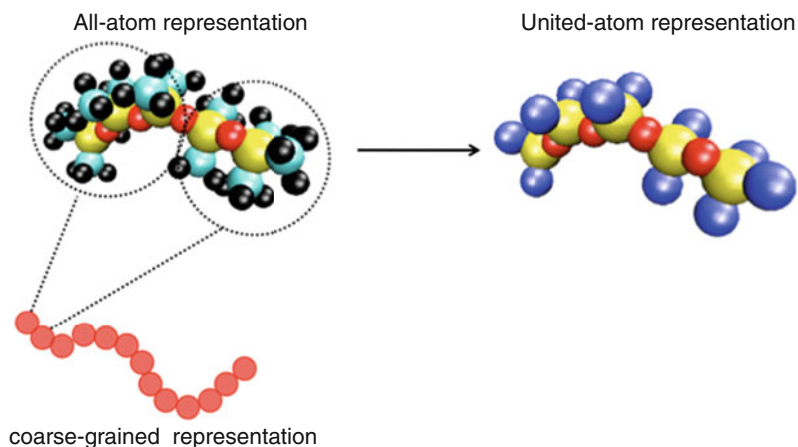


Fig. 1 Atomistic, united-atom, and coarse-grained representations of organic molecules

membranes; however, there are also coarse-grained force fields for proteins (as reviewed in Tozzini (2005) and some more recent examples Betancourt and Omovie (2009), Bereau and Deserno (2009), de Jong et al. (2013), and Kar et al. (2013)), carbohydrates (Lopez et al. 2009), and DNA (Tepper and Voth 2005; Khalid et al. 2008).

When developing a coarse-grained model for a system, there are two important decisions to be made: (1) how many atoms to combine (coarse grain) into a single interaction site and (2) how to parameterize the coarse-grained force field. In deciding the number of atoms to combine into a single interaction site, one must consider the obvious trade-off of how much detail are you able to sacrifice in order to simulate larger length and/or time scale phenomena and still be able to actually accurately model the phenomena of interest. The least amount of coarse graining that has been used is represented by what is called a “united-atom” representation of a molecule where all “heavy” atoms (generally all non-hydrogen elements in a molecule) are represented and the “light” (i.e., hydrogen) atoms are grouped with the heavy atom to which they are bonded into one interaction site. United-atom versions of many of the popular all-atom force fields listed are in section “Classical Force Fields” exist and have been successfully used in several studies. In addition to united-atom models, there are several existing coarse-graining methods that will combine a different number of atoms together into one interaction site.

A mapping function C is required to develop a description of the desired system in terms of the coarse-grained interaction sites. This mapping function uses a set of n atomic coordinates in the atomistic system ($\mathbf{r}^n = \mathbf{r}_1, \dots, \mathbf{r}_n$) and maps them to a unique configuration of coordinates for N interaction sites ($\mathbf{R}^N = \mathbf{R}_1, \dots, \mathbf{R}_N$) in the coarse-grained system. Then the mapping function generally has the form:

$$\mathbf{R} = \mathbf{C}\mathbf{r} \quad (36)$$

where \mathbf{R} and \mathbf{r} are vectors of length $3N$ and $3n$, respectively, and \mathbf{C} is a matrix with dimensions $3N \times 3n$ and contains the elements c_{Ii} that are used to link the descriptions of each coarse-grained interaction site, I , to the atoms in the all-atom configuration it represents:

$$\mathbf{R}_I = \sum_{i=1}^n c_{Ii} \mathbf{r}_i. \quad (37)$$

The most common form of mapping is where the coarse-grained interaction sites are mapped to the center of mass of the atoms within the all-atom configuration, and therefore $c_{Ii} = m_i / \sum_{i=1}^n m_i$.

In general, coarse-grained systems are governed by similar potential terms as are found in atomistic models such as nonbonded terms (both pairwise interactions and electrostatic interactions) and bond-stretching terms, and then in more sophisticated models even angle and dihedral terms will be included as well. Generally, all specific models are parameterized based on comparison to atomistic simulations and/or detailed experimental data. The “top-down” approach generally parameterizes the interactions without explicit consideration of a more detailed model. The interactions are often determined using physicochemical intuition or generic physical principles or certain emergent structural or thermodynamic properties that are observed on even larger scales. The most popular example of a “top-down” coarse-grained model is the Martini model (Marrink et al. 2004; de Jong et al. 2013; Lopez et al. 2009). The Martini model provides transferable potentials that describe effects of hydrophobic, van der Waals, and electrostatic interactions between sites, which are characterized based on their polarity and charge. The Martini potentials have been optimized to reproduce the partitioning of model compounds between aqueous and hydrophobic environments (Marrink et al. 2004, 2007). For more information about developing Martini force fields for molecules which have as of yet to be modeled, please refer to this website (Martini <http://md.chem.rug.nl/cgmartini/index.php/parametrizing-new-molecule>).

Effective coarse-grained potentials have also been extracted from atomistic simulations (often referred to as “bottom-up” approaches) using a variety of different methods including, but not limited to, those that we briefly introduce in this chapter as they are currently among the most popular methods: inverse Iterative Boltzmann Inversion approaches (Muller-Plathe 2002; Reith et al. 2003), Monte Carlo schemes (Lyubartsev 2005; Elezgaray and Laguerre 2006) and force matching (Ercolessi and Adams 1994; Izvekov and Voth 2005, 2006; Lu et al. 2013).

Iterative Boltzmann Inversion (IBI) is one of the most commonly used bottom-up approaches to parameterizing coarse-grained potentials because it utilizes a straightforward and generally robust algorithm. In general, this method is used to construct nonbonded pairwise additive potential functions between every type of interaction site in a given system. The potential functions are iteratively refined in order to match target radial distribution functions (see section “Spatial Distribution Functions”) that are determined from atomistic simulations, whose trajectories have been mapped into the coarse-grained representation of the systems. The general

philosophy behind the IBI method is similar to the philosophy behind deriving atomistic pair potentials which have been used to reproduce structure factors determined from neutron scattering experiments (Schommers 1973; Soper 1996). In the iterative process of refining the pair potentials, the initial guesses for each of these interactions are calculated from the potential of mean force, U_{PMF} , from the target radial distribution functions, $g(r)$:

$$U_{\text{PMF}}(r) = -k_B T \ln[g(r)] \quad (38)$$

where k_B and T are the Boltzmann constant and temperature, respectively. The intramolecular potential energies are calculated using the observed distributions P_q of each type of interaction (i.e., bond lengths, bond angles, and dihedral angles) from the reference atomistic simulations using an analogous approach as used for the pair potentials in Eq. (38):

$$U(q) = -k_B T \ln[P_q(q)] \quad (39)$$

where q is the bond length, bond angle, or dihedral angle, for the various terms of the intramolecular potential energy. Then using the initial guess for the nonbonded pair potentials (Eq. (38)) and these intramolecular potential energies, a coarse-grained simulation is performed and radial distribution functions $g_{\text{CG}}^n(r)$ are calculated and compared to the target radial distribution functions $g_t(r)$ to update the pair potential energy functions for the next iteration of the algorithm:

$$U^{(n+1)}(r) = U^n(r) + \lambda \Delta U^n(r) = U^n(r) + \lambda k_B T \ln \frac{g_{\text{CG}}^n(r)}{g_t(r)} \quad (40)$$

where λ scales the potential update, and should be chosen in order to ensure stability of the algorithm. This iterative process is repeated until the measured radial distribution functions in the CG simulations are within a user-defined tolerance of the target radial distribution functions. This method has been automated in a software package called VOTCA (<http://www.votca.org/home>) that can be used in combination with the popular molecular dynamics simulation packages GROMACS (Spoel et al. 2005; <http://www.gromacs.org/>), LAMMPS (<http://lammps.sandia.gov/>; Plimpton 1995), and DL-POLY (Smith et al. 2002; http://www.cse.scitech.ac.uk/ccg/software/DL_POLY/).

The Inverse Monte Carlo (IMC) algorithm is another iterative method that corrects a guessed interaction potential to reproduce a target radial distribution function. However, the IMC method applies a correction to the potential which is determined from statistical mechanical arguments that account for the fact that small changes in the potential energies at a given distance r can lead to variation in the radial distribution functions at all distances. Generally speaking, one can summarize this approach by stating that the Hamiltonian (potential energy) H of the system depends on a series of parameters $\lambda_1, \lambda_2, \dots, \lambda_M$, and the desire is to reproduce canonical averages of a given measurable $\langle S_1 \rangle, \langle S_2 \rangle, \dots, \langle S_M \rangle$

(e.g., radial distribution functions). In Inverse Monte Carlo, the desire is to calculate the values of the parameters used to calculate the potential energy λ_M from the canonical averages $\langle S_M \rangle$. In order to do so, in the vicinity of an arbitrary point in the space of the Hamiltonians, it can be shown that:

$$\Delta \langle S_\alpha \rangle = \sum_\gamma \frac{\delta \langle S_\alpha \rangle}{\delta \lambda_\gamma} \Delta \lambda_\gamma \quad (41)$$

where

$$\frac{\delta \langle S_\alpha \rangle}{\delta \lambda_\gamma} \Delta \lambda_\gamma = \left\langle \frac{\delta S_\alpha}{\delta \lambda_\gamma} \right\rangle - \beta \left[\left\langle \frac{\delta H}{\delta \lambda_\alpha} S_\gamma \right\rangle - \left\langle \frac{\delta H}{\delta \lambda_\alpha} \right\rangle \langle S_\gamma \rangle \right]. \quad (42)$$

Using Eqs. (41) and (42), then this problem can be solved iteratively.

The IBI and IMC methods are both designed to reproduce a target radial distribution function and therefore are commonly called structure-based coarse-graining methods. Another class of bottom-up coarse graining approaches are those that are designed to match the force distributions on the coarse-grained interaction sites to the force distributions measured within the atomistic trajectory on the atoms that are mapped into the various interaction sites. This force-matching approach was first used for obtaining atomistic potential energies by fitting forces and trajectories from ab initio calculations (Ercolessi and Adams 1994). It has since been extended to fitting coarse-grained potentials by Izvekov and Voth (Izvekov and Voth 2005, 2006; Lu et al. 2013), who call their method multi-scale coarse graining (MSCG). In this force-matching scheme, $\mathbf{f} = \mathbf{F}\boldsymbol{\phi}$, where \mathbf{f} is a vector composed of the atomistic force data. $\boldsymbol{\phi}$ is a vector containing all of the coarse-grained force-field parameters, and \mathbf{F} is a matrix that is related to the input atomistic configurations. Since this equation is a standard least squares matrix equation, it can be converted to

$$\mathbf{G}\boldsymbol{\phi} = \mathbf{b}, \quad (43)$$

where \mathbf{G} is the square matrix $\mathbf{F}^T\mathbf{F}$ and \mathbf{b} is the vector $\mathbf{F}^T\mathbf{f}$. Therefore, $\mathbf{G}\Delta\boldsymbol{\phi} = \delta\mathbf{b}$, which can be used to iterate the values of $\boldsymbol{\phi}$ in order to converge on the appropriate forces within the coarse-grained simulations.

We have attempted to introduce some of the more popular approaches that are utilized to develop coarse-grained potentials for use in simulating biological, polymeric, and colloidal systems in this section. However, by no means is this a complete summary of all of the approaches that have been used, so for more lengthy discussions of the various approaches, we would recommend some recent review articles including (Noid 2013; Saunders and Voth 2013; Brini et al. 2013; Riniker et al. 2012; de Pablo 2011). In general, the advantage of using a bottom-up approach is that the resulting force field will produce a higher level of accuracy and closer resemblance to atomistic simulations. However, these schemes produce force fields that are useful for a given state point and therefore are not transferable, whereas the advantages of the thermodynamic approach include that it produces a potential that has a broader range of applicability and also the thermodynamic approach does not require atomistic simulations to be done in the first place.

Interaction Potentials/Force Fields

Classical Force Fields

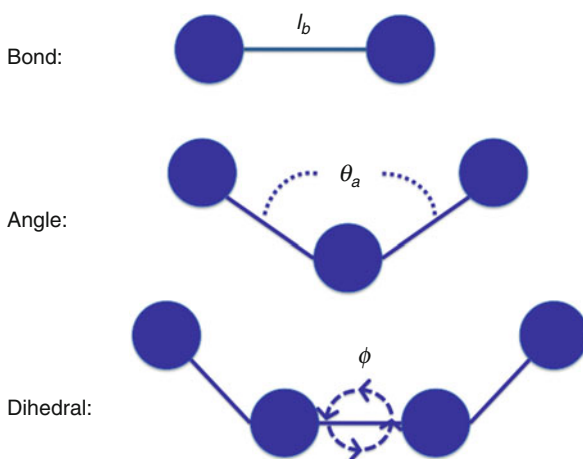
Classical, or empirical, force fields are generally used to calculate the energy of a system as a function of the nuclear positions of the particles within the system while ignoring the behavior of the individual electrons. As stated in section "Born–Oppenheimer Approximation," the Born–Oppenheimer approximation makes it possible to write the energy as a function of the nuclear coordinates. Another approximation that is key to the implementation of classical Force fields is that it is possible to model the relatively complex motion of particles within the system with fairly simple analytical models of inter- and intramolecular interactions. Generally, an empirical force field consists of terms that model the nonbonded interactions ($E_{\text{nonbonded}}$), which include both the van der Waals and Coulombic interactions, the bonded interactions (E_{bond}), the angle bending interactions (E_{angle}), and the dihedral (bond rotations) interactions (E_{dihedral}):

$$E(\mathbf{R}) = E_{\text{nonbonded}} + E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}}. \quad (44)$$

Figure 2 presents representative cartoons of the bond, angle, and dihedral interactions from a molecular perspective. The form that each of these individual terms takes is dependent on the Force field that you are using. There are several different force-field options available for various systems. The best way to find the most suitable force field for your specific problem is to conduct a literature and/or Internet search in order to find which force field has the capability to model the molecules you are interested in studying. However, if you are interested in modeling organic/biological molecules, there are several large force fields that may be a good place to start, including CHARMM (MacKerell et al. 1998), OPLS (Jorgensen and Swenson 1984), Amber (Cornell et al. 1995), and COMPASS (Sun et al. 1998). The CHARMM and AMBER force fields each have recently developed generalized versions of their force fields and tools to help researchers develop force-field parameters for small molecules that they may be interested in modeling for which parameters do not currently exist. In the AMBER force field, this general force field is called GAFF (generalized AMBER force field) (Wang et al. 2004b; <http://ambermd.org/antechamber/gaff.html>). While in the CHARMM force field, CGENFF is the generalized force field (Vanommeslaeghe et al. 2010; Yu et al. 2012; CGENFF <http://mackerell.umaryland.edu/~kenno/cgenff>; ParamChem <http://cgenff.paramchem.org>).

Likewise, there are several well-known large force fields that can be used for solids, like the BKS potential (van Beest et al. 1990) for oxides and the embedded-atom method (EAM) (Daw and Baskes 1983, 1984; Finnis and Sinclair 1984) and modified embedded-atom method (MEAM) (Baskes 1992) force fields. In addition to defining the functional forms used for the various terms in the general potential formulation, a force field will also define the variables used in the potential which

Fig. 2 Intramolecular terms of classical force fields: bond, angle, and dihedral interactions



are derived from a combination of quantum simulation results and experimental observations.

In the following sections, each of the terms in Eq. (44) will be discussed further and typical functional forms that are used in the previously mentioned force fields and others to represent each term will be shown. To this end, we limit the discussion to simple nonpolarizable force fields in which the individual atoms carry fixed charges. They capture many-body effects such as electronic polarization only in an effective way. More sophisticated, polarizable force fields have been developed over the past two decades (see, e.g., Ponder et al. 2010 and references therein); however, they are computationally substantially more demanding. Some of the most common polarizable force field methods will later be introduced at the end of section “Classical Polarizable Electrostatical Models”.

Nonbonded Interactions

There are two general forms of nonbonded interactions that need to be accounted for by a classical force field: (1) the van der Waals (vdw) interactions and (2) the electrostatic interactions.

van der Waals Interactions

In order to model the van der Waals interactions, we need a simple empirical expression that is not computationally intensive and that models both the dispersion and repulsive interactions that are known to act upon atoms and molecules. The most commonly used functional form of van der Waals energy (E_{vdw}) in classical force fields is the Lennard–Jones 12-6 function that has the form:

$$E_{\text{vdw}}(\mathbf{R}) = \sum_{I>J} 4\epsilon_{IJ} \left[\left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^{12} - \left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^6 \right], \quad (45)$$

where σ_{IJ} is the collision diameter and ϵ_{IJ} is the well depth of the interaction between atoms I and J . Both σ_{IJ} and ϵ_{IJ} are adjustable parameters that will have different values to describe the interactions between different pairs of particles (i.e., the values of σ and ϵ used to describe the interaction between two carbon atoms are different than the values of σ and ϵ used to describe the interaction between a carbon and an oxygen).

Equation (45) models both the attractive part (the R^{-6} term) and the repulsive part (the R^{-12} term) of the nonbonded interaction. Other formulations of the Lennard–Jones nonbonded potential commonly have the same power law description of the attractive part of the potential but will have different power law dependence for the repulsive part of the interaction, such as the Lennard–Jones 9-6 function:

$$E_{\text{vdW}}(\mathbf{R}) = \sum_{I>J} 4\epsilon_{IJ} \left[\left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^9 - \left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^6 \right]. \quad (46)$$

When the nonbonded interactions of a system that contains multiple particle types and multiple molecules are modeled using a Lennard–Jones-type nonbonded potential, it is necessary to be able to define the values of σ and ϵ that apply to the interaction between particles of type I and J . The parameters for these cross interactions are generally found using one of the two following mixing rules. One common mixing rule is the Lorentz–Berthelot rule where the value of σ_{IJ} is found from the arithmetic mean of the two pure values and the value of ϵ_{IJ} is the geometric mean of the two pure values:

$$\sigma_{IJ} = (\sigma_I + \sigma_J)/2 \quad (47)$$

$$\epsilon_{IJ} = \sqrt{\epsilon_I \epsilon_J} \quad (48)$$

The other commonly used mixing rule is the one that defines both σ_{IJ} and ϵ_{IJ} as the geometric mean of the values for the pure species:

$$\sigma_{IJ} = \sqrt{\sigma_I \sigma_J} \quad (49)$$

$$\epsilon_{IJ} = \sqrt{\epsilon_I \epsilon_J} \quad (50)$$

Most force fields use the Lorentz–Berthelot mixing rule; however, the OPLS force field is one force field that utilizes the geometric mixing rule.

In other nonbonded pairwise potentials, the repulsive portion of the interaction is modeled with an exponential term, which is in better agreement with the functional form of the repulsive term determined from quantum mechanics. One example of such a potential is the Buckingham potential (Buckingham 1938):

$$E_{\text{vdW}}(\mathbf{R}) = \sum_{I<J} A_{IJ} \exp(-B_{IJ} R_{IJ}) - \frac{C_{IJ}}{R_{IJ}^6}, \quad (51)$$

where A_{IJ} , B_{IJ} , and C_{IJ} are adjustable parameters that will have unique values for different types of particles. Another form of the nonbonded interaction is the Born–Mayer–Huggins potential (Fumi and Tosi 1964; Tosi and Fumi 1964):

$$E_{\text{vdw}}(\mathbf{R}) = \sum_{I < J} A_{IJ} \exp(B_{IJ}(\sigma_{IJ} - R_{IJ})) - \frac{C_{IJ}}{R_{IJ}^6} + \frac{D_{IJ}}{R_{IJ}^8}, \quad (52)$$

where A_{IJ} , B_{IJ} , C_{IJ} , D_{IJ} and σ_{IJ} are adjustable parameters that will have unique values for different types of particles. The Born–Mayer–Huggins potential (Eq. 52) is identical to the Buckingham potential (Eq. 51) when $\sigma = D = 0$.

All of the nonbonded potential functional forms that have been presented to this point are pairwise additive, i.e., they take into account the effect that one particle has on another particle based solely on the distance between the two particles. However, in some systems like metals and alloys as well as some covalently bonded materials like silicon and carbon, the nonbonded potential is a function of more than just the distance between two particles, i.e., many-body effects play an important role. In order to model these systems, the embedded-atom method (EAM) (Daw and Baskes 1983, 1984; Finnis and Sinclair 1984) and modified embedded-atom method (MEAM) (Baskes 1992) utilize an embedding energy, F_I , which is a function of the atomic electronic density ρ_I of the embedded atom I and a pair potential interaction ϕ_{IJ} such that:

$$E_I(\mathbf{R}) = F_I \left(\sum_{J \neq I} \rho_I(R_{IJ}) \right) + \frac{1}{2} \sum_{J \neq I} \phi_{IJ}(R_{IJ}). \quad (53)$$

The multi-body nature of the EAM potential is a result of the embedding energy term.

So while the EAM and MEAM potentials have a term to account for multi-body interactions, they are still only a pairwise potential, as are all the other nonbonded potentials presented to this point. However, there are multi-body potentials that will explicitly account for how the presence of a third, fourth, ... atom affects the nonbonded energy felt by any given atom. One example of a 3-body potential is the Stillinger–Weber potential (Stillinger and Weber 1985):

$$E(\mathbf{R}) = \sum_I \sum_{J > I} \phi_2(R_{IJ}) + \sum_I \sum_{J \neq I} \sum_{K > J} \phi_3(R_{IJ}, R_{IK}, \theta_{IJK}), \quad (54)$$

where there is a two-body term ϕ_2 :

$$\phi_2(R_{IJ}) = A_{IJ} \epsilon_{IJ} \left[B_{IJ} \left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^{p_{IJ}} - \left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^{q_{IJ}} \right] \exp \left(\frac{\sigma_{IJ}}{R_{IJ} - a_{IJ} \sigma_{IJ}} \right) \quad (55)$$

and a three-body term ϕ_3 :

$$\begin{aligned} \phi_3(R_{IJ}, R_{IK}, \theta_{IJK}) &= \lambda_{IJK} \epsilon_{IJK} \left[\cos \theta_{IJK} - \cos \theta_{0,IJK} \right]^2 \\ &\times \exp \left(\frac{\gamma_{IJ} \sigma_{IJ}}{R_{IJ} - a_{IJ} \sigma_{IJ}} \right) \\ &\times \exp \left(\frac{\gamma_{IK} \sigma_{IK}}{R_{IK} - a_{IK} \sigma_{IK}} \right). \end{aligned} \quad (56)$$

The Stillinger–Weber potential has generally been used for modeling crystalline silicon; however, more recently it has also been used for organic molecules as well. Another example of a three-body interatomic potential is the Tersoff potential (Tersoff 1988, 1989), which also was created initially in an attempt to accurately model silicon solids.

Electrostatic Interactions

Due to the fact that not all particles in a molecule have the same electronegativity, different particles will have stronger attractions to electrons than others. However, since classical force fields do not model the flow of electrons, the different particles within a molecule are assigned a partial charge that remains constant during the course of a simulation. Generally these partial charges q_I are assigned to the nuclear centres of the particles. The electrostatic interaction between particles in different molecules or particles that are separated by at least two other atoms in a given molecule is calculated as the sum of the contributions between pairs of these partial charges using Coulomb's law:

$$E_{\text{coul}} = \sum_I \sum_J \frac{q_I q_J}{4\pi \epsilon_0 R_{IJ}} \quad (57)$$

where the charges of each particle are q_I and q_J and ϵ_0 is the dielectric constant.

In practice, an Ewald sum (Ewald 1921) is generally used to evaluate the electrostatic interactions within a classical MD simulation. However, this is a very computationally expensive algorithm to implement and it results in a computational cost of $N^{3/2}$, where N is the number of particles in the system. In order to obtain better computational scaling, fast Fourier transforms (FFTs) have been used to calculate the reciprocal space summation required within the Ewald sum. By using the FFT algorithm, one can reduce the cost of the electrostatic algorithm to $N \log N$. The most popular FFT algorithm that has been adopted for use in classical MD simulations is the particle–particle particle–mesh (pppm) approach (Hockney and Eastwood 1981; Luty et al. 1994, 1995).

Classical Polarizable Electrostatical Models

In traditional classical molecular dynamics simulations with nonpolarizable force fields, the charge distribution is mimicked by the fixed partial charges. As a result the

electronic polarization effects are considered in a mean-field manner by optimizing the partial atomic charges such that they are representative of the condensed, typically aqueous, phase. In order to achieve a more accurate description of the response of the charge distribution to changes in the surrounding electrostatic field, models that explicitly account for electronic polarizability are essential.

There are several different approaches to including explicit electronic polarizability into classical force fields. The most computationally achievable methods to mimic the quantum-mechanical response of the electronic distribution within molecules to changes in the surrounding electric field is to extend the potential energy function such that it is capable of capturing this physics. There are three different general approaches most commonly used to do so and we will briefly discuss below: the *induced dipole model*, the *fluctuating charge model*, and the *classical Drude oscillator model*.

In the *induced dipole model*, generally an induced dipole, which is commonly calculated using a self-consistent field iterative procedure using isotropic atomic polarizabilities, is assigned to each atomic site (Warshel and Levitt 1976; Liu et al. 1998; Kaminski et al. 2002; Ren and Ponder 2002; Xie et al. 2007; Jorgensen et al. 2007; Shi et al. 2013). These induced dipoles are then used to calculate charge–dipole and dipole–dipole interactions, which are combined with the charge–charge interactions, to yield the electrostatic energy of the system. The induced dipole μ by an external electric field E from these models is calculated using $\mu = \alpha E$, where α is the polarizability of the atomic site.

The *fluctuating charge model* (Rick et al. 1994; Rick and Berne 1996; Stern et al. 2001; Patel and Brooks 2004; Patel et al. 2004) treats partial atomic charges as dynamical variables based on the electronegativity at each atomic site. Charges are propagated according to the principle of electronegativity equalization (Pauling and Yost 1932), and charge conservation constraints are assumed. Charges are generally allowed to only redistribute within molecules to avoid unphysical intermolecular charge transfer. During a molecular dynamics simulation, the charge variables used to model the fluctuating charges are usually propagated using an extended Lagrangian approach. In the fluctuating charge model, the induced dipole μ by an external electric field E is calculated by $\mu = (d^2/\eta)E$, where d is the distance between centers of two atoms and η is hardness of charge flow between them.

The *classical Drude oscillator model* (Lamoureux et al. 2003; Lamoureux and Roux 2003) is also commonly referred to as the shell (van Maaren and van der Spoel 2001) or charge-on-spring (Kunz and van Gunsteren 2009) model. In these models, a Drude particle or oscillator, which has a charge assigned to it, is attached to the center of the atom with which it is associated via a harmonic spring, and this setup is designed to mimic nuclear-electron motion (Dick and Overhauser 1958). In the Drude model, the induced dipole by an external electric field is calculated from $\mu = (q_D/k^2)E$, where q_D is the Drude charge and k is the spring constant of the harmonic spring that attaches the Drude particle to the center of its associated atom.

Each of these different models has been applied in various classical force fields in order to create polarizable versions of the force field. CHARMM has adopted the Drude model (Chowdhary et al. 2013; He et al. 2013; Patel et al. 2015; Savelyev

and MacKerell 2014; Lopes et al. 2013) and also has created a fluctuating charge model (Patel and Brooks 2004; Patel et al. 2004; Lucas et al. 2012; Zhong and Patel 2011). The induced dipole model has been used by the AMBER (Caldwell et al. 1990; Dang et al. 1991; Caldwell and Kollman 1995a, b) and AMOEBA (Ponder et al. 2010; Shi et al. 2013) force fields.

Bonded Interactions

The bonded interactions are needed to model the energetic penalty that will result from two covalently bonded atoms moving too close or too far away from one another. The most common functional form that is used to model the bond bending interactions is that of a harmonic term:

$$E_{\text{bond}} = \sum_{\text{bonds}} k_b (\ell_b - \ell_b^{(0)})^2 \quad (58)$$

where k_b is commonly referred to as the bond constant and is a measure of the bond stiffness and $\ell_b^{(0)}$ is the reference length or often referred to the equilibrium bond length. Each of these parameters will vary depending on the types of particles that the bond is joining.

Angle Bending Interactions

The angle bending interactions are also modeled in order to determine the energetic penalties of angles containing three different particles compressing or overextending such that they distort the geometry of a portion of a molecule away from its desired structure.

Again, the most common functional form to model the angle interactions is a harmonic expression:

$$E_{\text{angle}} = \sum_{\text{angles}} k_a (\theta_a - \theta_a^{(0)})^2 \quad (59)$$

where k_a is the angle constant and is a measure of the rigidity of the angle and $\theta_a^{(0)}$ is the equilibrium or reference angle.

Torsional Interactions

The torsional interactions are generally modeled using some form of a cosine series. The OPLS force-field uses the following expression for its torsional term:

$$E_{\text{dihed}} = \sum_{\text{dihedrals}} \frac{1}{2} K_d^{(1)} [1 + \cos(\phi)] + \frac{1}{2} K_d^{(2)} [1 - \cos(2\phi)] \\ + \frac{1}{2} K_d^{(3)} [1 + \cos(3\phi)] + \frac{1}{2} K_d^{(4)} [1 - \cos(4\phi)] \quad (60)$$

where $K_d^{(i)}$ are the force constants for each cosine term and ϕ is the measured dihedral angle. The CHARMM force field uses the following expression:

$$E_{\text{dihed}} = \sum_{\text{dihedrals}} K_d [1 + \cos(n\phi - d_d)], \quad (61)$$

where K_d is the force constant, n is the multiplicity of the dihedral angle ϕ , and d_d is the shift of the cosine that allows one to more easily move the minimum of the dihedral energy.

First-Principles Electronic Structure Methods

For the electronic ground state, i.e., $k = 0$, Kohn–Sham (KS) density functional theory is commonly used. In this case, the energy is given by

$$E_0 \approx E^{\text{KS}}[\rho] = T_s[\rho] + \int d\mathbf{r} v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} v_{\text{H}}(\mathbf{r})\rho(\mathbf{r}) + E_{\text{xc}}[\rho] \quad (62)$$

with the kinetic energy of *noninteracting electrons*, i.e., using a Slater determinant as a wave function ansatz,

$$\Psi^{\text{KS}} = \frac{1}{\sqrt{n!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_n(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_n(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_n) & \psi_2(\mathbf{x}_n) & \cdots & \psi_n(\mathbf{x}_n) \end{vmatrix} \quad (63)$$

$$T_s[\rho] = -\frac{1}{2} \sum_i^n f_i \int d\mathbf{r} \psi_i(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) \quad (64)$$

where f_i is the number of electrons occupying orbital ψ_i , the external potential including nucleus–nucleus repulsion and electron–nucleus attraction,

$$v_{\text{ext}}(\mathbf{r}) = \sum_{I=1}^{N-1} \sum_{J>I}^N \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} - \sum_{I=1}^N \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} \quad (65)$$

the Hartree potential (electron–electron interaction)

$$v_{\text{H}}(\mathbf{r}) = \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (66)$$

the exchange–correlation energy, E_{xc} , and the electron density

$$\rho(\mathbf{r}) = \sum_i^n f_i |\psi_i(\mathbf{r})|^2 \quad (67)$$

The orbitals which minimize the total, many-electron energy (62) are obtained by solving self-consistently the 1-electron Kohn–Sham equations,

$$\left[-\frac{1}{2} \nabla^2 + v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}(\mathbf{r}) + \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})} \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (68)$$

DFT is exact in principle, provided that $E_{xc}[\rho]$ is known, in which case E^{KS} (see Eq. (62)) is an exact representation of the ground state energy E_0 (see Eq. (8)). In practice, however, $E_{xc}[\rho]$ is not – and presumably never will be – known exactly; therefore, (semiempirical) approximations are used.

The starting point for most density functionals is the *local density approximation* (LDA), which is based on the assumption that one deals with a homogeneous electron gas. E_{xc} is split into an exchange term E_x and a correlation term E_c . Within the LDA the exchange functional is given exactly by Dirac (1930):

$$E_x^{\text{LDA}}[\rho] = \int \rho(\mathbf{r}) \epsilon_x^{\text{LDA}}(\rho(\mathbf{r})) d\mathbf{r} \quad (69)$$

where

$$\epsilon_x^{\text{LDA}}(\rho) = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{\frac{1}{3}} \rho(\mathbf{r})^{\frac{1}{3}} \quad (70)$$

The LDA correlation functional, on the other hand, can only be approximated. We give here the most commonly used expression by Vosko et al. (1980), derived from Quantum Monte Carlo calculations:

$$E_c^{\text{LDA}}[\rho] = \int \rho(\mathbf{r}) \epsilon_c^{\text{LDA}}(\rho(\mathbf{r})) d\mathbf{r} \quad (71)$$

where

$$\epsilon_c^{\text{LDA}}(\rho) = A \left\{ \ln \left(\frac{x^2}{X} \right) + \frac{2b}{Q} \tan^{-1} \left(\frac{Q}{2x+b} \right) - \frac{bx_0}{X(x_0)} \left[\ln \left(\frac{(x-x_0)^2}{X} \right) + \frac{2(b2x_0)}{Q} \tan^{-1} \left(\frac{Q}{2x+b} \right) \right] \right\} \quad (72)$$

with $X = x^2 + bx + c$, $x = \sqrt{r_s}$, $r_s = \sqrt[3]{\frac{3}{4\pi\rho(\mathbf{r})}}$, $Q = \sqrt{4c - b^2}$, $x_0 = -0.104098$, $A = 0.0310907$, $b = 3.72744$, $c = 12.9352$.

This simplest approximation, LDA, is often too inaccurate for chemically relevant problems. A notable improvement is usually offered by so-called *semilocal* or *gradient corrected* functionals (*generalized gradient approximation* (GGA)), in which E_x and E_c are expressed as functionals of ρ and the first variation of the density, $\nabla\rho$:

$$E_x^{\text{GGA}}[\rho, \nabla\rho] = \int \rho(\mathbf{r}) \epsilon_x^{\text{GGA}}(\rho(\mathbf{r}), \nabla\rho) d\mathbf{r} \quad (73)$$

$$E_c^{\text{GGA}}[\rho, \nabla\rho] = \int \rho(\mathbf{r}) \epsilon_c^{\text{GGA}}(\rho(\mathbf{r}), \nabla\rho) d\mathbf{r} \quad (74)$$

Popular examples are the BLYP (Becke 1988; Lee et al. 1988), BP (Becke 1988; Polák 1986), and BPW91 (Becke 1988; Perdew et al. 1992) functionals. The expressions for $\epsilon_{xc}^{\text{GGA}}(\rho(\mathbf{r}), \nabla\rho)$ are complex and shall not be discussed here.

In many cases, accuracy can be further increased by using so-called *hybrid functionals*, which contain an admixture of Hartree–Fock exchange to KS exchange. Probably the most widely used hybrid functional is the three-parameter B3LYP functional (Becke 1993),

$$E_{xc}^{\text{B3LYP}} = aE_x^{\text{LDA}} + (1-a)E_x^{\text{HF}} + b\Delta E_x^{\text{B}} + (1-c)E_c^{\text{LDA}} + cE_c^{\text{LYP}} \quad (75)$$

where $a = 0.80$, $b = 0.72$, $c = 0.81$, and E_x^{HF} is the Hartree–Fock exchange energy evaluated using KS orbitals.

New functionals are constantly proposed in search of better approximations to the exact E_{xc} . Often functionals are designed to remedy a particular shortcoming of previous functionals. For systems in which van der Waals interactions are important, for instance, Grimme’s empirical D3 dispersion correction (Grimme et al. 2010) may be used in combination with a GGA or hybrid functional. The total energy is then simply the sum of the DFT energy E^{KS} and the dispersion correction E^{disp} ,

$$E^{\text{DFT-D3}} = E^{\text{KS}} + E^{\text{disp}} \quad (76)$$

where E^{disp} contains 2-body and 3-body terms,

$$E^{\text{disp}} = E^{(2)} + E^{(3)} \quad (77)$$

Becke and Johnson proposed the following form for the 2-body contributions (Becke and Johnson 2005; Johnson and Becke 2005, 2006):

$$E^{(2)} = - \sum_{I>J} \sum_{n=6,8} s_n \frac{C_n^{IJ}}{R_{IJ}^n + f(R_{IJ}^0)^n} \quad (78)$$

with the damping function

$$f(R_{IJ}^0) = a_1 R_{IJ}^0 + a_2 \quad (79)$$

to avoid near singularities at short interatomic distances R_{IJ} , and

$$R_{IJ}^0 = \sqrt{\frac{C_8^{IJ}}{C_6^{IJ}}} \quad (80)$$

The functional-dependent scaling factors s_n and coefficients a_1 and a_2 can be obtained from the website of the Grimme group (www.thch.uni-bonn.de/tc) together with the dispersion coefficients C_n^{IJ} .

The reader is referred to a number of excellent textbooks (Parr and Yang 1989; Dreizler and Gross 1990; Koch and Holthausen 2000; Fiolhais et al. 2003) and online resources (Burke <http://dft.uci.edu/doc/g1.pdf>; Rappoport et al.) on DFT for further details on the theory and on how to choose the right functional for a specific application.

Building the System/Collecting the Ingredients

Setting Up an AIMD Simulation

Building a Molecule

In many cases, the coordinates of a molecular structure are available for download on the Web, from crystallographic databases (Reciprocal Net www.reciprocalnet.org/edumodules/commonmolecules; Cambridge Crystallographic Data Centre www.ccdc.cam.ac.uk; RCSB Protein Data Bank www.rcsb.org/pdb; Inorganic Crystal Structure Database www.fiz-karlsruhe.de/icsd.html; Toth Information Systems www.tothcanada.com) or journal supplements. For relatively small molecules, an initial guess structure can be built using molecular graphics software packages such as Molden (www.cmbi.ru.nl/molden/molden.html) or Avogadro (Hanwell et al. 2012; www.avogadro.cc).

Plane Waves and Pseudopotentials

The most common form of AIMD simulation employs DFT (see section “First-Principles Electronic Structure Methods”) to calculate atomic forces, in conjunction with periodic boundary conditions and a plane-wave basis set. Using a plane-wave basis has two major advantages over atom-centered basis functions: (i) there is no *basis set superposition error* (Boys and Bernardi 1970; Marx and Hutter 2000) and (ii) the Pulay correction (Pulay 1969, 1987) to the Hellmann–Feynman force, due to basis set incompleteness, vanishes (Marx and Hutter 2000, 2009).

Plane Wave Basis Set

As a consequence of Bloch’s theorem, in a periodic lattice, the Kohn–Sham orbitals (see Eq. 68) can be expanded in a set of plane waves (Ashcroft and Mermin 1976; Meyer 2006),

$$\psi_{\mathbf{k},j}(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{G}}^{\mathbf{k},j} e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}} \quad (81)$$

where \mathbf{k} is a wave vector within the Brillouin zone, satisfying Bloch’s theorem,

$$\psi(\mathbf{r} + \mathbf{T}) = e^{i\mathbf{k}\mathbf{T}} \psi(\mathbf{r}) \quad (82)$$

for any lattice vector \mathbf{T} ,

$$\mathbf{T} = N_1\mathbf{a}_1 + N_2\mathbf{a}_2 + N_3\mathbf{a}_3 \quad (83)$$

N_1, N_2, N_3 being integer numbers, and $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ the vectors defining the periodically repeated simulation box.

In Eq. (81), the summation is over all reciprocal lattice vectors \mathbf{G} which fulfill the condition $\mathbf{G} \cdot \mathbf{T} = 2\pi M$, M being an integer number. In practice, this plane-wave expansion of the Kohn–Sham orbitals is truncated such that the individual terms all yield kinetic energies lower than a specified cutoff value, E_{cut} ,

$$\frac{\hbar^2}{2m} |\mathbf{k} + \mathbf{G}|^2 \leq E_{\text{cut}} \quad (84)$$

The plane-wave basis set thus has the advantage over other basis sets that convergence can be controlled by a single parameter, namely, E_{cut} . Which cutoff value is the best compromise between accuracy and computational cost depends on the specific system studied, on the properties of interest, and, crucially, on the choice of pseudopotentials (see below).

In this periodic setup, the electron density (see Eq. 67) can be approximated by a sum over a mesh of N_{kpt} \mathbf{k} -points in the Brillouin zone (Chadi and Cohen 1973; Monkhorst and Pack 1976; Moreno and Soler 1992),

$$\rho(\mathbf{r}) \approx \frac{1}{N_{\text{kpt}}} \sum_{\mathbf{k}} f_{\mathbf{k},j} |\psi_{\mathbf{k},j}(\mathbf{r})|^2 \quad (85)$$

Since the volume of the Brillouin zone, $V_{\text{BZ}} = (2\pi)^3/V_{\text{box}}$, decreases with increasing volume of the simulation supercell, V_{box} , only a small number of \mathbf{k} -points need to be sampled for large supercells. For insulating materials (i.e., large band gap), a single \mathbf{k} -point is often sufficient, typically taken to be $\mathbf{k} = 0$ (Γ -point approximation).

Pseudopotentials

While plane waves are a good representation of delocalized Kohn–Sham orbitals in metals, a huge number of them would be required in the expansion (81) to obtain a good approximation of atomic orbitals, in particular near the nucleus where they oscillate rapidly. Therefore, in order to reduce the size of the basis set, only the valence electrons are treated explicitly, while the core electrons (i.e., the inner shells) are taken into account implicitly through pseudopotentials combining their effect on the valence electrons with the nuclear Coulomb potential. This *frozen-core approximation* is justified as typically only the valence electrons participate in chemical interactions. To minimize the number of basis functions, the pseudopotentials are constructed in such a way as to produce nodeless atomic valence wave functions. Beyond a specified cutoff distance from the nucleus, R_{cut} the nodeless pseudo-wave functions are required to be identical to the reference all-electron wave functions.

Normconserving Pseudopotentials

are generated subject to the condition that the pseudo-wave function has the same norm as the all-electron wave function and thus gives rise to the same electron

density. Although normconserving pseudopotentials have to fulfill a (small) number of mathematical conditions, there remains considerable freedom in how to create them. Hence, several different recipes exist (Hamann et al. 1979; Bachelet et al. 1982; Kerker 1980; Vanderbilt 1985; Troullier and Martins 1990, 1991; Goedecker et al. 1996; Hartwigsen et al. 1998).

Since pseudopotentials are generated using *atomic* orbitals as a reference, it is not guaranteed that they are *transferable* to any chemical environment. Generally, transferability is the better the smaller the cutoff radius R_{cut} is chosen. However, the reduction in the number of plane waves required to represent a particular pseudo-wave function – i.e., the *softness* of the corresponding pseudopotential – increases as R_{cut} gets larger. So R_{cut} has to be chosen carefully and there is always a trade-off between transferability and softness. An upper limit for R_{cut} is given by the shortest interatomic distances in the molecule or crystal the pseudopotential will be used for: one needs to make sure that the sum of the two cutoff radii of any two neighboring atoms is smaller than their actual spatial separation.

For each angular momentum l a separate pseudopotential $V_l^{\text{PS}}(r)$ is constructed. The total pseudopotential operator is written as

$$\hat{V}^{\text{PS}} = V_{\text{loc}}^{\text{PS}}(r) + \sum_l V_{\text{nl},l}^{\text{PS}}(r) \hat{P}_l \quad (86)$$

where the nonlocal part is defined as

$$V_{\text{nl},l}^{\text{PS}}(r) = V_l^{\text{PS}}(r) - V_{\text{loc}}^{\text{PS}}(r) \quad (87)$$

and the local part $V_{\text{loc}}^{\text{PS}}(r)$ is taken to be the pseudopotential $V_l^{\text{PS}}(r)$ for one specific value of l , typically the highest one for which a pseudopotential was created. The pseudopotential (86) is called semi-local, since the projector \hat{P}_l only acts on the l -th angular momentum component of the wave function, but not on the radius r . (**Note:** a pseudopotential is called nonlocal if it is l -dependent.)

To achieve higher numerical efficiency, it is common practice to transform the semi-local pseudopotential (86) to a fully nonlocal form,

$$\hat{V}^{\text{PS}} = V_{\text{loc}}^{\text{PS}}(r) + \sum_{ij} |\beta_i \rangle B_{ij} \langle \beta_j| \quad (88)$$

using the Kleinman-Bylander prescription (Kleinman and Bylander 1982).

Vanderbilt Ultrasoft Pseudopotentials

An ultrasoft type of pseudopotential was introduced by Vanderbilt (1990; Laasonen et al. 1993) to deal with nodeless valence states which are strongly localized in the core region. In this scheme the normconserving condition is lifted and only a small portion of the electron density inside the cutoff radius is recovered by the pseudo-wave function, the remainder is added in the form of so-called augmentation

charges. Complications arising from this scheme are the nonorthogonality of Kohn–Sham orbitals, the density dependence of the nonlocal pseudopotential, and the need to evaluate additional terms in atomic force calculations.

How to Obtain Pseudopotentials?

There are extensive pseudopotential libraries available for download with the simulation packages CPMD (Hutter et al.), CP2K (Hutter et al.), QuantumEspresso (www.quantum-espresso.org), Abinit (www.abinit.org), or online (www.physics.rutgers.edu/~dhv/uspp/; <http://cp2k.web.psi.ch/potentials>; www.quantum-espresso.org; www.abinit.org). However, before applying any pseudopotentials, they should always be tested against all-electron calculations. Pseudopotentials used in conjunction with a particular density functional should have been generated using the same functional.

In many cases, the required pseudopotential will not be available in any accessible library; in this case it may be generated using freely downloadable programs (Vanderbilt www.physics.rutgers.edu/~dhv/uspp/; Atomic Pseudopotentials Engine <http://tddft.org/programs/APE>; Hamann www.mat-simresearch.com/oncepsp-2.1.1.tar.gz; Opium – pseudopotential generation project <http://opium.sourceforge.net>; Martins www.abinit.org/downloads/psp-links/psp-links/MARTINS2ABINIT.tgz; Fuchs and Scheffler http://cpc.cs.qub.ac.uk/summaries/ADKA_v1_0.html; Dacapo pseudopotential library http://wiki.fysik.dtu.dk/dacapo/Pseudopotential_Library).

How to Set the Plane-Wave Cutoff?

The number of plane-wave basis functions, i.e., the value of the cutoff energy E_{cut} , required for a sufficiently accurate DFT calculation sensitively depends on the choice of pseudopotential. For a calculation on H₂O with normconserving pseudopotentials of the Troullier–Martins type (Troullier and Martins 1990, 1991), for example, a cutoff energy of 70 Ry (= 35 a.u.) is typically used (see, e.g., Sprik et al. 1996), while 25 Ry is usually sufficient with ultrasoft pseudopotentials. Goedecker–Teter–Hutter normconserving pseudopotentials (Goedecker et al. 1996; Hartwigsen et al. 1998; Krack 2005) normally require a slightly higher cutoff than Troullier–Martins pseudopotentials (CPMD Manual www.cpmc.org:81/manual/node73.html).

Strictly speaking, the best cutoff energy should be determined for every new system and set of pseudopotentials by carrying out a series of single point calculations with different E_{cut} and analyzing the convergence of the total energy and the properties of interest.

Setting Up a Classical MD Simulation

There are two general stages that make up the preparation to conduct force-field molecular dynamics simulations: (1) gathering preliminary information and (2) building the actual system.

Gathering Preliminary Information

Gathering the preliminary information before conducting the simulation is mostly focused on making sure that the simulation is possible. First, it is important to identify the type and number of molecules that you wish to model. Then, it is necessary to find the force field that will allow you to most accurately model the molecules and physical system that you want to simulate. A brief synopsis of some of the larger classical force-field parameter sets is given in section "Classical Force Fields." These force fields and references may be good starting points in searching for the correct classical force field to use for a given system, but the best way to find a specific force field is to just conduct a search for research articles that may have been conducted on the same system. If no force-field parameters exist for the system of interest, then you can use configurations and energies from quantum simulations to parameterize a given force field for your system. A methodology for how a force field was parameterized originally is presented in the relevant paper; however, this is a complicated exercise and is probably best left to the experts.

Building the System

After identifying that a force field exists for the system you wish to model, the next step is to build the initial configuration of the molecules within the system. The initial configuration will consist of initial spatial coordinates of each atom in each given molecule. When building a large system consisting of several molecules of various types, it is easiest to write a computer code that contains or reads in the molecular structure and coordinates of each molecule present in the system and then have the code replicate each molecule however many times is necessary in order to build the entire system. Alternatively, most of the molecular dynamics simulation packages previously mentioned have capabilities to build systems from a pdb file; however, these tools are often useful for only certain systems and force fields. There is unfortunately no one tool which can be used to build any system with any force field.

These initial configurations can represent a minimum energy structure either from another simulation (i.e., a final structure from an energy minimization in a quantum or classical Monte Carlo simulation can be used as the starting state for classical simulations), from experimental observation (i.e., the pdb database for crystallographic structures of proteins) or building the initial coordinates based upon the equilibrium bond distances and bond angles from the force field.

The placement of the molecules within the simulation box can be done in a number of different ways as well. The molecules can be placed on the vertices of a regular lattice or in any other regularly defined geometry that may be useful for conducting your simulation (i.e., in simulating the structural properties of micelles oftentimes the surfactant molecules will initially be placed on the vertices of a bucky ball such that they are in a spherical configuration). Also, molecules can be placed at random positions within the simulation box. The one advantage of placing molecules at regularly spaced positions is that it is easier to insure that there is no overlapping of molecules, whereas with the randomly placed molecules it can be quite difficult to ensure that a placed molecule does not overlap with another

molecule in the box (particularly for large or highly branched molecules). For more complex geometries, there are several software packages and online resources which are very useful in generating initial configurations (most of these output xyz and/or pdb files and which are generally useable within the majority of molecular dynamics codes). These tools include Enhanced Monte Carlo (<http://montecarlo.sourceforge.net/emc/Welcome.html>), VMD TopoTools (<http://sites.google.com/site/akohlmeyer/software/topotools>), Avogadro (www.avogadro.cc), Packmol ([ch7-2:packmol-wwww](http://ch7-2.packmol-wwww.org)), and AtomsK (http://pierrehiel.info/codes_atomsk.php?lang=eng), which are all freely available. Then if you are interested in carrying out simulations with the CHARMM or MARTINI force fields, you can use the online resource – CHARMM-GUI (<http://www.charmm-gui.org/>) – to create a variety of systems by choosing the molecules and arrangement of molecules (i.e., bilayer, micelle, ...) and it will output a pdb file and result force field files for you.

In addition to containing the initial spatial coordinates of all of the molecules in the system, the initial configuration must also contain some additional information about the atoms and molecules in the systems. Each atom in the configuration must contain a label of what atomic species (i.e., carbon, nitrogen, etc.) it represents. This label will be different for each simulation code used, but all of them will have some type of label as it will inform the simulation code what force-field values to use to represent the interactions of that atom. A list of all of the covalent bonds, the bond angles, and the dihedrals in the system will also need to be included in the initial configuration. The lists of the bonds, angles, and dihedrals contain an identifier for each atom that make up the bond, angle, or dihedral and then an identifier for the type that informs the simulation package which parameters to use in calculating the energy of the bond, angle, or dihedral. The final component of the initial configuration of a classical simulation is a list of all of the various types of atoms, bonds, angles, and dihedrals in the system along with their corresponding force-field parameters (i.e., ϵ and σ for atom types to describe their nonbonded interactions, force constants, and equilibrium values for bond, angle, and dihedral types).

Finally, after building the initial configuration, the simulation is about ready to be performed. The last step is to choose the simulation variables and set up the input to the simulation package in order to convey these selections.

These options and the decision process behind choosing from the various options will be presented in the following sections.

Preparing an Input File

Optimization Algorithms

Optimization algorithms are often used to find stationary points on a potential energy surface, i.e., local and global minima and saddle points. It is often useful to carry out a geometry optimization before starting an MD simulation, in particular when the initial coordinates are likely to be far away from their realistic values. The only place where they directly enter MD is in the case of Born–Oppenheimer AIMD,

in order to converge the SCF wave function for each MD step. It is immediately obvious that the choice of optimization algorithm crucially affects the speed of the simulation.

Steepest Descent

The steepest descent method is the simplest optimization algorithm. The initial energy $E[\Psi_0] = E(\mathbf{c}_0)$, which depends on the plane-wave expansion coefficients \mathbf{c} (see Eq. 81), is lowered by altering \mathbf{c} in the direction of the negative gradient,

$$\mathbf{d}_n = -\frac{\partial E(\mathbf{c}_n)}{\partial \mathbf{c}} \equiv -\mathbf{g}_n \quad (89)$$

$$\mathbf{c}_{n+1} = \mathbf{c}_n + \Delta_n \mathbf{d}_n \quad (90)$$

where $\Delta_n > 0$ is a variable step size chosen such that the energy always decreases and n is the optimization step index. The steepest descent method is very robust; it is guaranteed to approach the minimum. However, the rate of convergence ever decreases as the energy gets closer to the minimum, making this algorithm rather slow.

Conjugate Gradient Methods

The Conjugate Gradient method generally converges faster than the steepest descent method due to the fact that it avoids moving in a previous search direction. This is achieved by linearly combining the gradient vector and the last search vector,

$$\mathbf{c}_{n+1} = \mathbf{c}_n + \Delta_n \mathbf{d}_n \quad (91)$$

where

$$\mathbf{d}_n = -\mathbf{g}_n + \beta_n \mathbf{d}_{n-1} \quad (92)$$

Different recipes exist to determine the coefficient β_n (Jensen 2007) among which the Polak–Ribière formula usually performs best for non-quadratic functions,

$$\beta_n = \frac{\mathbf{g}_n(\mathbf{g}_n - \mathbf{g}_{n-1})}{\mathbf{g}_{n-1}\mathbf{g}_{n-1}} \quad (93)$$

In the case of a general non-quadratic function, such as the DFT energy, conjugacy is not strictly fulfilled and the optimizer may search in completely inefficient directions after a few steps. It is then recommended to restart the optimizer (setting $\beta = 0$). Convergence can be improved by multiplying \mathbf{g}_n with a preconditioner matrix, e.g., an approximate inverse of the second derivatives matrix (Hessian in the case of geometry optimization) $\tilde{\mathbf{H}}$. The method is then called the preconditioned conjugate gradient (PCG). In the CPMD code, the matrix $\tilde{\mathbf{H}}$ is approximated by

$$\tilde{\mathbf{H}} = \begin{cases} H_{GG'}^{\text{KS}} & \text{for } G \geq G_{\text{cut}} \\ H_{G_{\text{cut}}G_{\text{cut}}}^{\text{KS}} & \text{for } G < G_{\text{cut}} \end{cases} \quad (94)$$

where $H_{GG'}^{\text{KS}}$ is the Kohn–Sham matrix in the plane-wave basis and G_{cut} is a cutoff value for the reciprocal lattice vector \mathbf{G} (set to a default value of 0.5 a.u.).

Direct Inversion of the Iterative Subspace

Having generated a sequence of optimization steps \mathbf{c}_i , the direct inversion of the iterative subspace (DIIS) method (Pulay 1980, 1982; Császár and Pulay 1984; Hutter et al. 1994) is designed to accelerate convergence by finding the best linear combination of stored \mathbf{c}_i vectors,

$$\mathbf{c}_{n+1} = \sum_{i=1}^n a_i \mathbf{c}_i \quad (95)$$

Ideally, of course, \mathbf{c}_{n+1} is equal to the optimum vector \mathbf{c}_{opt} . Defining the error vector \mathbf{e}_i for each iteration as

$$\mathbf{e}_i = \mathbf{c}_i - \mathbf{c}_{\text{opt}} \quad (96)$$

Equation (95) becomes

$$\sum_{i=1}^n a_i \mathbf{c}_{\text{opt}} + \sum_{i=1}^n a_i \mathbf{e}_i = \mathbf{c}_{\text{opt}} \quad (97)$$

Equation (97) is satisfied if

$$\sum_{i=1}^n a_i = 1 \quad (98)$$

and

$$\sum_{i=1}^n a_i \mathbf{e}_i = 0 \quad (99)$$

Instead of the ideal case Eq. (99), in practice one minimizes the quantity

$$\left\langle \sum_{i=1}^n a_i \mathbf{e}_i \mid \sum_{j=1}^n a_j \mathbf{e}_j \right\rangle \quad (100)$$

subject to the constraint (98), which is equivalent to solving the system of linear equations

$$\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} & -1 \\ b_{21} & b_{22} & \cdots & b_{2n} & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} & -1 \\ -1 & -1 & \cdots & -1 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -1 \end{pmatrix} \quad (101)$$

where

$$b_{ij} = \langle \mathbf{e}_i | \mathbf{e}_j \rangle \quad (102)$$

and the error vectors are approximated by

$$\mathbf{e}_i = -\tilde{\mathbf{H}}^{-1}(\mathbf{c}_{\text{opt}})\mathbf{g}_i \quad (103)$$

using an approximate Hessian matrix $\tilde{\mathbf{H}}$, e.g., Eq. (94).

Controlling Temperature: Thermostats

If understanding the behavior of the system as a function of temperature is the aim of your study, then it is important to be able to control the temperature of your system. The temperature of the system is related to the time average of the kinetic energy, which can be calculated by

$$\langle \frac{1}{2} \sum_I M_I V_I^2 \rangle_{NVT} = \frac{3N-6}{2} k_B T \quad (104)$$

if the system is in thermal equilibrium. The right-hand side of Eq. (104) can be approximated by $\frac{3N}{2} k_B T$ for a large number of particles.

Below we introduce specific thermostating techniques for MD simulations at thermodynamic equilibrium, e.g., for calculating equilibrium spatial distribution and time-correlation functions. However, when MD simulations are performed on a system undergoing some *nonequilibrium* process involving exchange of energy between different parts of the system, e.g., when an energetic particle, such as an atom or a molecule, hits a crystal surface, or there is a temperature gradient across the system, one has to resort to specially developed techniques, see, for example, Kantorovich (2008), Kantorovich and Rompotis (2008), and Toton et al. (2010). In these methods, based on the so-called generalized Langevin equation, the actual system on which MD simulations are performed is considered in contact with one (or more) heat bath(s) kept at constant temperature(s), and the dynamics of the system of interest reflects the fact that there is an interaction and energy transfer between the system and the surrounding heat bath(s).

Rescale Thermostat

One obvious way to control the temperature of a system is to rescale the velocities of the atoms within the system (Woodcock 1971). The rescaling factor λ is determined from $\lambda \sqrt{T_{\text{target}}/T_0}$, where T_{target} and T_0 are the target and initial temperatures, respectively. Then, the velocity of each atom is rescaled such that $V_f = \lambda V_i$. In practice, the inputs generally required to use a rescale thermostat include:

- T_0 – Initial temperature,
- T_{target} – Target temperature,
- τ - Damping constant, (i.e., frequency with which to apply the thermostat),
- δT – Maximum allowable temperature difference from T_{target} before thermostat is applied,
- f_{rescale} – Fraction of temperature difference between current temperature and T_{target} is corrected during each application of thermostat.

If it is desired to have a strict thermostat (i.e., when first starting a simulation that might have particles very near one another), then δT and τ should have values of $\sim 0.01 T_{\text{target}}$ and 1 time step, respectively, and f_{rescale} should be near 1.0. However, if you wish to allow a more lenient thermostat, then the value of δT should be of the same order of magnitude as T_{target} , τ should be $\sim 10^2 - 10^3$ time steps, and $f_{\text{rescale}} \sim 0.01 - 0.1$.

Berendsen Thermostat

Another way to control the temperature is to couple the system to an external heat bath, which is fixed at a desired temperature. This is referred to as a Berendsen thermostat (Berendsen et al. 1984). In this thermostat, the heat bath acts as a reservoir of thermal energy that supplies or removes temperature as necessary. The velocities are rescaled each time step, where the rate of change in temperature is proportional to the difference in the temperature in the system $T(t)$ and the temperature of the external bath T_{bath} :

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_{\text{bath}} - T(t)) \quad (105)$$

which when integrated results in the change in temperature each time step:

$$\Delta T = \frac{\delta t}{\tau}(T_{\text{bath}} - T(t)). \quad (106)$$

In Eqs. 105 and 106, τ is the damping constant for the thermostat. In practice, the necessary inputs when using the Berendsen thermostat include:

- T_{bath} – temperature of the external heat bath,
- τ – damping constant for the thermostat.

Obviously the amount of control that the thermostat imposes on the simulation is controlled by the value of τ . If τ is large, then the coupling will be weak and the

temperature will fluctuate significantly during the course of the simulation, while if τ is small, then the coupling will be strong and the thermal fluctuations will be small. If $\tau = \delta t$, then the result will be the same as the rescale thermostat, in general.

Nosé–Hoover Thermostat

While the Berendsen thermostat is efficient for achieving a target temperature within your system, the use of a Nosé–Hoover thermostat properly reproduces a canonical ensemble once the system has reached a thermal equilibrium. The extended system method, which was originally introduced by Nose (1984a, b) and then further developed by Hoover (1985), introduces additional degrees of freedom into the Hamiltonian that describes the system, from which equations of motion can be determined.

The extended system method considers the external heat bath as an integral part of the system by including an additional degree of freedom in the Hamiltonian of the system that is represented by the variable s . As a result, the potential energy of the reservoir is

$$E_{\text{pot}} = (f + 1)k_B T \ln s, \quad (107)$$

where f is the number of degrees of freedom in the physical system and T is the target temperature. The kinetic energy of the reservoir is calculated by

$$E_{\text{kin}} = \frac{Q}{2} \left(\frac{ds}{dt} \right)^2, \quad (108)$$

where Q is a parameter with dimensions of energy \times (time)² and is generally referred to as the "virtual" mass of the extra degree of freedom s . The magnitude of Q determines the coupling between the heat bath and the real system, thus influencing the temperature fluctuations.

Utilizing Eqs. 107 and 108, and substituting the real variables for the corresponding

Nosé variables, the equations of motion are found to be as follows:

$$\ddot{\mathbf{R}}_I = \frac{\mathbf{F}_I}{M_I} - \gamma \dot{\mathbf{R}}_I, \quad (109)$$

$$\dot{\gamma} = \frac{f k_B}{Q} \left(T - T_{\text{target}} \right), \quad (110)$$

where $\gamma = \frac{\dot{s}}{s}$. quantity

$$\tau_{\text{NH}} = \sqrt{\frac{Q}{f k_B T_{\text{target}}}} \quad (111)$$

is an effective relaxation time or damping constant.

In practice, the inputs that are necessary when utilizing the Nosé–Hoover thermostat during a molecular dynamics simulation include:

- T_{target} – Target temperature,
- Q – Fictitious mass of the additional degree of freedom s .

The most significant variable in the above list is Q . Large values of Q may cause poor temperature control, with the infinite limit resulting in no energy exchange between the temperature bath and the real system, which is the case of conventional molecular dynamics simulations resulting in the microcanonical ensemble. However, if Q is too small, then the energy oscillates and the system will take longer in order to reach a thermal equilibrium. An in-depth discussion of the Nosé–Hoover thermostat can be found in Hunenberger (2005) and Tuckerman (2010), for example.

Nosé–Hoover Chains

The Nosé–Hoover thermostat discussed above has been demonstrated to run into ergodicity problems and does not generate the canonical ensemble in some pathological cases. To make the particle motion more “chaotic,” in the Nosé–Hoover chain thermostat, the simple Nosé–Hoover thermostat is itself coupled to another thermostat, which again can be coupled to a thermostat, and so forth. The equations of motion for a chain of m thermostats are given by (Martyna et al. 2002; Tuckerman 2010)

$$\ddot{\mathbf{R}}_I = \frac{\mathbf{F}_I}{M_I} - \gamma_1 \dot{\mathbf{R}}_I \quad (112)$$

$$\dot{\gamma}_1 = \frac{fk_B}{Q_1} [T - T_{\text{target}}] - \gamma_1 \gamma_2 \quad (113)$$

⋮

$$\dot{\gamma}_j = \frac{1}{Q_j} [Q_{j-1} \gamma_{j-1}^2 - kT] - \gamma_j \gamma_{j+1} \quad (114)$$

⋮

$$\dot{\gamma}_m = \frac{1}{Q_m} [Q_{m-1} \gamma_{m-1}^2 - kT] \quad (115)$$

The conserved quasi-Hamiltonian is

$$\tilde{H} = H + \sum_{j=1}^m \frac{Q_j}{2} \gamma_j^2 + 3NkT \int \gamma_1 dt + \sum_{j=2}^m kT \int \gamma_j dt \quad (116)$$

Compared to the input parameters required for the simple Nosé–Hoover thermostat, the user now needs to also specify the length m of the chain (e.g., $m = 4$ is the default value in the CPMD program).

Langevin Thermostat

The Langevin thermostat is a stochastic thermostat, in which the equations of motion of the particles in the system take the form:

$$M_I \ddot{\mathbf{R}}_I = \mathbf{F}_I + \mathbf{F}_{\text{friction}}(\dot{\mathbf{R}}_I). \quad (117)$$

Therefore, the force that a particle experiences is due to the gradient in the potential energy surface and a frictional force due to collisions of the particle with the particles in the thermal bath. This frictional force is defined as

$$\mathbf{F}_{\text{friction}}(\dot{\mathbf{R}}_I) = -\gamma_I M_I \dot{\mathbf{R}}_I + \boldsymbol{\eta}_I(t) \quad (118)$$

where γ_I is a proportionality constant, which is often referred to as the "friction coefficient." The term $\boldsymbol{\eta}_I(t)$ is a random force due to the stochastic collisions with the particles in the thermal bath, and it has a Gaussian probability distribution with no correlation to any past or present \mathbf{R} or $\dot{\mathbf{R}}$ and an autocorrelation function

$$\langle \eta_I(t) \eta_J(t') \rangle = 2\gamma_I k_B T_{\text{target}} \delta_{I,J} \delta(t - t'), \quad (119)$$

where k_B is Boltzmann's constant and T_{target} is the temperature of the thermal bath.

In practice, Eq. (117) is solved using a difference equation over a small time step δt , such that:

$$\Delta \mathbf{V}_I = \mathbf{A}_I(t) \delta t - \gamma_I \mathbf{V}_I(t) \delta t + \zeta \sqrt{2k_B T_{\text{target}} \gamma_I \delta t / M_I} \quad (120)$$

where ζ is a random number with unit variance.

When using a Langevin thermostat during a molecular dynamics simulation, the necessary inputs generally are:

- T_{target} – Target (or bath) temperature,
- γ – The friction coefficient.

The most significant variable in the above list is γ . Choosing too large a value of γ will result in the system becoming overdamped due to the friction term, and choosing too small a value of γ may result in too weak of a coupling between the thermal bath and the system and therefore poor temperature control.

Controlling Pressure: Barostats

It may be desired to study the behavior of the simulated system while the pressure is held constant (i.e., pressure-induced phase transitions). Many experimental measurements are made in conditions where the pressure and temperature are held constant and so it is of utmost importance to be able to accurately replicate these conditions in simulations.

One thing of note is that the pressure often fluctuates more than other quantities such as the temperature in an NVT molecular dynamics simulation or the energy in an NVE molecular dynamics simulation. This is due to the fact that the pressure is related to the virial term, which is the product of the positions of the particles in the system and the derivative of the potential energy function. These fluctuations will be observed in the instantaneous values of the system pressure during the course of the simulation, but the average pressure should approach the desired pressure. Since generally the temperature and number of atoms will also be held constant during constant pressure simulations, and the volume of the system will be allowed to change in order to arrive at the desired pressure. Therefore, less compressible systems will show larger fluctuations in the pressure than the systems that are more easily compressed.

Berendsen Barostat

Many of the approaches used for the controlling the pressure are similar to those that are used for controlling the temperature. One approach is to maintain constant pressure by coupling the system to a constant pressure reservoir as is done in the Berendsen barostat (Berendsen et al. 1984), which is analogous to the way temperature is controlled in the Berendsen thermostat. The pressure change in the system is determined by:

$$\frac{dP(t)}{dt} = \frac{1}{\tau_P}(P_0 - P(t)), \quad (121)$$

where τ_P is time constant of the barostat, P_0 is the desired pressure and $P(t)$ is the system pressure at any time t . In order to accommodate this change in pressure, the volume of the box is scaled by a factor of μ^3 each time step; therefore, the coordinates of each particle in the system are scaled by a factor of μ (i.e., $\mathbf{R}_I(t + \delta t) = \mu^{1/3}\mathbf{R}_I(t)$, where

$$\mu = \left[1 - \frac{\delta t}{\tau_P}(P - P_0) \right]^{\frac{1}{3}}. \quad (122)$$

In practice, the inputs for the Berendsen barostat will include:

- P_0 – Desired pressure,
- τ_P – Time constant of the barostat,

One other input that may be included in the use of the Berendsen barostat is to define which dimensions are coupled during the pressure relaxation. For example, you could define that the pressure is relaxed in a way that the changes in all three dimensions are coupled and therefore all of the dimensions change at the same rate. On the other hand, the pressure relaxation can be handled in an anisotropic manner, such that none of the dimensions are coupled and each dimension will have its own scaling factor that results from the individual pressure components.

Nosé–Hoover Barostat

Similar to the Nosé–Hoover thermostat, the extended system method has been applied to create a barostat (Hoover 1986) that is coupled with a Nosé–Hoover thermostat. In this case, the extra degree of freedom η corresponds to a “piston,” and it is added to the Hamiltonian of the system, which results in the following equations of motion:

$$\frac{d\mathbf{R}(t)}{dt} = \mathbf{V}(t) + \eta(t)(\mathbf{R}(t) - \mathbf{R}_{\text{COM}}), \quad (123)$$

$$\frac{d\mathbf{V}(t)}{dt} = \frac{\mathbf{F}(t)}{M} - [\chi(t) + \eta(t)]\mathbf{V}(t), \quad (124)$$

$$\frac{d\chi(t)}{dt} = \frac{1}{\tau_T^2} \left(\frac{T}{T_0} - 1 \right), \quad (125)$$

$$\frac{d\eta(t)}{dt} = \frac{1}{Nk_B T_0 \tau_P^2} V(t)(P - P_0), \quad (126)$$

$$\frac{dV(t)}{dt} = 3\eta(t)V(t) \quad (127)$$

where \mathbf{R}_{COM} are the coordinates of the center of mass of the system, η is the thermostat extra degree of freedom and can be thought of as a friction coefficient, τ_T is the thermostat time constant, χ is barostat extra degree of freedom and is considered a volume scaling factor, and τ_P is the barostat time constant. Equations 126 and 127 explicitly contain the volume of the simulation box, $V(t)$. Generally, this barostat is implemented using the approach described in Melchionna et al. (1993).

In addition to the variables that are a part of the equations of motion, there is a variable Q that represents the “mass” of the “piston.” This is analogous to the “mass” variable in the Nosé–Hoover thermostat. In practice, the required input for the Nosé–Hoover barostat will include:

- P_0 – Desired pressure,
- T_0 – Desired temperature,
- τ_P – Time constant of the barostat,
- τ_T – Time constant of the thermostat,
- Q – The “mass” of the piston.

Like in the case of the Nosé–Hoover thermostat, care must be taken when selecting the value of the variable Q . A small value of Q is representative of a piston with small mass and thus will have rapid oscillations of the box size and pressure, whereas a large value of Q will have the opposite effect. The infinite limit of Q results in normal molecular dynamics behavior.

Setting the Timestep

Born–Oppenheimer MD/Classical MD

Since BOMD is a classical MD in the sense that the nuclei are classical particles, the same rules concerning the choice of timestep apply to both BOMD and atomistic or coarse-grained force-field MD. The largest possible timestep, δt , is partially determined by the fastest oscillation in the system – in many molecules this would be a bond-stretching vibration involving hydrogen, e.g., CH, NH, or OH. It is immediately plausible that δt must be smaller than the shortest vibrational period in order to resolve that motion and for the numerical integrator (see section “Classical Molecular Dynamics”) to be stable. Let us assume a particular molecule has an OH vibration at $\omega' = 3500 \text{ cm}^{-1}$. Then corresponding to a period T can be calculated to be about 10 fs:

$$\omega' = \frac{\hbar\omega}{hc} = \frac{h\nu}{hc} = \frac{\nu}{c} = \frac{1}{cT} \quad (128)$$

where we have used $\nu = 1/T$ and c is the speed of light (remember to use c in cm/s if ω' is in cm^{-1}). Solving (128) for T , we get

$$T = \frac{1}{c\omega'} \quad (129)$$

According to the Nyquist–Shannon sampling theorem, the lowest allowed sampling frequency is twice the maximum oscillation frequency of the simulated system, which would mean for the present example that the maximum timestep is 5 fs.

Another important consideration is the stability limit of the numerical integrator used in the MD simulation (Sutmann 2006). For a one-dimensional harmonic oscillator, the propagation of nuclear position and momentum can be written in matrix form as

$$\begin{pmatrix} \omega R_{n+1} \\ V_{n+1} \end{pmatrix} = \mathbf{M} \begin{pmatrix} \omega R_n \\ V_n \end{pmatrix} \quad (130)$$

In order for the propagation to be stable, the eigenvalues λ of matrix \mathbf{M} , which are obtained from the secular equation,

$$\det(\mathbf{M} - \lambda\mathbf{1}) = 0 \quad (131)$$

must satisfy the condition $|\lambda| < 1$. This then defines the maximum timestep.

The velocity Verlet algorithm is usually implemented in the form

$$V_{n+\frac{1}{2}} = V_n + \frac{\delta t}{2M} F_n \quad (132)$$

$$R_{n+1} = R_n + \delta t V_{n+\frac{1}{2}} \quad (133)$$

$$V_{n+1} = V_{n+\frac{1}{2}} + \frac{\delta t}{2M} F_{n+1} \quad (134)$$

Using the force expression for harmonic oscillator,

$$F = -\omega^2 MR \quad (135)$$

we obtain

$$V_{n+\frac{1}{2}} = V_n - \frac{\omega^2 \delta t}{2} R_n \quad (136)$$

$$R_{n+1} = R_n + \delta t V_{n+\frac{1}{2}} \quad (137)$$

$$V_{n+1} = V_{n+\frac{1}{2}} - \frac{\omega^2 \delta t}{2} R_{n+1} \quad (138)$$

The total propagation matrix \mathbf{M} (see Eq. 130) can be written as the product of the three propagation matrices corresponding to steps (136–138),

$$\mathbf{M} = \begin{pmatrix} 1 & 0 \\ -\frac{\alpha}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{\alpha}{2} & 1 \end{pmatrix} = \begin{pmatrix} 1 - \frac{\alpha^2}{2} & \alpha \\ -\alpha(1 - \frac{\alpha^2}{4}) & 1 - \frac{\alpha^2}{2} \end{pmatrix} \quad (139)$$

where $\alpha = \omega \delta t$. The eigenvalues of \mathbf{M} are then given by

$$\lambda = 1 - \frac{\alpha^2}{2} \pm \alpha \sqrt{1 - \frac{\alpha^2}{4}} \quad (140)$$

The stability condition is satisfied if (Sutmann 2006)

$$\omega^2 \delta t^2 < 2 \quad (141)$$

Solving (141) for δt , we obtain

$$\delta t < \frac{\sqrt{2}}{\omega} = \frac{T}{\pi \sqrt{2}} \approx \frac{T}{4.44} \quad (142)$$

In our example of the OH vibration at 3500 cm^{-1} , this would dictate a maximum timestep of 2 fs. However, although such a choice guarantees numerical stability, it

results in deviations from the exact answer (Sutmann 2006). Therefore, in practice smaller timesteps – typically around 1 fs – are often used.

Car–Parrinello MD

Although in CPMD the nuclei are still treated as classical particles, the choice of timestep can no longer be based solely on the highest nuclear frequency ω_n^{\max} . We also need to consider the fictitious dynamics of the electronic degrees of freedom. In fact, the optimum simulation timestep is closely linked to the value of the fictitious electron mass μ as we will see in the following.

The fictitious mass μ has to be chosen small enough to guarantee adiabatic separation of electronic and nuclear motion. This means that the frequency spectrum of the electronic degrees of freedom (Marx and Hutter 2009; Pastore et al. 1991)

$$\omega_{ph} = \sqrt{\frac{2(\epsilon_p - \epsilon_h)}{\mu}} \quad (143)$$

must not overlap with the vibrational spectrum of the nuclear system. The above formula (143) can be derived assuming that the fictitious orbital masses behave like classical harmonic oscillators whose energy, ΔE , is equal to the energy difference between an unoccupied and an occupied orbital,

$$\Delta E = \epsilon_p - \epsilon_h = \frac{1}{2} \mu \omega^2 \quad (144)$$

It is important to realize that μ has the dimension kg m^2 , as this is required to give the fictitious kinetic energy term in the Car–Parrinello Lagrangian (16), $\sum_i \mu_i |\dot{\psi}_i|^2$, the dimension $\text{kg m}^2/\text{s}^2$ (i.e., energy). Hence, ΔE in (144) also has the correct dimension of $\text{kg m}^2/\text{s}^2$ (i.e., energy). Equation (143) can now be easily obtained by solving (144) for ω .

The lowest electronic frequency according to Eq. (143) is

$$\omega^{\min} = \sqrt{\frac{2(\epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}})}{\mu}} \quad (145)$$

The highest electronic frequency is determined by the plane-wave cutoff energy E_{cut} ,

$$\omega^{\max} \approx \sqrt{\frac{2E_{\text{cut}}}{\mu}} \quad (146)$$

Thus, the maximum oscillation period, which is inversely proportional to ω^{\max} , obeys the relation

$$T^{\max} = \frac{2\pi}{\omega^{\max}} \approx \pi \sqrt{\frac{2\mu}{E_{\text{cut}}}} \quad (147)$$

According to Eq. (147) the maximum oscillation period can be increased by simply increasing μ . However, this would also result in a lowering of ω_e^{\min} (see Eq. 145) and therefore in a smaller separation $\omega_e^{\min} - \omega_n^{\max}$ between the nuclear and electronic spectra.

Let us discuss the above using some realistic numbers. In the case of the H₂O molecule, for example, the HOMO–LUMO gap with the BLYP functional is about 5.7 eV. Assuming a typical value of 400 a.u. for μ , the minimum electronic frequency (Eq. 145) is ca. 6900 cm⁻¹. The highest energy molecular vibrational mode in a CPMD simulation using these parameter values is the asymmetric stretch at about 3500 cm⁻¹. This means that electronic and nuclear spectra are well separated. A basis set cutoff of $E_{\text{cut}} = 70$ Ry (= 35 a.u.), which is typical for simulations with normconserving pseudopotentials (see section "Pseudopotentials"), leads to a maximum electronic frequency (Eq. 146) of approximately 92,000 cm⁻¹ corresponding to a vibrational period of 15 a.u. According to (142), the maximum CPMD timestep is 3.4 a.u. In fact, for water, a timestep/fictitious mass combination of 4 a.u./400 a.u. has been shown to be a good compromise between efficiency and accuracy (Kuo et al. 2004). When ultrasoft pseudopotentials are employed, a smaller cutoff energy of 25 Ry is usually sufficient (see section "Pseudopotentials"). This increases the maximum electronic oscillation period to 25 a.u. (Eq. (147)) and the maximum simulation timestep to 5.6 a.u. (142).

If we were to increase μ to 1000 a.u., we could afford a larger timestep of about 6 a.u. (according to Eq. (147)). However, ω_e^{\min} (Eq. (145)) would become ca. 4500 cm⁻¹, dangerously close to ω_n^{\max} . A simple trick that is often used to be able to afford larger timesteps is to replace all hydrogen atoms by deuterium atoms thus downshifting ω_n^{\max} to about 2500 cm⁻¹. If the aim of the simulation is to learn about a particular reaction mechanism, changing the isotope mass does not affect the outcome. Dynamical properties, such as diffusion or vibrations, on the other hand, are mass dependent.

For systems with a small or even vanishing (e.g., metals) band gap, it is increasingly difficult or impossible to achieve adiabatic separation of electronic and nuclear degrees of freedom following the above considerations. A solution to this problem is the use of separate thermostats for the two subsystems (Marx and Hutter 2009; Sprik 1991).

Besides the numerical stability and adiabaticity issues discussed so far, one also has to bear in mind that the approximations made in the Car–Parrinello method do introduce deviations from a Born–Oppenheimer simulation that increase with the value of the fictitious mass μ . The physical explanation for this is that in CPMD, nuclear motion is dampened since they have to carry with them the heavy fictitious mass of the orbitals. This results, for instance, in a red-shifted vibrational spectrum. It has been suggested that this error can be compensated by a priori reduction of the nuclear masses used in a CPMD simulation (Pastore et al. 1991; Blochl 1994; Tangney and Scandolo 2002; Marx and Hutter 2009). In the rigid ion approximation, the difference in mass is (Tangney and Scandolo 2002)

$$\Delta M_I = \frac{2m_e}{3\hbar^2} \sum_{\eta} \mu_{\eta} T^{I,\eta} \quad (148)$$

where $T^{I,\eta}$ is the quantum-mechanical kinetic energy of an electron with the fictitious mass μ_η in orbital η of atom I . Vibrational frequencies obtained from a CPMD simulation can be corrected using this mass shift (Tangney and Scandolo 2002). For a system containing only a single element (e.g., a silicon crystal), a corrected frequency $\omega_{\text{corrected}}$ may be obtained from the CPMD frequency ω_{CP} using the relation (Tangney and Scandolo 2002)

$$\omega_{\text{corrected}} = \omega_{\text{CP}} \sqrt{1 + \frac{\Delta M}{M}} \quad (149)$$

Postprocessing

Data Analysis

Spatial Distribution Functions

For a system of N particles in a volume V at temperature T , the probability of molecule 1 being in the volume element $d\mathbf{R}_1$ around the position \mathbf{R}_1 , molecule 2 being in $d\mathbf{R}_2$, ..., molecule N being in $d\mathbf{R}_N$ is given by (McQuarrie 1992)

$$P^{(N)}(\mathbf{R})d\mathbf{R} = P^{(N)}(\mathbf{R}_1, \dots, \mathbf{R}_N)d\mathbf{R}_1, \dots, d\mathbf{R}_N = \frac{e^{-E(\mathbf{R})/kT}}{Z_N} \quad (150)$$

with the configuration integral

$$Z_N = \int_V e^{-E(\mathbf{R})/kT} d\mathbf{R} \quad (151)$$

where $E(\mathbf{R})$ is the potential energy of the system at configuration \mathbf{R} (cf. Eqs. 8 and 10).

For a subset of n molecules, the probability of molecule 1 being in $d\mathbf{R}_1$, ..., molecule n being in $d\mathbf{R}_n$ is

$$P^{(n)}(\mathbf{R}_1, \dots, \mathbf{R}_n) = \frac{\int \dots \int e^{-E(\mathbf{R})/kT} d\mathbf{R}_{n+1} \dots d\mathbf{R}_N}{Z_N} \quad (152)$$

The probability of *any* molecule being in $d\mathbf{R}_1$, ..., *any* molecule n being in $d\mathbf{R}_n$ is

$$\rho^{(n)}(\mathbf{R}_1, \dots, \mathbf{R}_n) = \frac{N!}{(N-n)!} P^{(n)}(\mathbf{R}_1, \dots, \mathbf{R}_n) \quad (153)$$

In a liquid the probability of finding any one molecule in $d\mathbf{R}_1$, $\rho^{(1)}(\mathbf{R}_1)d\mathbf{R}_1$, is independent of \mathbf{R}_1 . Therefore,

$$\frac{1}{V} \int \rho^{(1)}(\mathbf{R}_1)d\mathbf{R}_1 = \rho^{(1)} = \frac{N}{V} = \rho \quad (154)$$

The dependence of the molecules of a liquid on all the other molecules, in other words, their correlation, is captured by the *correlation function* $g^{(n)}(\mathbf{R}_1, \dots, \mathbf{R}_n)$, which is defined by

$$\rho^{(n)}(\mathbf{R}_1, \dots, \mathbf{R}_n) = \rho^n g^{(n)}(\mathbf{R}_1, \dots, \mathbf{R}_n) \quad (155)$$

Using Eq. (153) we can thus write

$$g^{(n)}(\mathbf{R}_1, \dots, \mathbf{R}_n) = \frac{V^n N!}{N^n (N-n)!} \frac{\int \dots \int e^{-E(\mathbf{R})/kT} d\mathbf{R}_{n+1} \dots d\mathbf{R}_N}{Z_N} \quad (156)$$

The two-body correlation function $g^{(2)}(\mathbf{R}_1, \mathbf{R}_2)$ is of particular interest as it can be determined in X-ray diffraction experiments. In the following we shall only consider the dependence of $g^{(2)}$ on the interparticle distance $R = R_{12} = |\mathbf{R}_1 - \mathbf{R}_2|$, i.e., we have averaged over any angular dependence, and call $g^{(2)}(R_{12}) = g(R)$ the *radial distribution function*. The quantity $\rho g(R) d\mathbf{R}_I$ is proportional to the probability of finding another particle, I , in $d\mathbf{R}_I$ if the reference particle is at the origin. Spherical integration yields

$$\int \rho g(R) 4\pi R^2 dR = N - 1 \approx N \quad (157)$$

showing that $\rho g(R) 4\pi R^2 dR$ is the number of particles in the spherical volume element between R and $R + dR$ about the central particle. The radial distribution function $g(R)$ is proportional to the local density $\rho(R) = \rho g(R)$ about a certain molecule. In a fluid, $g(R) \rightarrow 1$ as $R \rightarrow \infty$, i.e., there is no long-range order and we "see" only the average particle density. At very short range, i.e., $R \rightarrow 0$, $g(R) \rightarrow 0$, due to the repulsiveness of the molecules. Examples from a CPMD simulation of liquid water are shown in Fig. 3.

The radial distribution function $g(R)$ provides a useful measure of the quality of a simulation as it can be compared to experimental – X-ray or neutron diffraction – data obtained by Fourier transform of the structure factor

$$h(k) = \rho \int [g(R) - 1] e^{i\mathbf{k}\mathbf{R}} d\mathbf{R} \quad (158)$$

where \mathbf{k} is the wave vector.

In addition to characterizing the structure of a liquid, the radial distribution function may also be used to calculate thermodynamic properties such as the total energy,

$$E = \frac{3}{2} NkT + 2\pi N\rho \int_0^\infty u(R) g(R) R^2 dR \quad (159)$$

the pressure,

$$p = \rho kT - \frac{2}{3} \pi \rho^2 \int_0^\infty \frac{du(R)}{dR} g(R) R^3 dR \quad (160)$$

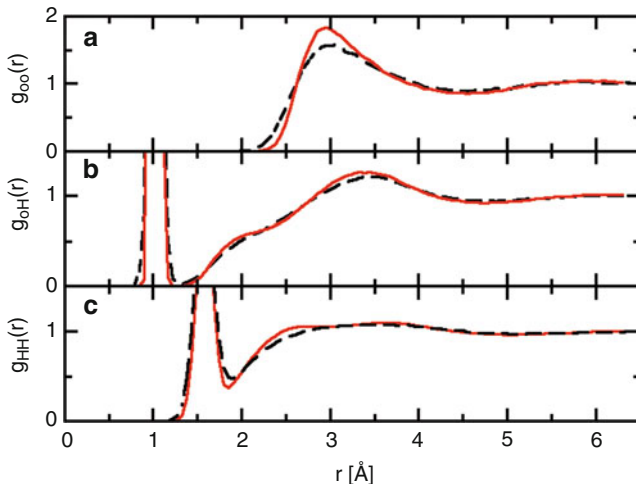


Fig. 3 Radial distribution function of liquid water from CPMD simulations at 900 and 1200 K, respectively

and the chemical potential,

$$\mu = kT \ln(\rho\Lambda^3) + 4\pi\rho \int_0^1 d\xi \int_0^\infty u(R)g(R, \xi)R^2 dR \quad (161)$$

where

$$\Lambda = \sqrt{\frac{h^2}{2\pi mkT}} \quad (162)$$

is the thermal de Broglie wavelength. By varying the coupling parameter ξ between 0 and 1 one can effectively take a molecule in and out of the system. It should be stressed that Eqs. (159)–(161) have been derived assuming a pairwise additive intermolecular potential $u(R)$.

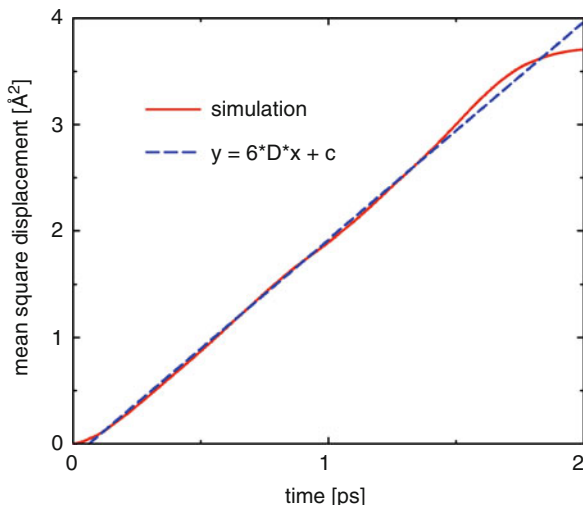
We now define the *potential of mean force*, i.e., the interaction between n fixed molecules averaged over the configurations of the remaining molecules $n + 1, \dots, N$, as

$$w^{(n)}(\mathbf{R}_1, \dots, \mathbf{R}_n) = -kT \ln g^{(n)}(\mathbf{R}_1, \dots, \mathbf{R}_n) \quad (163)$$

The mean force acting on molecule J is then obtained from

$$f_J^{(n)} = -\nabla_J w^{(n)} \quad (164)$$

Fig. 4 Mean square displacement of liquid water from CPMD simulations at 900 K and linear fit to determine the diffusion constant D using Eq. (168)



Time Correlation Functions

The classical *time autocorrelation function* of some vectorial function

$$\mathbf{A}(t) = \mathbf{A}(P(t), Q(t)) = \mathbf{A}(P, Q; t) \quad (165)$$

where $Q(t)$ and $P(t)$ are the generalized coordinate and momentum, respectively, is defined as

$$C(t) = \langle \mathbf{A}(0)\mathbf{A}(t) \rangle = \int \dots \int dP dQ \mathbf{A}(P, Q; 0)\mathbf{A}(P, Q; t) f(P, Q) \quad (166)$$

where $f(P, Q)$ is the equilibrium phase space distribution function.

From the velocity autocorrelation function, for example, one can calculate the diffusion coefficient as

$$D = \frac{1}{3} \int_0^{\infty} \langle \mathbf{V}_I(0)\mathbf{V}_I(t) \rangle dt \quad (167)$$

where \mathbf{V}_I is the velocity of particle I . Alternatively, one can obtain the diffusion coefficient for long times from the associated *Einstein relation*,

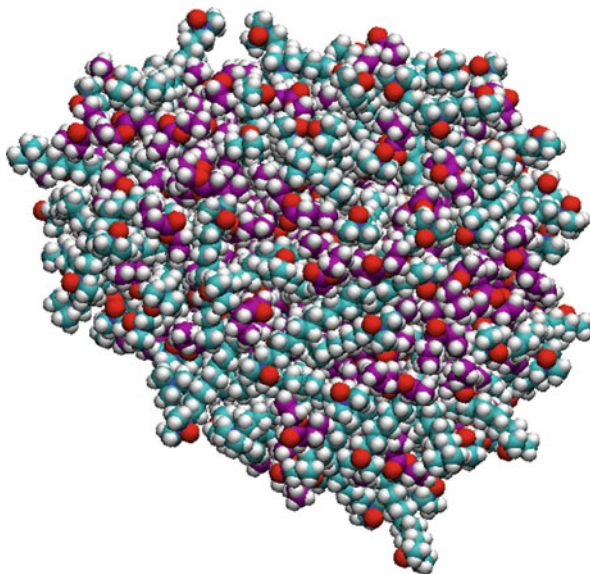
$$6tD = \langle |\mathbf{R}_I(t) - \mathbf{R}_I(0)|^2 \rangle \quad (168)$$

In practice, D is then determined from a linear fit to the *mean square displacement* (rhs of Eq. 168) as one sixth of the slope. An example is shown in Fig. 4.

Another common application of correlation functions is the calculation of IR absorption spectra. The line-shape function, $I(\omega)$, is given by the Fourier transform of the autocorrelation function of the electric dipole moment \mathbf{M} ,

$$I(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \langle \mathbf{M}(0)\mathbf{M}(t) \rangle e^{-i\omega t} dt \quad (169)$$

Fig. 5 A snapshot of a micelle formed from DDAO molecules and oil molecules formed using the VMD software package (<http://www.ks.uiuc.edu/Research/vmd/>)



Visualization

Due to the nature of MD simulations, one of the most productive forms of analysis of a simulation is to be able to visualize the trajectory of the molecules of interest. This is particularly useful since experimental techniques are not able to produce visual pictures of atomistic interactions and therefore it is something that only simulations (at this point) are able to provide. In order to visualize a simulation trajectory, there are several different very powerful computer packages that are commonly used. These software packages include VMD (<http://www.ks.uiuc.edu/Research/vmd/>), PyMol (<http://pymol.sourceforge.net/>), RasMol (<http://rasmol.org/>), and several others (<http://www.umass.edu/microbio/rasmol/otherof.htm>). Figure 5 shows an example of the type of pictures that can be made using the visualization software.

Each of these codes will generally accept the trajectory in any number of standard inputs (i.e., pdb, xyz,...) and then will generate snapshots which can be rendered individually or as a movie. In addition to providing the visualization, these codes have become progressively powerful analysis codes in their own right. They now have the ability to measure bond lengths, angles, and dihedrals as a function of time and determine the solvent accessible surface area, hydrogen bond network, and many other useful structure-related and dynamic properties of the system.

Bibliography

ABINIT. www.abinit.org.

Allen, M. P., & Tildesley, D. J. (1987). *Computer simulation of liquids*. Oxford: Clarendon Press.

Amara, P., Field, M. J., Alhambra, C., & Gao, J. (2000). *Theoretical Chemistry Accounts*, 104, 336.

- Anderson, J. A., Lorenz, C. D., & Travesset, A. (2008). *Journal of Computational Physics*, 227, 5342.
- Aqvist, J., & Warshel, A. (1993). *Chemical Reviews*, 93, 2523.
- Artacho, E., Anglada, E., Dieguez, O., Gale, J. D., Garcia, A., Junquera, J., Martin, R. M., Ordejón, P., Pruneda, J. M., Sánchez-Portal, D., & Soler, J. M. (2008). *Journal of Physics: Condensed Matter*, 20, 064208.
- Ashcroft, N. W., & Mermin, N. D. (1976). *Solid state physics*. Philadelphia: Saunders College Publishing.
- Assfeld, X., & Rivail, J. L. (1996). *Chemical Physics Letters*, 263, 100.
- Assfeld, X., Ferré, N., & Rivail, J. L. (1998). In J. Gao & M. A. Thompson (Eds.), *Combined quantum mechanical and molecular mechanical methods* (ACS Symposium Series, Vol. 712, p. 234). Washington, DC: American Chemical Society.
- Atomsk website. http://pierrehirel.info/codes_atomsk.php?lang=eng.
- Atomic Pseudopotentials Engine. <http://tddft.org/programs/APE>.
- Aulbur, W. G., Jonsson, L., & Wilkins, J. W. (2000). *Solid State Physics*, 54, 1.
- Avogadro. www.avogadro.cc.
- Bachelet, G. B., Hamann, D. R., & Schlüter, M. (1982). *Physical Review B*, 26, 4199.
- Balint-Kurti G. G. (2008). *International Reviews in Physical Chemistry*, 27, 507.
- Baskes, M. I. (1992). *Phys. Rev. B*, 46, 2727.
- Becke, A. D. (1988). *Physical Review A*, 38, 3098.
- Becke, A. D. (1993). *Journal of Chemical Physics*, 98, 5648.
- Becke, A. D., & Johnson, E. R. (2005). *Journal of Chemical Physics*, 122, 154101.
- van Beest, B. W. H., Kramer, G. J., & van Santen, R. A. (1990). *Physical Review Letters*, 64, 1955.
- Bereau, T., & Deserno, M. (2009). *Journal of Chemical Physics*, 130, 235106.
- Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., Nola, A. D., & Haak, J. R. (1984). *Journal of Chemical Physics*, 81, 3684.
- Betancourt, M. R., & Omovie, S. J. (2009). *Journal of Chemical Physics*, 130, 195103.
- Bloch, P. E. (1994). *Physical Review B*, 50, 17953.
- Bloch, P. E., Forst, C. J., & Schimpl, J. (2003). *Bulletin of Materials Science*, 26, 33.
- Bockstedte, M., Kley, A., Neugebauer, J., & Scheffler, M. (1997). *Computer Physics Communications*, 107, 187.
- Boeck, S. (2009). *Development and application of the S/PHI/nX library*. Saarbrücken: Südwestdeutscher Verlag für Hochschulschriften.
- Bowler, D. R., Choudhury, R., Gillan, M. J., & Miyazaki, T. (2006). *Physica Status Solidi B*, 243, 989.
- Bredow, T., & Jug, K. (2005). *Theoretical Chemistry Accounts*, 113, 1.
- Brini, E., Algaer, E. A., Ganguly, P., Li, C., Rodríguez-Roperio, F., & van der Vegt, N. F. A. (2013). *Soft Matter*, 9, 2108.
- Brooks, B. R., Brooks III, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., & Karplus, M. (2009). *Journal of Computational Chemistry*, 30, 1545.
- Buckingham, R. A. (1938). *Proceedings of the Royal Society of London A*, 168, 264.
- Burke, K. *The ABC of DFT*. <http://dft.uci.edu/doc/g1.pdf>.
- Bulo, R. E., Ensing, B., Sikkema, J., & Visscher, L. (2009). *Journal of Chemical Theory and Computation*, 5, 2212.
- Boys, S. F., & Bernardi, F. (1970). *Molecular Physics*, 19, 553.
- Caldwell, J., & Kollman, P. A. (1995). *Journal of Physical Chemistry*, 99, 6208.
- Caldwell, J., & Kollman, P. A. (1995). *Journal of the American Chemical Society*, 117, 4177.
- Caldwell, J., Dang, L. X., & Kollman, P. A. (1990). *Journal of the American Chemical Society*, 112, 9144.
- Cambridge Crystallographic Data Centre. www.ccdc.cam.ac.uk.

- Case, D. A., Cheatham III, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., & Woods, R. (2005). *Journal of Computational Chemistry*, 26, 1668.
- Car, R., & Parrinello, M. (1985). *Physical Review Letters*, 55, 2471.
- CASTEP. www.tcm.phy.cam.ac.uk/castep/.
- CGENFF website. <http://mackerell.umaryland.edu/~kenno/cgenff>.
- CHARMM-GUI website. <http://www.charmm-gui.org/>.
- Chadi, D. J., & Cohen, M. L. (1973). *Physical Review B*, 8, 5747.
- ChemShell. www.chemshell.org/.
- Chowdhary, J., Harder, E., Lopes, P. E. M., Huang, L., MacKerell, J. A. D., & Roux, B. (2013). *Journal of Physical Chemistry B*, 117, 9142.
- Clark, S. J., Segall, M. D., Pickard, C. J., Hasnip, P. J., Probert, M. J., Refson, K., & Payne, M. C. (2005). *Zeitschrift für Kristallographie*, 220, 567.
- CONQUEST. www.conquest.ucl.ac.uk/.
- Cooke, I. R., Kremer, K., & Deserno, M. (2005). *Physical Review E*, 72, 011506.
- Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., & Kollman, P. A. (1995). *Journal of the American Chemical Society*, 117, 5179.
- CP - PAW. <https://orion.pt.tu-clausthal.de/paw/>.
- CPMD Manual. www.cpmc.org:81/manual/node73.html.
- Császár, P., & Pulay, P. (1984). *Journal of Molecular Structure*, 114, 31.
- DACAPO. <http://dcwww.camp.dtu.dk/campos/Dacapo/>.
- Dacapo pseudopotential library. http://wiki.fysik.dtu.dk/dacapo/Pseudopotential_Library.
- Dang, L. X., Rice, J. E., Caldwell, J., & Kollman, P. A. (1991). *Journal of the American Chemical Society*, 113, 2481.
- Daw, M. S., & Baskes, M. I. (1983). *Physical Review Letters*, 50, 1285.
- Daw, M. S., & Baskes, M. I. (1984). *Physical Review Letters*, 29, 6443.
- Day, P. N., Jensen, J. H., Gordon, M. S., Webb, S. P., Stevens, W. J., Krauss, M., Garmer, D., Basch, H., & Cohen, D. (1996). *Journal of Chemical Physics*, 105, 1968.
- Dirac, P. A. M. (1930). *Proceedings of the Cambridge Philosophical Society*, 26, 37.
- Doltsinis, N. L. (2006). Molecular dynamics beyond the born-oppenheimer approximation: Mixed quantum-classical approaches. In J. Grotendorst, S. Blügel, & D. Marx (Eds.), *Computational nanoscience: Do it yourself!* FZ Jülich: NIC. www.fz-juelich.de/nic-series/volume31/doltsinis1.pdf.
- Doltsinis, N. L., & Marx, D. (2002). *Physical Review Letters*, 88, 166402.
- Doltsinis, N. L., & Marx, D. (2002). *Journal of Chemical Theory and Computation*, 1, 319–349.
- Dick, B. G., & Overhauser, A. W. (1958). *Physical Review*, 112, 90.
- Dreizler, R. M., & Gross, E. K. U. (1990). *Density-functional theory*. Berlin: Springer.
- Elezgaray, J., & Laguerre, M. (2006). *Computer Physics Communications*, 175, 264.
- Elsner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., Suhai, S., & Seifert, G. (1998). *Physical Review B*, 58, 7260.
- Enhanced Monte Carlo website. <http://montecarlo.sourceforge.net/emc/Welcome.html>.
- Ercolessi, F., & Adams, J. B. (1994). *Europhysics Letters*, 26, 583.
- Ewald, P. P. (1921). *Annals of Physics*, 369, 253.
- Ferré, N., Assfeld, X., & Rivail, J. L. (2002). *Journal of Computational Chemistry*, 23, 610.
- FHI98md. www.fhi-berlin.mpg.de/th/fhimd/.
- Finnis, M. W., & Sinclair, J. E. (1984). *Philosophical Magazine A*, 29, 6443.
- Fiolhais, C., Nogueira, F., & Marques, M. (Eds.). (2003). *A primer in density functional theory*. Berlin: Springer.
- Fornili, A., Sironi, M., & Raimondi, M. (2003). *THEOCHEM*, 632, 157.
- Fornili, A., Moreau, Y., Sironi, M., & Assfeld, X. (2006a). *Journal of Computational Chemistry*, 27, 515.

- Fornili, A., Loos, P.-F., Sironi, M., & Assfeld, X. (2006b). *Chemical Physics Letters*, 427, 236.
- Fuchs, M., & Scheffler, M. FHI98PP ab initio pseudopotentials. http://cpc.cs.qub.ac.uk/summaries/ADKA_v1_0.html.
- Fumi, F. G., & Tosi, M. P. (1964). *Journal of Physics and Chemistry of Solids*, 25, 31.
- Gao, J., Amara, P., Alhambra, C., & Field, M. J. (1998). *Journal of Physical Chemistry A*, 102, 4714.
- Garcia-Viloca, M., & Gao, J. (2004). *Theoretical Chemistry Accounts*, 111, 280.
- Generalized AMBER Force Field website. <http://ambermd.org/antechamber/gaff.html>.
- Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Chiarotti, G. L., Cococcioni, M., Dabo, I., Corso, A. D., Fabris, S., Fratesi, G., de Gironcoli, S., Gebauer, R., Gerstmann, U., Gougoussis, C., Kokalj, A., Lazzeri, M., Martin-Samos, L., Marzari, N., Mauri, F., Mazzarello, R., Paolini, S., Pasquarello, A., Paulatto, L., Sbraccia, C., Scandolo, S., Sclauzero, G., Seitsonen, A. P., Smogunov, A., Umari, P., & Wentzcovitch, R. M. (2009). *Journal of Physics: Condensed Matter*, 21, 395502.
- Goedecker, S., Teter, M., & Hutter, J. (1996). *Physical Review B*, 54, 1703.
- Goetz, R., Compper, G., & Lipowsky, R. (1999). *Physical Review Letters*, 82, 221.
- Gordon, M. S., Freitag, M. A., Bandyopadhyay, P., Jensen, J. H., Kairys, V., & Stevens, W. J. (2001). *Journal of Physical Chemistry A*, 105, 105.
- Grigorenko, B. L., Nemukhin, A. V., Topol, I. A., & Burt, S. K. (2002). *Journal of Physical Chemistry A*, 106, 10663.
- Grimme's DFT-D3 Site. www.thch.uni-bonn.de/tc.
- Grimme, S., Anthony, J., Ehrlich, S., & Krieg, H. (2010). *Journal of Chemical Physics*, 132, 154104.
- Groenhof, G. (2013). In L. Monticelli & E. Salonen (Eds.), *Biomolecular simulations: Methods and protocols*. New York: Springer.
- Hamann, D. *ONCVSPSP very soft norm-conserving generator*. www.mat-simresearch.com/oncvpsp-2.1.1.tar.gz.
- Hamann, D. R., Schlüter, M., & Chiang, C. (1979). *Physical Review Letters*, 43, 1494.
- Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). *Journal of Chemical Information*, 4, 17.
- Hartwigsen, C., Goedecker, S., & Hutter, J. (1998). *Physical Review B*, 58, 3641.
- He, X., Lopes, P. E. M., & MacKerell, J. A. D. (2013). *Biopolymers*, 99, 724.
- Heyden, A., Lin, H., & Truhlar, D. G. (2007). *Journal of Physical Chemistry B*, 111, 2231.
- Hockney, R. W. (1970). *Methods in Computational Physics*, 9, 136.
- Hockney, R., & Eastwood, J. (1981). *Computer simulation using particles*. New York: McGraw-Hill.
- Hofer, T. S., Pribil, A. B., Randolf, B. R., & Rode, B. M. (2005). *Journal of the American Chemical Society*, 127, 14231.
- Hoover, W. G. (1985). *Physical Review A*, 31, 1695.
- Hoover, W. G. (1986). *Physical Review A*, 34, 2499.
- Hünenberger, P. (2005). Thermostat algorithms for molecular dynamics simulations. In C. Holm & K. Kremer (Eds.), *Advanced computer simulation approaches for soft matter sciences I* (Advances in polymer science, Vol. 173, pp. 105–147). Berlin: Springer.
- Hutter, J., et al. (2016). A general program to perform molecular dynamics simulations. See <http://cp2k.berlios.de>.
- Hutter, J., et al. (2016). Car–Parrinello molecular dynamics: An *ab initio* electronic structure and molecular dynamics program. See www.cpmd.org.
- Hutter, J., Lüthi, H. P., & Parrinello, M. (1994). *Computation Materials Science*, 2, 244.
- Inorganic Crystal Structure Database. www.fiz-karlsruhe.de/icsd.html.
- Izvekov, S., & Voth, G. A. (2005). *Journal of Physical Chemistry B*, 109, 2469.
- Izvekov, S., & Voth, G. A. (2006). *Journal of Chemical Theory and Computation*, 2, 637.
- Jensen, F. (2007). *Introduction to computational chemistry*. Chichester: Wiley.

- Jensen, J. H., Day, P. N., Gordon, M. S., Basch, H., Cohen, D., Garmer, D. R., Krauss, M., & Stevens, W. J. (1994). In D. A. Smith (Ed.), *Modeling the hydrogen bond* (ACS Symposium Series, Vol. 569, p. 139). Washington, DC: American Chemical Society.
- Johnson, E. R., & Becke, A. D. (2005). *Journal of Chemical Physics*, *123*, 024101.
- Johnson, E. R., & Becke, A. D. (2006). *Journal of Chemical Physics*, *124*, 174104.
- de Jong, D. H., Singh, G., Bennett, W. F. D., Arnarez, C., Wassenaar, T. A., Schaffer, L. V., Periole, X., Tieleman, D. P., & Marrink, S. J. (2013). *Journal of Chemical Theory and Computation*, *9*, 687.
- Jorgensen, J. D. M. W. L., & Swenson, C. J. (1984). *Journal of the American Chemical Society*, *106*, 6638.
- Jorgensen, W. L., Jensen, K. P., & Alexandrova, A. N. (2007). *Journal of Chemical Theory and Computation*, *3*, 1987.
- Jung, J., Choi, C. H., Sugita, Y., & Ten-no, S. (2007). *Journal of Chemical Physics*, *127*, 204102.
- Kairys, V., & Jensen, J. H. (2000). *Journal of Physical Chemistry A*, *104*, 6656.
- Kaminski, G. A., Stern, H. A., Berne, B. J., Friesner, R. A., Cao, Y. X., Murphy, R. B., Zhou, R., & Halgren, T. A. (2002). *Journal of Computational Chemistry*, *23*, 1515.
- Kantorovich, L. (2008). *Physical Review B*, *78*, 094304.
- Kantorovich, L., & Rompotis, N. (2008). *Physical Review B*, *78*, 094305.
- Kar, P., Gopal, S. M., Cheng, Y. -M., Predeus, A., & Feig, M. (2013). *Journal of Chemical Theory and Computation*, *9*, 3769.
- Kendall, R. A., Apra, E., Bernholdt, D. E., Bylaska, E. J., Dupuis, M., Fann, G. I., Harrison, R. J., Ju, J., Nichols, J. A., Nieplocha, J., Straatsma, T. P., Windus, T. L., & Wong, A. T. (2000). *Computer Physics Communications*, *128*, 260.
- Kerdcharoen, T., & Morokuma, K. (2002). *Chemical Physics Letters*, *355*, 257.
- Kerdcharoen, T., Liedl, K. R., & Rode, B. M. (1996). *Chemical Physics*, *211*, 313.
- Kerker, G. (1980). *Journal of Physics C*, *13*, L189.
- Khalid, S., Bond, P. J., Holyoake, J., Hawtin, R. W., & Sansom, M. S. P. (2008). *Journal of the Royal Society Interface*, *5*, S241.
- Kleinman, L., & Bylander, D. M. (1982). *Physical Review Letters*, *48*, 1425.
- Koch, W., & Holthausen, M. C. (2000). *A chemist's guide to density functional theory*. New York: Wiley.
- Kolos, W. (1970). *Advances in Quantum Chemistry*, *5*, 99.
- Krack, M., (2016). *Goedecker-Teter-Hutter pseudopotentials*. <http://cp2k.web.psi.ch/potentials>.
- Krack, M. (2005). *Theoretical Chemistry Accounts*, *114*, 145.
- Kremer, K., & Grest, G. (1990). *Journal of Chemical Physics*, *92*, 5057.
- Kresse, G., & Furthmüller, J. (1996). *Physical Review B*, *54*, 11169.
- Kunz, A.-P. E., & van Gunsteren, W. F. (2009). *Journal of Physical Chemistry A*, *113*, 11570.
- Kuo, I.-F. W., Mundy, C. J., McGrath, M. J., Siepmann, J. I., VandeVondele, J., Sprik, M., Hutter, J., Chen, B., Klein, M. L., Mohamed, F., Krack, M., & Parrinello, M. (2004). *Journal of Physical Chemistry B*, *108*, 12990.
- Kutzelnigg, W. (1997). *Molecular Physics*, *90*, 909.
- Laasonen, K., Pasquarello, A., Car, R., Lee, C., & Vanderbilt, D. (1993). *Physical Review B*, *47*, 10142.
- Laio, A., VandeVondele, J., & Rothlisberger, U. (2002). *Journal of Chemical Physics*, *116*, 6941.
- Lamoureux, G., & Roux, B. (2003). *Journal of Chemical Physics*, *119*, 3025.
- Lamoureux, G., MacKerell, J. A. D., & Roux, B. (2003). *Journal of Chemical Physics*, *119*, 5185.
- Lee, C., Yang, W., & Parr, R. C. (1988). *Physical Review B*, *37*, 785.
- Lindahl, B. H. E., & van der Spoel, D. (2001). *Journal of Molecular Modeling*, *7*, 306.
- Liu, Y. -P., Kim, K., Berne, B. J., Friesner, R. A., & Rick, S. W. (1998). *Journal of Chemical Physics*, *108*, 4739.
- Loos, P. -F., & Assfeld, X. (2007). *Journal of Chemical Theory and Computation*, *3*, 1047.
- Lopes, P. E. M., Huang, J., Shim, J., Luo, Y., Li, H., Roux, B., & MacKerell, J. A. D. (2013). *Journal of Chemical Theory and Computation*, *9*, 5430.

- Lopez, C. A., Rzepiela, A., de Vries, A. H., Dijkhuizen, L., Huenenberger, P. H., and Marrink, S. J. (2009). *Journal of Chemical Theory and Computation*, 5, 3195.
- Lu, L., Dama, J. F., & Voth, G. A. (2013). *Journal of Chemical Physics*, 139, 121906.
- Lucas, T. R., Bauer, B. A., & Patel, S. (2012). *Biochimica et Biophysica Acta, Biomembranes*, 1818, 318.
- Luty, B. A., Davis, M. E., Tironi, I. G., & van Gunsteren, W. F. (1994). *Molecular Simulation*, 14, 11.
- Luty, B. A., Tironi, I. G., & van Gunsteren, W. F. (1995). *Journal of Chemical Physics*, 103, 3014.
- Lyubartsev, A. P. (2005). *European Biophysics Journal*, 35, 53.
- van Maaren, P. J., & van der Spoel, D. (2001). *Journal of Physical Chemistry B*, 105, 2618.
- Martins, J. L. (2005). Troullier-Martins pseudopotential generator. www.abinit.org/downloads/psp-links/psp-links/MARTINS2ABINIT.tgz.
- Marrink, S. J., de Vries, A. H., & Mark, A. E. (2004). *Journal of Physical Chemistry B*, 108, 750.
- Marrink, S. J., Risselada, H. J., Yefimov, S., Tieleman, D. P., & de Vries, A. H. (2007). *Journal of Physical Chemistry B*, 111, 7812.
- MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher III, W. E., Raux, B., Schlenkrich, M., Smith, J. C., Store, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., & Karplus, M. (1998). *Journal of Physical Chemistry B*, 102, 3586.
- Martini FF development website. <http://md.chem.rug.nl/cgmartini/index.php/parametrizing-new-molecule>.
- Martyna, G. J., Klein, M. L., & Tuckerman, M. E. (2002). *Journal of Chemical Physics*, 97, 2635–2643.
- Marx, D., & Hutter, J. (2000). Ab initio molecular dynamics: Theory and implementation. In J. Grotendorst (Ed.), *Modern methods and algorithms of quantum chemistry*. Jülich: NIC. www.theochem.rub.de/go/cprev.html.
- Marx, D., & Hutter, J. (2009). *Ab initio molecular dynamics: Basic theory and advanced methods*. Cambridge: Cambridge University Press.
- McQuarrie, D. A. (1992). *Statistical mechanics*. London: Academic.
- Melchionna, S., Ciccotti, G., & Holian, B. L. (1993). *Molecular Physics*, 78, 533.
- Mervis, J. (2001). *Science*, 293, 1235.
- Meyer, B. (2006). The pseudopotential plane wave approach. In J. Grotendorst, S. Blügel, & D. Marx (Eds.), *Computational nanoscience: Do it yourself!* FZ Jülich: NIC. www.fz-juelich.de/nic-series/volume31/meyer1.pdf.
- Molden – a pre- and post processing program of molecular and electronic structure. www.cmbi.ru.nl/molden/molden.html.
- Monard, G., Loos, M., Théry, V., Baka, K., & Rivail, J. L. (1996). *International Journal of Quantum Chemistry*, 58, 153.
- Monkhorst, H. J., & Pack, J. D. (1976). *Physical Review B*, 13, 5188.
- Moreno, J., & Soler, J. M. (1992). *Physical Review B*, 45, 13891.
- Müller-Plathe, F. (2002). *ChemPhysChem*, 3, 754.
- Murphy, R. B., Philipp, D. M., & Friesner, R. A. (2000). *Chemical Physics Letters*, 321, 113.
- Nemukhin, A. V., Grigorenko, B. L., Bochenkova, A. V., Topol, I. A., & Burt, S. K. (2002). *THEOCHEM*, 581, 167.
- Nemukhin, A. V., Grigorenko, B. L., Topol, I. A., & Burt, S. K. (2003). *Journal of Computational Chemistry*, 24, 1410.
- Nielsen, S. O., Moore, P. B., & Ensing, B. (2010). Adaptive multiscale molecular dynamics of macromolecular fluids. *Physical Review Letters*, 105, 237802.
- Noid, W. G. (2013). *Journal of Chemical Physics*, 139, 090901.
- Nose, S. (1984). *Journal of Chemical Physics*, 81, 511.
- Nose, S. (1984). *Molecular Physics*, 52, 255.
- NWChem. www.emsl.pnl.gov/docs/nwchem.

- ONETEP. www.tcm.phy.cam.ac.uk/onetep/.
- Opium – pseudopotential generation project. <http://opium.sourceforge.net>.
- Orsi, M., & Essex, J. W. (2011). *PLoS One*, 6, e28637.
- de Pablo, J. J. (2011). *Annual Review of Physical Chemistry*, 62, 555.
- Packmol website. <http://ime.unicamp.br/~martinez/packmol/>.
- ParamChem website. <http://cgenff.paramchem.org>.
- Patel, S., & Brooks, C. L. (2004). *Journal of Computational Chemistry*, 25, 1.
- Patel, S., MacKerell, J. A. D., & Brooks, C. L. (2004). *Journal of Computational Chemistry*, 25, 1504.
- Patel, D. S., He, X., & MacKerell, J. A. D. (2015). *Journal of Physical Chemistry B*, 119, 637.
- Parr, R. G., & Yang, W. (1989). *Density functional theory of atoms and molecules*. Oxford: Oxford University Press.
- Pastore, G., Smargiassi, E., & Buda, F. (1991). *Physical Review A*, 44, 6334.
- Pauling, L., & Yost, D. M. (1932). *Proceedings of the National Academy of Sciences of the United States of America*, 18, 414.
- Perdew, J. P., Chevary, J. A., Vosko, S. H., Jackson, K. A., Pederson, M. R., Singh, D. J., & Fiolhais, C. (1992). *Physical Review B*, 46, 6671.
- Pezeshki, S., & Lin, H. (2011). *Journal of Chemical Theory and Computation*, 7, 3625–3634.
- Pezeshki, S., Davis, C., Heyden, A., & Lin, H. (2014). *Journal of Chemical Theory and Computation*, 10, 4765–4776.
- Phillip, D. M., & Friesner, R. A. (1999). *Journal of Computational Chemistry*, 20, 1468.
- Phillips, J. C., Braun, R., Wang, W., Cumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., & Schulten, K. (2005). *Journal of Computational Chemistry*, 26, 1781.
- PINY. <http://homepages.nyu.edu/~mt33/PINY.MD/PINY.html>.
- Polák, R. (1986). *Chemical Physics*, 103, 277.
- Ponder, J. W., Wu, C., Ren, P., Pande, V. S., Chodera, J. D., Schnieders, M. J., Haque, I., Mobley, D. L., Lambrecht, D. S., DiStasio, R. A., Jr., Head-Gordon, M., Clark, G. N. I., Johnson, M. E., & Head-Gordon, T. (2010). *Journal of Physical Chemistry B*, 114, 2549.
- Plimpton, S. J. (1995). *Journal of Computational Physics*, 117, 1.
- Pu, J., Gao, J., & Truhlar, D. G. (2004). *Journal of Physical Chemistry A*, 108, 5454.
- Pu, J., Gao, J., & Truhlar, D. G. (2004). *Journal of Physical Chemistry A*, 108, 632.
- Pu, J., Gao, J., & Truhlar, D. G. (2005). *ChemPhysChem*, 6, 1853.
- Pulay, P. (1969). *Molecular Physics*, 17, 197.
- Pulay, P. (1980). *Chemical Physics Letters*, 73, 393.
- Pulay, P. (1982). *Journal of Computational Chemistry*, 3, 556.
- Pulay, P. (1987). *Advances in Chemical Physics*, 69, 241.
- PWscf. www.pwscf.org/.
- QuantumEspresso. www.quantum-espresso.org.
- Rappoport, D., Crawford, N. R. M., Furche, F., & Burke, K. (2008). *Which functional should I choose?* <http://dft.uci.edu/pubs/RCFB08.pdf>.
- RCSB Protein Data Bank. www.rcsb.org/pdb.
- Reciprocal Net – a distributed crystallography network for researchers, students and the general public. www.reciprocalnet.org/edumodules/commonmolecules.
- Reed, D. A. (2003). *Computer*, 36, 62.
- Refson, K. (2000). *Computer Physics Communications*, 126, 310.
- Reith, D., Pütz, M., & Müller-Plathe, F. (2003). *Journal of Computational Chemistry*, 24, 1624.
- Ren, P., & Ponder, J. W. (2002). *Journal of Computational Chemistry*, 23, 1497.
- Rick, S. W., & Berne, B. J. (1996). *Journal of the American Chemical Society*, 118, 672.
- Rick, S. W., Stuart, S. J., & Berne, B. J. (1994). *Journal of Chemical Physics*, 101, 6141.
- Riniker, S., Allison, J. R., & van Gunsteren, W. F. (2012). *Physical Chemistry Chemical Physics*, 14, 12423.
- Saunders, M. G., & Voth, G. A. (2013). *Annual Review of Biophysics*, 42, 73.
- Saveliev, A., & MacKerell, J. A. D. (2014). *Journal of Computational Chemistry*, 35, 1219.
- Schommers, W. (1973). *Physics Letters A*, 43, 197.

- Scott, W. R. P., Hünenberger, P. H., Tironi, I. G., Mark, A. E., Billeter S. R., Fennel, J., Torda, A. E., Huber, T., Krüger, P., & van Gunsteren, W. F. (1999). *Journal of Physical Chemistry A*, 103, 3596.
- Segall, M. D., Lindan, P. L. D., Probert, M. J., Pickard, C. J., Hasnip, P. J., Clark, S. J., & Payne, M. C. (2002). *Journal of Physics: Condensed Matter*, 14, 2717.
- Senn, H. M., & Thiel, W. (2009). *Angewandte Chemie International Edition*, 48, 1198.
- Shelley, J. C., Shelley, M. Y., Reeder, R. C., Bandyopadhyay, S., Moore, P. B., & Klein, M. L. (2001). *Journal of Physical Chemistry B*, 105, 9785.
- Shi, Y., Xia, Z., Zhang, J., Best, R., Wu, C., Ponder, J. W., & Ren, P. (2013). *Journal of Chemical Theory and Computation*, 9, 4046.
- Shillcock, J. C., & Lipowsky, R. (2002). *Journal of Chemical Physics*, 117, 5048.
- Shirts, M., & Pande, V. S. (2000). *Science*, 290, 1903.
- Shurki, A., & Warshel, A. (2003). In V. Dagett (Ed.), *Protein simulations* (Advances in protein chemistry, Vol. 66, p. 249). San Diego: Academic.
- SIESTA. www.uam.es/siesta/.
- Sironi, M., Genoni, A., Civera, M., Pieraccini, S., & Ghitti, M. (2007). *Theoretical Chemistry Accounts*, 117, 685.
- Skylaris, C. -K., Haynes, P. D., Mostofi, A. A., & Payne, M. C. (2005). *Journal of Chemical Physics*, 122, 084119.
- Smith, W., Yong, C. W., & Rodger, P. M. (2002). *Molecular Simulation*, 28, 385.
- Soler, J. M., Artacho, E., Gale, J. D., Garcia, A., Junquera, J., Ordejón, P., & Sánchez-Portal, D. (2002). *Journal of Physics: Condensed Matter*, 14, 2745.
- Soper, A. K. (1996). *Chemical Physics*, 202, 295.
- S/PHI/nX. www.sphinxlib.de.
- Spoel, D. V. D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen H. J. C. (2005). *Journal of Computational Chemistry*, 26, 1701.
- Språk, M. (1991). *Journal of Physical Chemistry*, 95, 2283.
- Språk, M., Hutter, J., & Parrinello, M. (1996). *Journal of Chemical Physics*, 105, 1142.
- Stern, H. A., Rittner, F., Berne, B. J., & Friesner, R. A. (2001). *Journal of Chemical Physics*, 115, 2237.
- Stillinger, F., & Weber, T. A. (1985). *Physical Review B*, 31, 5262.
- Sun, H., Ren, P., & Fried, J. R. (1998). *Computational and Theoretical Polymer Science*, 8, 229.
- Sutmann, G. (2002). Classical molecular dynamics. In J. Grotendorst, D. Marx, & A. Muramatsu (Eds.), *Quantum simulations of complex many-body systems: From theory to algorithms*. FZ Jülich: NIC. For downloads see <http://www.fz-juelich.de/nic-series/volume10/sutmann.pdf>.
- Sutmann, G. (2006). In J. Grotendorst, S. Blügel, & D. Marx (Eds.), *Computational nanoscience: Do it yourself!* FZ Jülich: NIC. www.fz-juelich.de/nic-series/volume31/sutmann.pdf.
- Swope, W. C., Anderson, H. C., Berens, P. H., & Wilson, K. R. (1982). *Journal of Chemical Physics*, 76, 637.
- Tangney, P., & Scandolo, S. (2002). *Journal of Chemical Physics*, 116, 14.
- Tepper, H. L., & Voth, G. A. (2005). *Journal of Chemical Physics*, 122, 124906.
- Tersoff, J. (1988). *Physical Review B*, 37, 6991.
- Tersoff, J. (1989). *Physical Review B*, 39, 5566.
- Thiel, W. (2009). QM/MM methodology: Fundamentals, scope, and limitations. In J. Grotendorst, N. Attig, S. Blügel, & D. Marx (Eds.), *Multiscale simulation methods in molecular sciences*. FZ Jülich: NIC. www.fz-juelich.de/nic-series/volume42/thiel.pdf.
- Théry, V., Rinaldi, D., Rivail, J. L., Maigret, B., & Ferenczy, G. G. (1994). *Journal of Computational Chemistry*, 15, 269.
- Tosi, M. P., & Fumi, F. G. (1964). *Journal of Physics and Chemistry of Solids*, 25, 45.
- Toth Information Systems. www.tothcanada.com.
- Toton, D., Lorenz, C. D., Rompotis, N., Martsinovich, N., & Kantorovich, L. (2010). *Journal of Physics: Condensed Matter*, 22, 074205.
- Tozzini, V. (2005). *Current Opinion in Structural Biology*, 15, 144.
- Troullier, N., & Martins, J. L. (1990). *Solid State Communications*, 74, 613.

- Troullier, N., & Martins, J. L. (1991). *Physical Review B*, 43, 1993.
- Tuckerman, M. E. (2002). In J. Grotendorst, D. Marx, & A. Muramatsu (Eds.), *Quantum simulations of complex many-body systems: From theory to algorithms*. FZ Jülich: NIC. for downloads see <http://www.fz-juelich.de/nic-series/volume10/tuckerman1.pdf>.
- Tuckerman, M. E. (2010). *Statistical mechanics: Theory and molecular simulation*. New York: Oxford University Press.
- Tuckerman, M. E., & Hughes, A. (1998). In B. J. Berne, G. Ciccotti, & D. F. Coker (Eds.), *Classical and quantum dynamics in condensed phase simulations*. Singapore: World Scientific.
- Tuckerman, M. E., Berne, B. J., Martyna, G. J., & Klein, M. L. (1993). Efficient molecular-dynamics and hybrid Monte-Carlo algorithms for path-integrals. *Journal of Chemical Physics*, 99, 2796.
- Vanderbilt, D. (1985). *Physical Review B*, 32, 8412.
- VandeVondele, J., Krack, M., Mohamed, F., Parrinello, M., Chassaing, T., & Hutter, J. (2005). *Computer Physics Communications*, 167, 103.
- VandeVondele, J., Iannuzzi, M., & Hutter, J. (2006). Large scale condensed matter calculations using the Gaussian and augmented plane waves method. In M. Ferrario, G. Ciccotti, & K. Binder (Eds.), *Computer simulations in condensed matter systems: From materials to chemical biology volume 1* (Lecture notes in physics, Vol. 703, p. 287). Springer, Berlin/Heidelberg.
- Vanommeslaeghe, K., Hatcher, E., Acharya, C., Kundu, S., Zhong, S., Shim, J., Darian, E., Guvench, O., Lopes, P., Vorobyov, I., & MacKerell, J. A. D. (2010). *Journal of Computational Chemistry*, 31, 671.
- Vanderbilt, D. (1990). *Physical Review B*, 41, 7892.
- Vanderbilt Ultra-Soft Pseudopotential Site. www.physics.rutgers.edu/~dhv/usp/.
- VASP. <http://cms.mpi.univie.ac.at/vasp/>.
- Verlet, L. (1967). *Physical Review*, 159, 98.
- VMD TopoTools website. <http://sites.google.com/site/akohlmeyer/software/topotools>.
- Vosko, S. H., Wilk, L., & Nusair, M. (1980). *Canadian Journal of Physics*, 58, 1200.
- Votca website. <http://www.votca.org/home>.
- Waller, M. P., Kumbhar, S., & Yang, J. (2014). A density-based adaptive quantum mechanical/molecular mechanical method. *ChemPhysChem*, 15, 3218.
- Wang, J. M., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004a). *Journal of Computational Chemistry*, 24, 1157.
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004b). *Journal of Computational Chemistry*, 25, 1157.
- Warshel, A. (1991). *Computer modeling of chemical reactions in enzymes and solutions*. New York: Wiley.
- Warshel, A. (2003). *Annual Review of Biophysics and Biomolecular Structure*, 32, 425.
- Warshel, A., & Levitt, M. (1976). *Journal of Molecular Biology*, 103, 227.
- Woodcock, L. V. (1971). *Chemical Physics Letters*, 10, 257.
- Worth, G. A., Meyer, H. D., Koeppel, H., Cederbaum, L. S., & Burghardt, I. (2008). *International Reviews in Physical Chemistry*, 27, 569.
- Xie, W., Pu, J., MacKerell, J. A. D., & Gao, J. (2007). *Journal of Chemical Theory and Computation*, 3, 1878.
- Yu, W., He, X., Vanommeslaeghe, K., & MacKerell, J. A. D. (2012). *Journal of Computational Chemistry*, 33, 2451.
- Zhang, Y. (2005). *Journal of Chemical Physics*, 122, 024114.
- Zhang, Y. (2006). *Theoretical Chemistry Accounts*, 116, 43.
- Zhang, Y., Lee, T. -S., & Yang, W. (1999). *Journal of Chemical Physics*, 110, 46.
- Zhong, Y., Bauer, B. A., Patel, S. (2011). *Journal of Computational Chemistry*, 32, 3339.

Sanjay Kumar

Contents

Introduction	398
Models	398
Continuum Models	399
Lattice Models	403
Polymer and Critical Phenomena	404
Methods and Techniques	405
Generating Function Technique	405
Exact Enumeration Technique	407
Monte Carlo Simulation	408
Molecular Dynamics	409
Applications	414
Conclusions	417
Bibliography	417

Abstract

Single molecule force spectroscopy constitutes a robust method for probing the unfolding of biomolecules. Knowledge gained from statistical mechanics is helping to build our understanding about more complex structure and function of biopolymers. Here, we have review some of the models and techniques that have been employed to study force-induced transitions in biopolymers. We briefly describe the merit and limitation of these models and techniques. In this context, we discuss statistical models of polymer along with numerical techniques, which may provide enhanced insight in understanding the unfolding of biomolecules.

S. Kumar (✉)

Department of Physics, Banaras Hindu University, Varanasi, India

e-mail: yashankit@yahoo.com

Introduction

Recent technological developments of experimental techniques, for example, optical tweezers, atomic force microscope, etc., made it possible to apply a force of the order of pN to manipulate the single biomolecules (Kumar and Li 2010; Rief et al. 1999, 1997; Tskhovrebova et al. 1997). Many interesting results, for example, structural, functional, and elastic properties of biomolecules, information about the kinetics of biomolecular reactions, and detection of molecular intermediates have been obtained (Bustamante et al. 2000; Cecconi et al. 2005; Smith et al. 1992, 1996). Moreover, these experiments also provided a platform to verify theoretical predictions based on the models developed in the framework of statistical mechanics (Bhattacharjee 2000; Giri and Kumar 2006; Kumar 2009; Kumar and Giri 2007; Kumar et al. 2005; Lubensky and Nelson 2000; Marenduzzo et al. 2001a; Zhou et al. 2006).

Many biological reactions involve large conformational changes which provide well-defined mechanical reaction coordinates, for example, the end-to-end distance of a polymer, that can be used to follow the progress of the reaction (Bustamante et al. 2004; Kumar et al. 2007). Such processes have been modeled by a simple two-state model (Bustamante et al. 2004). The applied force “tilts” the free energy surface along the reaction coordinate by an amount linearly dependent on the end-to-end distance. The kind of transitions induced by the applied force are the folding–unfolding transition of proteins, the stretching and unzipping transition of double-stranded DNA, or the ball–string transition of a polymer (Kumar and Li 2010). From thermodynamics point of view, the change in energy of system can be categorized into components related to the heat exchanged and the work done on or performed by the system. If the change in energy of the system is quite low then the system remains in quasi-static equilibrium. The aim of this chapter is to provide the concepts from statistical mechanics to describe the force-induced transitions of biopolymers. In section “[Models](#)”, we discuss some of the models of polymers, which have been used to study biopolymers in the presence of a force f . Section “[Methods and Techniques](#)” dwells with few techniques used in statistical mechanics for polymers. Among these techniques, we choose an exact enumeration technique to study force-induced desorption of polymer adsorbed on the surface in section “[Applications](#)”. The method developed for homo-polymers can also be extended to study the protein unfolding and DNA unzipping in quasi static equilibrium. The chapter ends with some results based on model studies which are interesting for future experiments in section “[Conclusions](#)”.

Models

Macromolecules from living organisms are called biopolymers which vary in their size, shape, and function. Polymers including biopolymers are made up of long chains of “monomer” units (de Gennes 1979; des Cloizeaux and Jannink 1990;

Grosberg and Khokhlov 1994; Vanderzande 1998). The monomers can be of different natures. In DNA and RNA they are called nucleotides, and amino acids in proteins. Building blocks of simple artificial polymers can be group of just a few atoms, for example, CH_2 in polyethylene or a complex structure in amino acids. The difference between polymers and biopolymers is in their structure. Biopolymers have generally well-defined structures. The most intensive theoretical study related to biomolecules can be performed with the help of all-atom simulations. All-atom models, which provide the most detailed description on the atomistic level, include the local interaction and interaction involved in non-bonded monomers. The later include the (6–12) Lennard-Jones potential, the electrostatic interaction, and the interaction with environment (Kumar and Li 2010). However, such study for a long chain is computationally demanding and is difficult to handle analytically. In order to get the key features, physicists attempt to simplify complex structures of biopolymers as much as possible. Such process is termed as coarse graining. They model polymer chains as threads or necklaces made of beads on a string which describes some essential properties of biopolymers (de Gennes 1979; des Cloizeaux and Jannink 1990; Grosberg and Khokhlov 1994; Vanderzande 1998) but not always. In the literature, coarse-grained models are divided into two broad categories: (1) Continuum models and (2) Lattice models, which are briefly discussed below.

Continuum Models

We briefly describe three models, namely, the Gaussian Chain model, Freely Jointed Chain (FJC) model, and Worm Like Chain (WLC) model, which have been extensively used to describe the force-extension curves of biopolymers. The advantages of these models come from their simplicity and allowing one to derive analytical expressions in a simple form. In these models a polymer chain consists of N beads (monomers) of contour length L . A point in d -dimensional space represents each monomer and the distance between two consecutive monomers is $R_{i-1} - R_i$ (see Fig. 1). The energy with force f (along x -direction) in Gaussian model is expressed as (Doi and Edwards 1986)

$$F = \frac{3k_B T}{2b} \int_0^{L=Nb} ds \left(\frac{\partial \mathbf{r}(s)}{\partial s} \right)^2 - f \hat{x}_0 \cdot \int_0^{Nb} ds \frac{\partial \mathbf{r}(s)}{\partial s}, \quad (1)$$

where b is the effective bond length known as the Kuhn length and $\mathbf{r}(s)$ describes the local state at arc-length point s . For example, in the case of a flexible Gaussian chain, \mathbf{r} is a three-dimensional position vector, whereas for WLC case it represents the unit tangent vector. k_B is the Boltzmann constant and T the temperature of the system. Using the path integral technique (Doi and Edwards 1986; Kleinert 1990), the end-to-end distance distribution function of a chain under the force can be written as

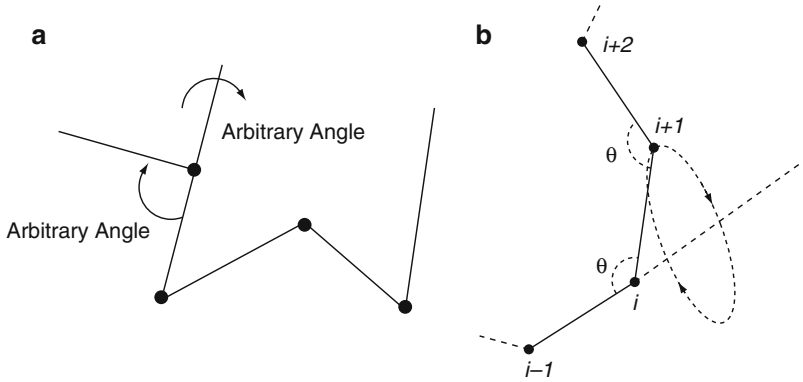


Fig. 1 (a) Schematic of a freely jointed chain. (b) Freely rotating chain with fixed bond angle. These models do not incorporate excluded volume effects in their description

$$P_N(\mathbf{R}, f) = \left(\frac{3}{2\pi N b^2} \right)^{(3/2)} \exp \left\{ -\frac{3}{2N b^2} \left(\mathbf{R} - \frac{N b^2 f \hat{x}_0}{3k_B T} \right)^2 \right\}. \quad (2)$$

The x-component of \mathbf{R} can be expressed as

$$\begin{aligned} P_N(R_x, f) &= \int dR_y dR_z P_N(\mathbf{R}, f) \\ &= \left(\frac{1}{2\pi N b^2} \right)^{(3/2)} \exp \left\{ -\frac{3}{2N b^2} \left(R_x - \frac{N b^2 f}{3k_B T} \right)^2 \right\}, \end{aligned} \quad (3)$$

which gives the expression for the extension x in the presence of applied force f (Dai et al. 2003)

$$x(f) = \frac{N b}{3} \frac{f b}{k_B T}. \quad (4)$$

This is a linear force relation. The model describes the response of a single polymer chain under low force. The major limitation of the model is the property that the distance between two monomers can be extended without any limit. This shortcoming of Gaussian chain is removed in the FJC model, where the distance between two consecutive monomers (bond length) is kept fixed while the rotational angle occurs with equal probability (Fig. 1b). The free energy of the system can be written as

$$F = -f \hat{x}_0 \cdot \sum_{n=1}^N \mathbf{r}_n, \quad (5)$$

where \mathbf{r}_n are bond vectors with constant length $|\mathbf{r}_n| = b$.

The distribution function of end-to-end vector \mathbf{R} in the presence of force f is defined as

$$P_N(\mathbf{R}, f) = \frac{\exp[\beta f \hat{x}_0 \cdot \mathbf{R}] P_N(\mathbf{R}, 0)}{\int d\mathbf{R} \exp[\beta f \hat{x}_0 \cdot \mathbf{R}] P_N(\mathbf{R}, 0)}, \quad (6)$$

where $\beta = 1/k_B T$ and $P_N(\mathbf{R}, 0)$ is the end-to-end distance distribution function in the absence of force f , which is given as

$$P_N(\mathbf{R}, 0) = \frac{1}{2^{N+1} (N-2)! \pi b^2 R} \sum_{n=0}^{[(N-R/b)/2]} (-1)^n \binom{N}{n} \times (N-2n-R/b)^{N-2}. \quad (7)$$

The component of end-to-end distance along the x -direction may be obtained by integrating Eq. 6 with respect to R_y and R_z as (Dai et al. 2003)

$$P_N(R_x, f) = \frac{\exp(\beta f R_x)}{[4\pi \sinh(\beta f b) / \beta f b]^N} P_N(R_x, 0), \quad (8)$$

where the value of $P_N(R_x, 0)$ can be obtained from Eq. 7.

The average extension along the force direction is

$$x(f) = bN \left[-\frac{1}{\beta f b} + \coth(\beta f b) \right]. \quad (9)$$

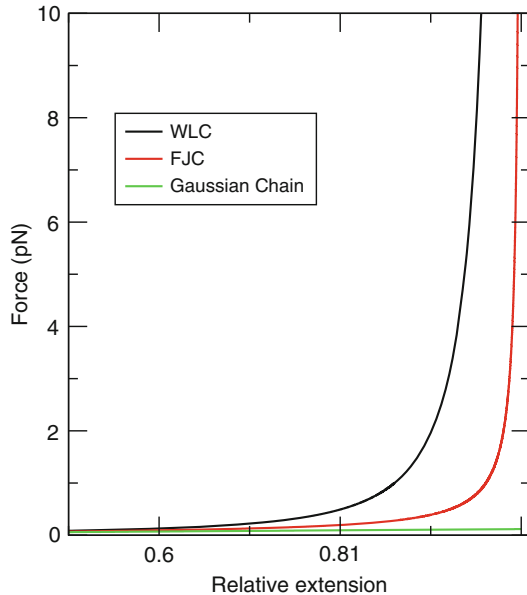
One of the important characteristics of biopolymers not described by the FJC model is the stiffness of the chain. A more realistic model for many systems is that of the Freely Rotating Chain, or Worm Like Chain (WLC), which includes stiffness in its description. This model builds an extension of the FJC model with the assumptions that the bond angles are fixed at certain angle, but are free to rotate. This gives rise to a uniform distribution of dihedral angles. The end-to-end distance within the WLC model can be calculated by transfer matrix methods, numerical simulations, etc. One can find variational expressions for x as a function of f (Chatney et al. 2004; Marko and Siggia 1995; Rosa et al. 2003a). The one used extensively in the literature is (Marko and Siggia 1995)

$$f = \frac{k_B T}{2b} \left[\frac{1}{(1-x/L)^2} - 1 + \frac{4x}{L} \right]. \quad (10)$$

This equation also reduces to the Gaussian chain result in the small force regime i.e., $f < k_B T/b$ and for large force $f > k_B T/b$, it acquires the form

$$x = L \left(1 - \sqrt{\frac{k_B T}{2fb}} \right). \quad (11)$$

Fig. 2 Force-extension curves for the Gaussian chain, FJC, and WLC models



One can see from this equation that the difference between the contour length L and the extension x varies as $f^{-0.5}$. The force-extension curves for the Gaussian chain, FJC and WLC are shown in Fig. 2. The behavior of a biopolymer subjected to a large force as described by the WLC model is contrary to the behavior as described by the FJC model, whereas at low force all models show identical behavior. The Kuhn length of the WLC model corresponds to twice the persistence length. One of the important shortcomings is that all these models have only one free parameter, known as the persistence length or Kuhn length, and hence are not suitable when it comes to describing the entire force-extension curve involving many intermediates.

It is important to recall that all these models ignore crucial excluded volume effect, i.e., the space occupied by a monomer is not available to other monomers (de Gennes 1979). Apart from these, in a polymer solution three types of interactions may be present (de Gennes 1979): (1) monomer–monomer interactions, (2) monomer–solvent interaction, (3) solvent–solvent interaction.

If monomer–monomer (or solvent–solvent) interaction is more than the monomer–solvent interaction in the solution, such solvent is known as a poor solvent. However, if monomer–solvent interaction favors in the solution, the solvent is referred as a good solvent. It is possible by lowering the temperature, one can go from the expanded coil state to a compact globule state and such transition is known as a coil–globule transition and transition point is called θ -point. Generally, force-induced transitions have been studied by these models where excluded volume effects and attractive interactions between chain segments have been ignored and is thus well suited only for modeling the stretching of polymers in a good solvent (Kumar and Li 2010).

Lattice Models

Any model designed to represent a polymer chain over a full range of physical conditions must include important effects such as excluded volume and attractive interactions as a starting point. In a simplified assumption, a linear biopolymer chain in a solvent can be described by a walk on a lattice in which a step or vertex of the walk represents a monomer. If the walk is allowed to cross itself without restriction such walk is referred as Random Walk, which is same as that of the Gaussian chain (de Gennes 1979; Vanderzande 1998). The excluded volume effect has been incorporated in the Random Walk model of polymer with a constraint that a lattice site cannot be visited more than once. This kind of walk is known as Self-Avoiding Walks (SAWs), which simulates a linear polymer chain in a good solvent (Fig. 3a). A polymer chain in a poor solvent is modeled by Self-Attracting Self-Avoiding Walks (SASAWs) by including self-attraction among non-bonded monomers. This model exhibits collapse transition including the existence of θ -point (de Gennes 1979; Vanderzande 1998). Various kinds of underlying lattices have been proposed in the literature to study the conformational properties of linear polymer chains. The choice of the lattice depends on the mathematical convenience, nature of the system,

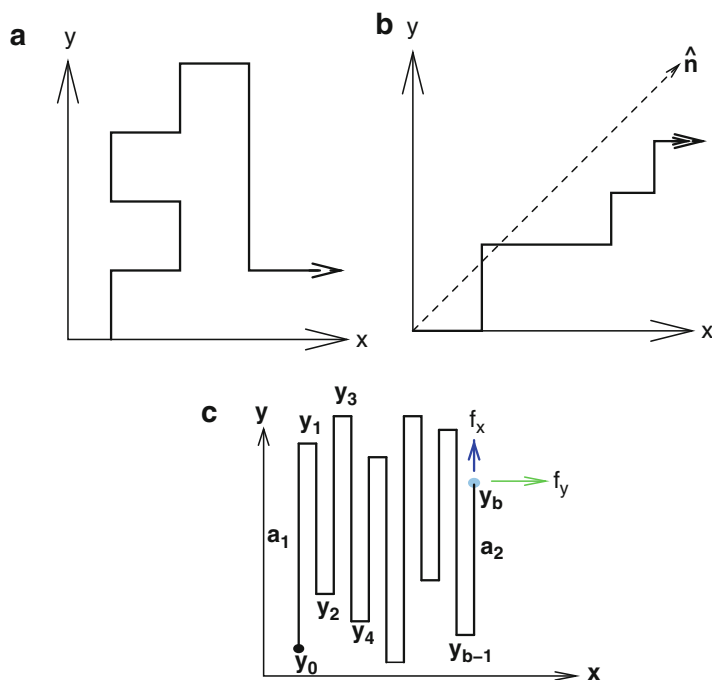


Fig. 3 Schematic diagrams of SASAW, DW, and PDSAW model. A force has been applied at one end (along x-axis) keeping other end fixed. In order to study the role of anisotropy, one may apply force along y-direction (Fig. 3c)

and the interactions present in the system. As far as universality is concerned, the nature of underlying lattice and the detail of interactions do not matter much and many important properties associated with polymers can be derived, which are in qualitative agreement with the experiments. Sometimes quantitative agreement has also been achieved particularly in determining the critical exponents (de Gennes 1979; Vanderzande 1998). Therefore, SAWs with suitable interactions on lattice has been studied extensively in describing the various properties of polymers and biopolymers.

In addition to self-avoidance, one can also introduce additional constraint on the walks, for example, certain direction(s) is (are) not accessible to the walker, such walks are called Directed Walks (DWs) (Privman and Svrakic 1989; Vanderzande 1998). To define a directed walk, a preferred direction \hat{n} on the lattice is assigned. Walkers are allowed to take only those steps in the non-negative direction of \hat{n} . The directed walk can be seen as a model of polymer that is subject to some external force in the direction of \hat{n} , for example, flow in which the polymer is immersed or an electric field acting on electrically charged polymers. If the direction \hat{n} is assigned as shown in Fig. 3b and walks in the non-negative projection of \hat{n} are not allowed then such walks are called Fully Directed Walks (FDWs). If the direction \hat{n} is assigned, say, along the x-axis (Fig. 3c), then the walks are said to be the Partial Directed Walks (PDSAWs). For example in two dimensions, the walker can take step in $\pm y$ directions but only in $+x$ direction. The major advantage of the directed walk model of polymer is that it can be solved analytically and many important results may be derived exactly. The drawback is that it is too far from the real chain.

Polymer and Critical Phenomena

It is known that the certain quantities associated with polymers modeled by SAWs for, example, number of distinct conformation (C_N), number of closed polygons (P_N), and end-to-end distance of chain (R_e) of N monomers scale as (de Gennes 1979)

$$\begin{aligned} C_N &\sim \mu^N N^{\gamma-1}, \\ P_N &\sim \mu^N N^{\alpha-3}, \\ R_e &\sim N^{\nu}, \end{aligned} \quad (12)$$

where μ is the connectivity constant giving the number of choices per step for an infinitely long walk. From the phase transition and critical phenomena, we also know that certain physical quantities like susceptibility (χ), specific heat (C), and correlation length (ξ) near the transition point T_C scale as

$$\begin{aligned} \chi &\sim \chi_0 \left| \frac{T - T_c}{T_c} \right|^{-\gamma}, \\ C &\sim C_0 \left| \frac{T - T_c}{T_c} \right|^{-\alpha}, \\ \xi &\sim \xi_0 \left| \frac{T - T_c}{T_c} \right|^{-\nu}. \end{aligned} \quad (13)$$

A relation between polymer statistics and phase transition was established by de Gennes (1979) and des Cloiseaux (1974) showing a correspondence between the polymer chain modeled by SAWs and the n -vector spin model of magnetization in the limit $n \rightarrow 0$. Similarities between correlation length (ξ) and the end-to-end distance (R_e) can be noticed by comparing Eqs. 12 and 13. Correspondence between $1/N$ and $\frac{T-T_c}{T_c}$ is viewed as $\frac{T-T_c}{T_c} \rightarrow 0$ and $N \rightarrow \infty$. This equivalence allowed polymer science to benefit from the vast knowledge accumulated in the study of critical phenomena. For example, the following relations, which are quite non-trivial to derive directly, are also valid here (de Gennes 1979):

$$\begin{aligned}\alpha + 2\beta + \gamma &= 2, \\ 2 - \alpha &= d\nu, \\ \gamma &= \nu(2 - \eta).\end{aligned}\tag{14}$$

Here, α , β , γ , and η are the critical exponents and d represents the dimensionality of the system. The above equivalence may serve an important role in describing the long range behavior of the polymers.

Methods and Techniques

Methods and techniques used in critical phenomena can also be applied to polymers and biopolymers to gain further insight of the system. In the following, we discuss some of the techniques of statistical mechanics which have been used in analyzing the results of single molecule force spectroscopy.

Generating Function Technique

The generating function technique is a very powerful technique, which may be adapted to study the conformational properties of biopolymers (Forgacs et al. 1995; Privman and Svrakic 1989). Here, we use a simple example to illustrate how the generating function approach can be applied and exact results can be derived. Let us consider a directed walk model of a polymer chain (Privman and Svrakic 1989; Vanderzande 1998). For simplicity, we consider partial directed walks (PDWs) in which the walker is allowed to move along x - and $\pm y$ -axes only (Fig. 3c). As discussed above biopolymers are in general semi-flexible. Stiffness is introduced into the polymer chain by putting an energy cost ε_b on every bend of the walk thus giving rise to an associated Boltzmann weight, $k = e^{-\frac{\varepsilon_b}{k_B T}}$. For $k = 1$ ($\varepsilon_b = 0$) the chain is said to be flexible, while for $0 < k < 1$, ($0 < \varepsilon_b < \infty$) the chain is said to be semi-flexible. The grand canonical partition function of such a chain can be written as (Mishra et al. 2003; Privman and Svrakic 1989)

$$\mathcal{Z}(z, k) = \sum_{N=0}^{\infty} \sum_{\text{all walks}} z^N k^{N_b}.\tag{15}$$

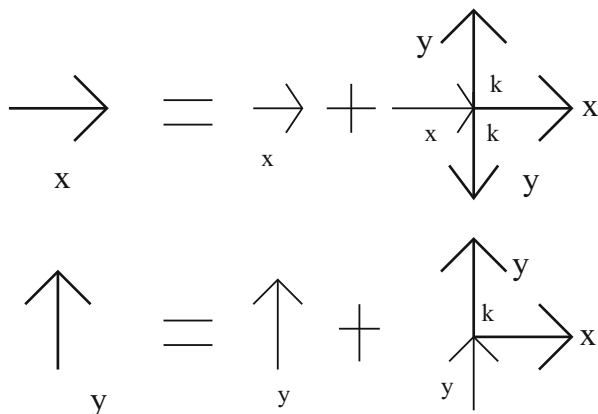


Fig. 4 The diagrammatic representations of the recursion relations Eqs. 16 and 17 for PDWs. The thick arrows X and Y denote all possible walks with the initial step (fugacity) along $+x$ and $\pm y$ directions

Here, z is the fugacity of the walk and N_b is the number of bends in a given configuration. In two dimensions, the grand canonical partition function defined by Eq. 15 may be expressed as a sum of two components of PDWs. The recursion relations for the case of a semi-flexible polymer chain are

$$X = z + z(X + 2kY), \tag{16}$$

and

$$Y = z + z(kX + Y). \tag{17}$$

Schematic representations of the above equations (for 2D) have been shown in Fig. 4. It may be noted that the first term of Eqs. 16 and 17 is the fugacity of the walk and remain constant as z . Solving Eqs. 16 and 17 we get

$$X = \frac{z + (2k - 1)z^2}{1 - 2z + z^2 - 2z^2k^2}, \tag{18}$$

$$Y = \frac{z + (k - 1)z^2}{1 - 2z + z^2 - 2z^2k^2}. \tag{19}$$

The partition function of the system can, therefore, be written as

$$\mathcal{Z}(z, k) = X + 2Y = \frac{(4k - 3)z^2 + 3z}{1 - 2z + z^2 - 2z^2k^2}. \tag{20}$$

The critical point for polymerization of an infinite chain is found from the relation

$$1 - 2z + z^2 - 2z^2k^2 = 0. \quad (21)$$

This leads to an expression for the critical value of the step fugacity as a function of k , $z_c = \frac{1}{(1+\sqrt{2k})}$. In the limit $k \rightarrow 1$, it reduces to the well-known value for the flexible polymer chain in 2D (Privman and Svrakic 1989). The approach has been successfully applied to study the DNA unzipping and unfolding of collapse polymer (Marenduzzo et al. 2001a, b; Rosa et al. 2003a, b).

Exact Enumeration Technique

The conformational properties of a polymer chain and phase transition phenomena can be understood if one has the complete information about the partition function of the chain. In the lattice model, the canonical partition function of a polymer chain is calculated (Domb and Lebowitz 1989) by enumerating all possible walks of a given length. For example, we show total number of conformations (C_N) for N step walk in Table 1. The grand canonical partition function defined in Eq. 15 can be written as

$$\mathcal{Z}(z) = \sum_N C_N z^N, \quad (22)$$

where z is the fugacity associated with each step of the walk. As the singularity of the partition function is associated with critical phenomena, the partition function defined in Eq. 22 will follow $\mathcal{Z}(z) \sim (1 - \mu z)^{-\gamma}$ in the thermodynamic limit. In

Table 1 Values of C_N for different N for the square lattice. The value of C_N can be obtained numerically or one may see the following webpage: <http://www.ms.unimelb.edu.au/iwan/saw/series/sqsaw.ser>

N	C_N	N	C_N
1	4	16	17245332
2	12	17	46466676
3	36	18	124658732
4	100	19	335116620
5	284	20	897697164
6	780
7	2172
8	5916
9	16268
10	44100
11	120292
12	324932
13	881500	53	99121668912462180162908
14	2374444	54	263090298246050489804708
15	6416596	55	698501700277581954674604

most of the cases, where model is not analytically solvable, one uses numerical techniques (Domb and Lebowitz 1989) to calculate the partition function. Once the partition function is known, other thermodynamic variables can be calculated.

Since in a system of finite size the true phase transition cannot take place, therefore, one has to use suitable extrapolation scheme to calculate C_N in the limit $N \rightarrow \infty$. For this purpose suitable techniques, for example, ratio method, Pade approximants, differential approximation, etc. (Domb and Lebowitz 1989) can be used. In ratio method, the approximate value of μ for $N \rightarrow \infty$ can be calculated by taking the ratios of consecutive terms of the series. These quantities should, for large N , be a linear function of $1/N$

$$\log \frac{C_{N+1}}{C_N} \simeq \log \mu + (\gamma - 1) \mathcal{O}(1/N), \quad (23)$$

where $\mathcal{O}(1/N)$ lower order correction. A simple fit then gives estimates for μ and γ (Domb and Lebowitz 1989). However, there is an odd-even effect in $\frac{C_{N+1}}{C_N}$. One can avoid this by using the square root of the successive ratios of C_N as $\mu = \sqrt{\frac{C_{N+2}}{C_N}}$ and extrapolate it to $N \rightarrow \infty$. Linear extrapolation of μ with $\frac{1}{N}$ is shown in Fig. 5.

Monte Carlo Simulation

The Monte Carlo (MC) technique is extremely simple in principle and has been greatly expanded and applied to a wide variety of problems (Binder 1995; Grass-

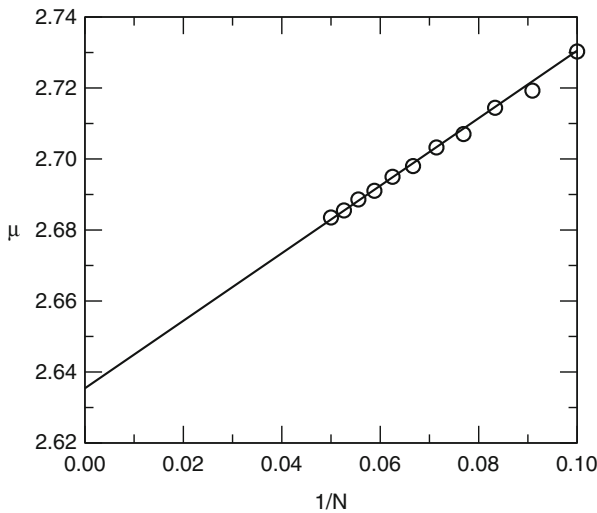


Fig. 5 Variation of connectivity constant (μ) with $1/N$ for SAW on square lattice. The linear extrapolation gives $\mu = 2.636$ which is quite close to the best estimated value

berger et al. 1999; Landau and Binder 2005). The typical goal of Monte Carlo simulations is first to find out the ground state of a given system (optimization problem) and then calculate ensemble averages through random sampling (equilibrium problem). The core of this approach is a set of predefined moves, which are traditionally on the lattice but may be defined in free space. At each time step a move is selected and may be accepted or rejected based on some criterion, for example, Metropolis method (Binder 1995), which has been used extensively, is defined as follows:

1. Start with some initial configuration I at a given temperature T
2. Define some rule to go from the present configuration to a new one
3. Compute the energy difference $\Delta E = E_{\text{new}} - E_{\text{old}}$, where E_{new} and E_{old} are energies of the system after and before the move
4. If $\Delta E < 0$, accept the move with new configuration. Otherwise, generate a random number r between 0 and 1 and accept the trial configuration if $\exp(-\beta\Delta E) > r$. Such choice of acceptance of move is known as Metropolis method
5. Repeat Steps 2–4 till enough configurations have been sampled
6. Repeat Steps 1–5 for different temperatures

The way in which temperature is decreased is known as the cooling process. Ideally one should try to devise an optimal way to find the annealing schedule, as the decrease rate and the number of MC steps per T can be varied during the numerical simulation. An important parameter to monitor during the annealing is the number of accepted moves that do not violate physical constraints. At high T , this number is very high but at low temperatures almost all moves may be rejected, and therefore, caution should be taken in sampling. One of the biggest advantages of the MC is that it can be very fast especially if the moves are selected carefully. The disadvantage is that it is not generally suitable for studying the dynamics since it is dependent on the move set, and method fails at low temperature. There are a number of good reviews and books available, and serious readers are advised to go through them (Binder 1997; Leckband and Israelachvili 2001; Muller et al. 2006; Muller-Plathe 1997).

Molecular Dynamics

One of the important tools that has been employed in statistical mechanics is the molecular dynamics (MD) simulations (Allen and Tildesley 1987; Frenkel and Smit 2002). Quite frequently, this technique has been used to study the biomolecules (Kumar and Li 2010). Before studying the structural and dynamical properties of biomolecules, it is important to note that these molecules exhibit a wide range of time scales over which specific processes take place. For example, local motion, which involves atomic fluctuation, side chain motion, and loop motion occurs in the length scale of 0.01–5 Å and the time involved in such process is of the order of 10^{-15} to 10^{-9} s. The motion of a helix, protein domain, or subunit falls under the

rigid body motion whose typical length scales are in between 1 Å and 10 Å and time involved in such motion is in between 10^{-9} and 1 s. Large-scale motion consists of helix–coil transition or folding–unfolding transition, which is more than 5 Å in length and, time involved is about 10^{-7} to 10^{-4} s. Hence the basic goal of molecular dynamics is to understand the role of different length and time scales involved in describing the physical phenomena. Since MD simulation provides information at microscopic level, it is desirable to use the basic concepts of statistical mechanics to derive the macroscopic observable like pressure, energy, heat capacity, etc.

In MD one generates a sequence of points in phase space $\Omega(p_i, q_i)$ where p_i is the momentum of the i th particle and q_i is the position as a function of time. These points belong to the same ensemble and correspond to different conformations (Allen and Tildesley 1987; Frenkel and Smit 2002). The ergodic hypothesis states that the ensemble average is equivalent to time average. Hence if one allows system to evolve in time indefinitely then system will pass through all possible states and thus corresponds to the time average. Therefore, the goal is to generate enough representative point such that equality is satisfied.

The method is based on Newton's second law of motion, that is, Frenkel and Smit (2002), Allen and Tildesley (1987)

$$\mathcal{F}_i = m_i \frac{d^2 r}{dt^2} = -\frac{dE_p}{dr_i}, \quad (24)$$

where \mathcal{F}_i is the applied force on the i -th particle of mass m_i and E_p is the potential energy of the system. The potential energy, in general, is a complicated function and there is no analytic solution to the equation of motion and therefore, one has to solve these equations numerically.

There are various algorithms/numeric schemes for integrating Newton's equations of motion (Allen and Tildesley 1987; Frenkel and Smit 2002). They are generally derived from the Taylor expansion. The expansion for the positions (\mathbf{r}), velocity (\mathbf{v}), and acceleration (\mathbf{a}) of the particles around some moment of time t are

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \dots, \quad (25)$$

$$v(t + \Delta t) = v(t) + a(t)\Delta t + \frac{1}{2}b(t)\Delta t^2 + \dots, \quad (26)$$

$$a(t + \Delta t) = a(t) + b(t)\Delta t + \dots, \quad (27)$$

where Δt is the time step used in the simulation. In the following, we will discuss some of the common schemes of integration algorithm, which have been extensively used in literature.

- **Verlet algorithm:**

The Verlet algorithm uses positions and accelerations at time t and the positions from time $t - \Delta t$ to calculate new positions at time $t + \Delta t$. In order to derive the Verlet algorithm one can write

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2, \quad (28)$$

$$r(t - \Delta t) = r(t) - v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2. \quad (29)$$

The sum of these equations gives

$$r(t + \Delta t) = 2r(t) - r(t - \Delta t) + a(t)\Delta t^2. \quad (30)$$

The algorithm does not use the explicit form of velocities. The usefulness of the algorithm are it is simple to implement, and does not require large memory. The algorithm is of moderate precision.

- **Leap Frog algorithm:**

In this algorithm, the velocities are first computed at time $t + \frac{\Delta t}{2}$ and then used to calculate the positions, r , at time $t + \Delta t$. The velocities leap over the positions and subsequently the positions leap over the velocities.

$$r(t + \Delta t) = r(t) + v\left(t + \frac{\Delta t}{2}\right)\Delta t, \quad (31)$$

$$v(t + \Delta t) = v\left(t - \frac{\Delta t}{2}\right) + a(t)\Delta t. \quad (32)$$

The approximate form of velocities at time t is given by

$$v(t) = \frac{1}{2}\left[v\left(t + \frac{\Delta t}{2}\right) + v\left(t - \frac{\Delta t}{2}\right)\right]. \quad (33)$$

The major gain in this process is that the velocities are explicitly calculated, however, the shortcoming of the algorithm is that velocities are not calculated at the same time as the positions.

- **The Velocity Verlet algorithm:**

It may be noted that in the Leap Frog algorithm, velocities are not defined at the same time as the positions. As a consequence kinetic and potential energies are also not defined at the same time, and hence one cannot directly compute the total energy in the Leap Frog algorithm. It is, however, possible to implement in the Verlet algorithm a form that uses positions and velocities computed at same instant of times.

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2, \quad (34)$$

$$v(t + \Delta t) = v(t) + \frac{1}{2}[a(t) + a(t + \Delta t)]\Delta t. \quad (35)$$

It has to be emphasized here that the Velocity Verlet algorithm is not memory consuming, because it is not required to keep track of the velocity at every time step during the simulation. Moreover, the long-term results of Velocity Verlet are equivalent to the Semi-implicit Euler method, and there is no compromise on precision.

- **Beeman's algorithm:**

This algorithm is related to the Verlet algorithm and yields the same trajectories. However, it provides better estimate of the velocities and looks different than the Verlet algorithm.

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{2}{3}a(t)\Delta t^2 - \frac{1}{6}a(t - \Delta t)\Delta t^2, \quad (36)$$

$$v(t + \Delta t) = v(t) + v(t)\Delta t + \frac{1}{3}a(t)\Delta t + \frac{5}{6}a(t)\Delta t - \frac{1}{6}a(t - \Delta t)\Delta t. \quad (37)$$

The advantage of this algorithm is that it gives accurate expression for the velocities and better energy conservation. The disadvantage is that because of the complex expressions computation is much more time-consuming.

The procedure for a molecular dynamics simulation is subject to many user-defined variables. However, one should consider certain criteria in choosing these algorithms. It should conserve energy, momentum and computationally efficient so that long time step integration can be performed. Because the evaluation of the atomic positions is not performed on a continuous basis, but at intervals of a femtosecond (fs). Since this is the time-scale of stretches of the bonds with hydrogen atoms, these stretches should be constrained in order to permit time steps of 2 fs. The algorithm that permits this increase in simulation speed is called SHAKE (Frenkel and Smit 2002). Application of constraints on all bond lengths and bond angles would increase the permitted time step even more, without too much loss of information. Since most of the computational time is spent on evaluating the non-bonded interactions between atoms, the evaluation time step for non-bonded atom pairs can be increased to a small extent. Using a cutoff distance beyond which atoms are no longer considered to interact can significantly reduce the amount of non-bonded atom pairs.

It may be noted that MD is a scheme for studying the time evolution of a classical system of N particles in volume V , where total energy is a constant of motion. As pointed out earlier, time averages are equivalent to ensemble averages, the observables, say A evolves according to constant energy simulation and hence,

$$\langle A \rangle_{N,V,E} = \lim_{\tau \rightarrow \infty} \int_0^\tau A(t) dt. \quad (38)$$

It is possible to run many NVE simulations with different energies and result could be used in a Boltzmann-weighted average as

$$\langle A \rangle_{NVT} = \sum_E \langle A \rangle_{NVE} e^{-\beta E}. \quad (39)$$

However, such process is not very practical, because simulations are quite expensive in terms of time. Moreover, we require a lot of E values to have enough statistics, and to conserve energy during simulation is a tough task. Thus, it is more appropriate to run simulation in constant temperature ensemble, that is, NVT. In the following, we will discuss some popular temperature control schemes (Allen and Tildesley 1987; Frenkel and Smit 2002).

1. Scaling velocities (Berendsen thermostat):

Since in the beginning, the system is not equilibrated and hence remains in the incorrect state. The potential energy is converted to thermal energy or thermal energy is being consumed during equilibration. In order to stop the temperature from drifting, one of the possible ways is to constrain instantaneous velocities scaled by a factor at each step

$$\mathbf{v}'_i = \sqrt{\frac{T_d}{T_i}} \mathbf{v}_i, \quad (40)$$

where T_d is desired temperature and T_i is instantaneous temperature given by $T_i = \frac{2}{3Nk_B} \sum_i^N \frac{p_i^2}{2m_i}$ after i -th step. This influences the dynamics severely and does not give correct canonical ensemble because it does not generate fluctuations in temperature which are also present in the canonical ensemble. The method is good for initialization of phase.

A weaker formulation of this approach is the Berendsen thermostat where to keep temperature constant, system is coupled to an external heat bath of temperature T_0 . The velocities are scaled in such a way that the rate of change of temperature is proportional to the difference in temperature between system and bath, that is,

$$\frac{dT}{dt} = \frac{T_0 - T(t)}{c}, \quad (41)$$

where c is coupling parameter of the system with a bath. The scaling factor for velocities is given by

$$\lambda = \sqrt{1 + \frac{\Delta t}{c} \left(\frac{T_0}{T(t - \Delta t/2)} - 1 \right)}. \quad (42)$$

2. Adding stochastic forces and/or velocities (Langevin thermostat):

This thermostat models the influence of a heat bath by adding to the velocity of each particle a small random (white) noise and a frictional force directly proportional to the velocity of that particle. These two factors are balanced to give a constant temperature. Since each particle is coupled to a local heat bath, the heat trapped in localized modes can be removed by using this model.

$$ma = -\xi v + f(r) + f'. \quad (43)$$

Here $f(r)$ and f' are conservative force and random force respectively. ξ is the friction coefficient. The drawback of this thermostat is that momentum transfer is destroyed. So, it is not advisable to use Langevin thermostat in the simulations where one wishes to study the diffusion processes.

3. Nosé–Hoover thermostat:

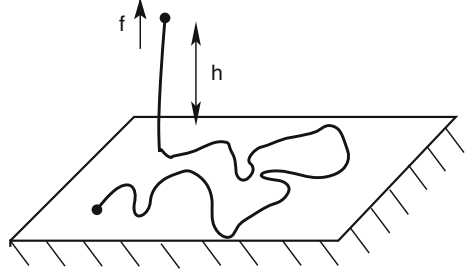
The Berendsen thermostat is quite efficient for relaxing a system to the target temperature. However, once your system has reached equilibrium, it is important to probe a correct canonical ensemble. The Nosé–Hoover thermostat is an extended-system method for controlling the temperature of simulated system (Huenenberger 2005; Thijssen 1999). It allows temperatures to fluctuate about an average value, and uses a friction factor to control particle velocities. This particular thermostat can oscillate when a system is not in equilibrium. Therefore, it is recommended to use a weak-coupling method for initial system preparation (e.g., Berendsen thermostat), followed by data collection under the Nosé–Hoover thermostat. This thermostat produces a correct kinetic ensemble.

In this context, molecular dynamics simulation of nucleic acids have been considered as a bigger challenge because of the negative backbone charge and the poly-electrolytes behavior. In oligonucleotide dynamics simulations, particular attention should be paid to the atomic charges. The negatively charged phosphate groups may very well influence the trajectory. In contrast with molecular mechanics where a structural minimum is the end result, molecular dynamics offers so much information that is hard to quantify. Excellent reviews on molecular dynamics and its use in biochemistry and biophysics are numerous (see, e.g., Adcock and McCammon 2006 and references therein).

Applications

FJC and WLC models have been used in analyzing the force-extension curve obtained from SMFS experiments. In this context, molecular dynamics and Monte Carlo simulations have been found to be quite useful in understanding the phenomena. Since there is a lot of good literature available in this subject, we are not going

Fig. 6 Schematic diagram of the polymer chain adsorbed on the surface under the application of external force



to discuss them here again. In this section, we shall illustrate that lattice model along with exact enumeration technique which may provide enhanced insight in the mechanism involved in the force-induced transitions because exact density of states are available and the entire phase diagram can be probed exactly for small chains.

We consider SAWs that start from a point on an impenetrable surface and experience a force f in a direction perpendicular to the surface at the other end as shown in Fig. 6. The applied force, because of its direction, favors desorption and one expects a critical force (f_c) for the desorption. At a given temperature (T) when the applied force f is less than $f_c(T)$ the polymer will be adsorbed, while for $f > f_c(T)$ the polymer will be desorbed. The curve $f_c(T)$, therefore, gives the boundary that separates the desorbed phase from the adsorbed phase in the force–temperature (f, T) plane (Mishra et al. 2005). Here, $C_N(N_s, h)$ corresponds to the number of SAWs of N steps having N_s number of monomers on the surface and h , the height of the end-monomer away from the surface. In this case, the partition function may be defined as

$$Z_N(\omega, u) = \sum_{N_s, h} C_N(N_s, h) \omega^{N_s} u^h, \quad (44)$$

where $\omega = e^{-\varepsilon_s/k_B T}$ and $u = e^{f/k_B T}$ are the Boltzmann weights for the surface interaction ($\varepsilon_s < 0$) and the applied force respectively. In the following, we set the Boltzmann constant $k_B = 1$ and $\varepsilon_s = -1$. For a fixed force f , we locate the adsorption–desorption transition temperature from the maximum of fluctuation in number of adsorbed monomers (i.e., $\langle N_s^2 \rangle - \langle N_s \rangle^2$) (see Fig. 7), where $\langle N_s \rangle$ and $\langle N_s^2 \rangle$ are defined as:

$$\langle N_s \rangle = \frac{1}{Z_N(\omega, u)} \sum_{N_s, h} N_s C_N(N_s, h) \omega^{N_s} u^h, \quad (45)$$

and

$$\langle N_s^2 \rangle = \frac{1}{Z_N(\omega, u)} \sum_{N_s, h} N_s^2 C_N(N_s, h) \omega^{N_s} u^h. \quad (46)$$

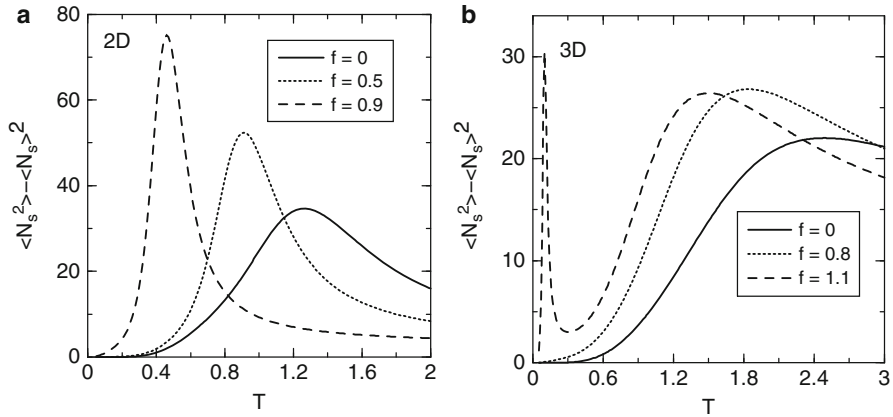


Fig. 7 The dependence of critical force $f_c(T)$ on T in (a) 2D and (b) 3D. The star corresponds to results obtained from the extrapolated values of the reduced free energy, and cross corresponds to the value obtained from finite size data of a step $N = 20$ (3D) and $N = 31$ (2D) respectively

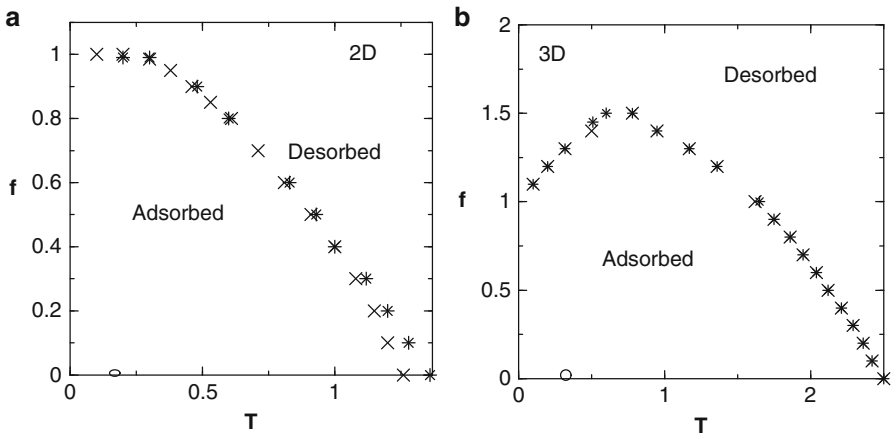


Fig. 8 The dependence of critical force $f_c(T)$ on T in (a) 2D and (b) 3D. The star corresponds to results obtained from the extrapolated values of the reduced free energy, and cross corresponds to the value obtained from finite size data of a step $N = 20$ (3D) and $N = 31$ (2D) respectively

The force–temperature phase diagram is shown in Fig. 8. The occurrence of two maxima in fluctuation curve gives the signature of re-entrance in 3D, but absent in 2D (Mishra et al. 2005). Using the phenomenological argument and the probability distribution analysis, it was shown that the ground state entropy is responsible for the re-entrance which is absent in 2D (Mishra et al. 2005).

It is possible to obtain better estimates of phase boundaries by extrapolating μ for the large N . The reduced free energy per monomer in this case is defined as

$$G(\omega, u) = \lim_{N \rightarrow \infty} \frac{1}{N} \log Z_N(\omega, u) = \log \mu_N(\omega, u). \quad (47)$$

$\mu_N(\omega, u)$ can be estimated from the partition functions found from the data of exact enumerations for finite N by extrapolating to large N . The fluctuation in terms of reduced free energy is given as $\frac{\partial^2 G}{\partial(\log \omega)^2}$.

The method described above has been applied to protein unfolding (Kumar and Giri 2005, 2007; Kumar et al. 2007), stretching of DNA (Kumar and Mishra 2008; Mishra et al. 2009), and unzipping of DNA (Giri and Kumar 2006; Kumar et al. 2005) and, many useful information about biomolecules have been derived which are in qualitative agreement with experiments.

Conclusions

We have discussed some basic models of biopolymers and few techniques which have been used extensively in the past to understand the mechanism involved in force-induced transitions. In particular, we showed that lattice model along with exact enumeration technique may be used to interpret the results of SMFS. It is important to point out here that all the single molecule experiments have been performed to understand the structure and function of biopolymers in vivo by analyzing it in vitro. As a result, the effect of cellular environment has been ignored in all these studies. It is known that the interior of the cell contains different kinds of biomolecules like sugar, nucleic acids, lipids, etc. These macromolecules occupy about 40% of the total volume with steric repulsion among themselves. This confined environment induces phenomena like “molecular confinement” and “molecular crowding” and has major thermodynamic and kinetic consequences on the cellular processes. Recently Singh et al. (2009a, b) have used exact enumeration technique and showed that the molecular crowding has significant impact on the force-induced transitions. In order to have better understanding of force-induced transitions, it is advisable to perform SMFS experiments in the environment similar to a cell.

Acknowledgments We would like to thank D. Giri, A. R. Singh, and G. Mishra for many helpful discussions. Financial assistance from the Department of Science and Technology, New Delhi is gratefully acknowledged.

Bibliography

- Adcock, S. A., & McCammon, J. A. (2006). Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chemical Reviews*, 106, 1589–1615.
- Allen, M. P., & Tildesley, D. J. (1987). *Computer simulations of liquids*. Oxford: Oxford Science.
- Bhattacharjee, S. M. (2000). Unzipping DNAs: Towards the first step of replication. *Journal of Physics A*, 33, L423–L428.

- Binder, K. (1995). *Monte Carlo and molecular dynamics simulations in polymer science*. New York: Oxford University Press.
- Binder, K. (1997). Applications of Monte Carlo methods to statistical physics. *Reports on Progress in Physics*, 60, 487–559.
- Bustamante, C., Smith, S. B., Liphardt, J., & Smith, D. (2000). Single-molecule studies of DNA mechanics. *Current Opinion in Structural Biology*, 10, 279–285.
- Bustamante, C., Chemla, Y. R., Forde, N. R., & Izhaky, D. (2004). Mechanical processes in biochemistry. *The Annual Review of Biochemistry*, 73, 705–748.
- Ceccconi, C., Shank, E. A., Bustamante, C., & Marqusee, S. (2005). Direct observation of the three-state folding of a single protein molecule. *Science*, 309, 2057–2060.
- Chatney, D., Cocco, S., Monasson, R., & Thieffry, D. (2004). *Multiple aspects of DNA and RNA: From biophysics to bioinformatics: Lecture notes of the Les Houches Summer School*. Amsterdam: Elsevier.
- Cloizeaux, J. D. (1974). Langrangian theory for a self-avoiding random chain. *Physical Review A*, 10, 1665–1669.
- Dai, L., Liu, F., & Ou-Yang, Z.-C. (2003). Maximum-entropy calculation of the end-to-end distance distribution of force-stretched chains. *Journal of Chemical Physics*, 119, 8124.
- de Gennes, P. G. (1979). *Scaling concepts in polymer physics*. Ithaca/London: Cornell University Press.
- des Cloizeaux, J., & Jannink, G. (1990). *Polymers in solution*. Oxford: Clarendon.
- Doi, M., & Edwards, S. F. (1986). *The theory of polymer dynamics*. Oxford: Clarendon.
- Domb, C., & Lebowitz, J. L. (1989). *Phase transition and critical phenomena* (Vol. 13). New York: Academic.
- Forgacs, G., Lipowsky, R., & Nieuwenhuizen, T. M. (1995). *The behaviour of interfaces in ordered and disordered systems* (Vol. 14). Oxford: Clarendon.
- Frenkel, D., & Smit, B. (2002). *Understanding molecular simulation*. London: Academic.
- Giri, D., & Kumar, S. (2006). Effects of the eye phase in DNA unzipping. *Physical Review E*, 73, 050903(R).
- Grassberger, P., Nadler, W., & Barkema, G. T. (1999). *The Monte Carlo approach to biopolymers and protein folding*. Singapore: World Scientific.
- Grosberg, A. Y., & Khokhlov, A. R. (1994). *Statistical physics of macromolecules*. New York: American Institute of Physics.
- Huelsenberger, P. (2005). Thermostat algorithms for molecular dynamics simulations. *Advances in Polymer Science*, 173, 105–149.
- Kleinert, H. (1990). *Path integrals in quantum mechanics, statistics, and polymer physics*. Singapore: World Scientific.
- Kumar, S. (2009). Can reentrance be observed in force induced transitions? *Europhysics Letters*, 85, 38003.
- Kumar, S., & Giri, D. (2005). Force-induced conformational transition in a system of interacting stiff polymers: Application to unfolding. *Physical Review E*, 72, 052901.
- Kumar, S., & Giri, D. (2007). Does changing the pulling direction give better insight into biomolecules? *Physical Review Letters*, 98, 048101.
- Kumar, S., Giri, D., & Bhattacharjee, S. M. (2005). Force induced tripple point for interacting polymers. *Physical Review E*, 71, 051804.
- Kumar, S., Jensen, I., Jacobsen, J. L., & Guttman, A. J. (2007). Role of conformational entropy in force induced biopolymer unfolding. *Physical Review Letters*, 98, 128101–128104.
- Kumar, S., & Li, M. (2010). Biomolecules under mechanical force. *Physics Reports*, 486, 1–74.
- Kumar, S., & Mishra, G. (2008). Force-induced stretched state: Effects of temperature. *Physical Review E*, 78, 011907.
- Landau, D. P., & Binder, K. (2005). *A guide to Monte Carlo simulations in statistical physics*. New York: Cambridge University Press.
- Leckband, D., & Israelachvili, J. (2001). Intermolecular forces in biology. *Quarterly Review of Biophysics*, 34, 105–267.

- Lubensky, D. K., & Nelson, D. R. (2000). Pulling pinned polymers and unzipping DNA. *Physical Review Letters*, *85*, 1572–1575.
- Marenduzzo, D., Bhattacharjee, S. M., Maritan, A., Orlandini, E., & Seno, F. (2001a). Dynamical scaling of the DNA unzipping transition. *Physical Review Letters*, *88*, 028102.
- Marenduzzo, D., Trovato, A., & Maritan, A. (2001b). Phase diagram of force-induced DNA unzipping in exactly solvable models. *Physical Review E*, *64*, 031901.
- Marko, J., & Siggia, E. (1995). Stretching DNA. *Macromolecules*, *28*, 8759–8770.
- Mishra, G., Giri, D., & Kumar, S. (2009). Stretching of a single stranded DNA: Evidence for structural transition. *Physical Review E*, *79*, 031930.
- Mishra, P. K., Kumar, S., & Singh, Y. (2003). A simple and exactly solvable model for a semi flexible polymer chain interacting with a surface. *Physica A*, *323*, 453–465.
- Mishra, P. K., Kumar, S., & Singh, Y. (2005). Force-induced desorption of a linear polymer chain adsorbed on an attractive surface. *Europhysics Letters*, *69*, 102–108.
- Muller, M., Katsov, K., & Schick, M. (2006). Biological and synthetic membranes: What can be learned from a coarse-grained description? *Physics Reports*, *434*, 113–176.
- Muller-Plathe, F. (1997). Coarse-graining in polymer simulation: From the atomistic to the mesoscopic scale and back. *ChemPhysChem*, *3*, 754–769.
- Privman, V., & Svratkic, N. M. (1989). *Directed models of polymers, interfaces, and clusters*. Berlin: Springer.
- Rief, M., Clausen-Schaumann, H., & Gaub, H. E. (1999). Sequence-dependent mechanics of single DNA molecules. *Nature Structural Biology*, *6*, 346–349.
- Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M., & Gaub, H. E. (1997). Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science*, *276*, 1109–1112.
- Rosa, A., Hoang, T. X., Marenduzzo, D., & Maritan, A. (2003a). Elasticity of semiflexible polymers with and without self-interactions. *Macromolecules*, *36*, 10095–10102.
- Rosa, A., Marenduzzo, D., Maritan, A., & Seno, F. (2003b). Mechanical unfolding of directed polymers in a poor solvent: Critical exponents. *Physical Review E*, *67*, 041802.
- Singh, A. R., Giri, D., & Kumar, S. (2009a). Force induced unfolding of bio-polymers in a cellular environment: A model study. *Journal of Chemical Physics*, *131*, 065103.
- Singh, A. R., Giri, D., & Kumar, S. (2009b). Effects of molecular crowding on stretching of polymers in poor solvent. *Physical Review E*, *79*, 051801.
- Smith, S. B., Finzi, L., & Bustamante, C. (1992). Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science*, *258*, 1122–1126.
- Smith, S. B., Cui, Y., & Bustamante, C. (1996). Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. *Science*, *271*, 795–799.
- Thijssen, J. M. (1999). *Computational physics*. Cambridge: Cambridge University.
- Tskhovrebova, L., Trinick, K., Sleep, J. A., & Simons, R. M. (1997). Elasticity and unfolding of single molecules of the giant muscle protein titin. *Nature*, *387*, 308–312.
- Vanderzande, C. (1998). *Lattice models of polymers*. Cambridge: Cambridge University Press.
- Zhou, H. J., Zhou, J., Ou-Yang, Z. C., & Kumar, S. (2006). Collapse transition of two-dimensional flexible and semiflexible polymers. *Physical Review Letters*, *97*, 158302.

Part II

**Applications of Computational Methods to
Model Systems**

Anna Kaczmarek-Kedziera and Heribert Reis

Jon Baker

Contents

Introduction	424
Molecular Structure	425
He_2^+ : An Illustrative Example	426
The Newton–Raphson Step	428
The Hessian Matrix and Hessian Updates	430
Transition State Searches	432
Choice of Coordinates	433
The Modified Newton–Raphson Step	435
GDIIS	438
Geometry Optimization and Symmetry	439
Performance for Minimization	440
Minimization: An Example	443
Transition State Searches: An Example	451
Using Optimized Potential Scans in Transition State Searches	458
Comparison of Experimental and Theoretical Geometries	460
Geometry Optimization of Molecular Clusters	463
Geometry Optimization in the Presence of External Forces	464
Molecular Vibrations	465
1,2-Dichloroethane: An Illustrative Example	471
Scaled Quantum Mechanical Force Fields	478
1,2-Dichloroethane: A Further Analysis	479
Density Functional Theory and Weight Derivatives	487
Conclusions	491
Bibliography	492

J. Baker (✉)

Parallel Quantum Solutions, Fayetteville, AR, USA

e-mail: baker@pqs-chem.com

Abstract

This chapter deals with two very important aspects of modern ab initio computational chemistry: the determination of molecular structure and the calculation, and visualization, of vibrational spectra. It deals primarily with the practical aspects of determining molecular structure and vibrational spectra computationally. Both minima (i.e., stable molecules) and transition states are discussed, as well as infrared (IR), Raman, and vibrational circular dichroism (VCD) spectra, all of which can now be computed theoretically.

Introduction

This chapter deals with two very important aspects of modern ab initio computational chemistry: the determination of molecular structure and the calculation, and visualization, of vibrational spectra. The two things are intimately related as, once a molecular geometry has been found (as a stationary point on a potential energy surface at whatever level of theory is being used) it has to be characterized, which usually means that it has to be confirmed that the structure is a genuine minimum. This of course is done by vibrational analysis, i.e., by computing the vibrational frequencies and checking that they are all real.

A large percentage of the total expenditure in CPU cycles devoted to computational chemistry (variously estimated at between 60% and 85%) is spent optimizing geometries. In order to calculate various molecular properties, one first needs a reliable molecular structure so this is perhaps not surprising. Algorithms for geometry optimization are now highly advanced and usually very efficient and most of the quantum chemistry programs available for general use have solid and reliable geometry optimization modules. They also nearly all have analytical second derivatives, at least for the most common theoretical methods, which makes it relatively straightforward to compute vibrational frequencies once a structure has been found.

In this chapter I deal primarily with the practical aspects of determining molecular structure and vibrational spectra computationally. I consider both minima (i.e., stable molecules) and transition states, as well as infrared (IR), Raman, and vibrational circular dichroism (VCD) spectra, all of which can now be computed theoretically. The program used to carry out the calculations presented here is the PQS package developed by Parallel Quantum Solutions (PQS 2010), although any modern general purpose package (e.g., Gaussian, Turbomole, GAMESS) would do just as well. As the name implies, all the major ab initio functionality of this package is fully parallel, including energies, gradients, and second derivatives. PQS was chosen because (a) it is the package I actually use in all my application work; and (b) I am one of its principal authors. A review of the capabilities and parallel efficiency of the PQS package was published recently (Baker et al. 2009).

I have elected to use a standard level of ab initio theory for all of the examples presented in this chapter, namely, density functional theory (Hohenberg and Kohn

1964; Kohn and Sham 1965) (DFT) using the B3LYP (Becke 1993) (see also Hertwig and Koch (1997)) functional and the 6-31G* (Ditchfield et al. 1971) basis set (B3LYP/6-31G*). DFT is now the method of choice for routine chemical applications, and B3LYP – despite the large number of functionals developed since – is still one of the most popular. Many of the techniques and pitfalls in locating stable geometries are essentially independent of method, although DFT has its own issues as a result of the numerical quadrature required to handle many of the integrals; this will be discussed in some detail later. DFT is so popular that Hartree–Fock theory, which used to be the standard approach throughout the 1980s, has now almost disappeared other than as a precursor for higher level post-SCF calculations. At the time of writing, the latter are beginning to make a bit of a comeback as the limitations of DFT are being reached, in particular its inability to systematically improve the wave function (and hence the results). A good introduction to DFT for those, like me, whose background is in traditional quantum chemistry is the now classic 1993 paper from (Johnson et al. 1993).

The Born–Oppenheimer approximation (Born and Oppenheimer 1927) (see also Wikipedia (2010)) is used throughout. This extremely important approximation, namely that electrons, being so much less massive, respond instantaneously to the motion of the nuclei, underpins the whole concept of the potential energy surface (PES). Stable arrangements of nuclei (stable or meta-stable molecules) correspond to local minima on the PES, first-order saddle points connecting two different minima correspond to transition states, and the “valleys” joining transition states to minima correspond to “reaction paths.” Electronic excitation is represented by a “jump” from one potential energy surface to another. We will be spending all of our time on the lowest energy (ground-state) PES.

Molecular Structure

From a theoretical point of view the determination of a “molecular structure” means determining the geometry (i.e., the positions of atoms relative to one another) at a minimum (the bottom of a well) on the ground-state potential energy surface. Although a prerequisite, locating a minimum does not automatically mean that you have found the structure of a stable molecule; this depends on the barrier height, i.e., the energy required to get out of the well. The barrier height can be determined by locating the transition state, which for our purposes is defined as the highest energy point on the lowest energy path between reactants and products. (In a barrierless reaction, this will effectively be the energy of the products, assuming the reactant is taken to have the lower energy.) As a rough rule of thumb, if the barrier height is less than about 12–15 kcal mol⁻¹, the system is kinetically unstable at room temperature.

Determining the geometry of a stable molecule therefore requires knowledge of both the minimum itself and the barrier height (i.e., effectively locating the transition state) for all possible decomposition reactions. Of course, the structures of many stable molecules can be derived from “chemical intuition” based on

preexisting knowledge, but for new structures in particular the transition state is equally important and the practicing theoretician must be able to find both. In this section we discuss the methods available to do so.

He₂²⁺: An Illustrative Example

The simplest possible molecular orbital approach to the bonding in the hydrogen molecule (H₂) would consider the overlap between two 1s atomic orbitals on each hydrogen producing two molecular orbitals (MOs), one of which is bonding (the in-phase combination) and the other antibonding (the out-of-phase combination). H₂ is stable because its two electrons occupy the bonding orbital, giving an overall energy lowering compared to two separate H atoms.

Similar considerations applied to helium would suggest, quite rightly, that He₂ would not be stable, since both the bonding and antibonding MOs would be occupied, giving no net bonding. (In fact, as is well known, the antibonding MO is more antibonding than the bonding MO is bonding.) What happens if the electrons in the antibonding MO are removed? On the one hand, there should be a strong bonding interaction, while on the other the resultant positive charge should cause the two nuclei to repel one another.

Figure 1 shows the ground-state PES of He₂²⁺. This was obtained via a potential scan, varying the He–He distance from 0.6 to 4.0 Å in steps of 0.01 Å and computing the energy at each scanned distance. (An optimized potential scan – scanning one particular variable or combination of variables while optimizing all remaining degrees of freedom – is a useful tool for locating transition states, as will be seen later.) The upper curve shows the energy at the restricted RB3LYP/6-31G* level, while the lower curve is the unrestricted equivalent (UBLYP/6-31G*). The two curves coalesce at He–He distances less than about 0.96 Å; at distances greater than this the restricted wave function is energetically unstable, i.e., a lower energy can be obtained by switching to the unrestricted formalism.

During the restricted (upper) scan full symmetry ($D_{\infty h}$) was maintained and the system is a pure singlet. For the unrestricted scan, symmetry was turned off and the highest occupied (HOMO) and lowest unoccupied (LUMO) MOs, both alpha and beta spin, were allowed to mix; this results – as can be seen – in an ultimately much lower energy, but the spin symmetry is lost and the system wave function is no longer a pure singlet. (Spin contamination gradually rises with increasing He–He distance until at 4.0 Å $\langle S^2 \rangle$ is 1.000 and the multiplicity is 2.236, i.e., the system is somewhat “greater” than a doublet.)

At B3LYP/6-31G* the energy of two separate He⁺ ions is $-3.9863 E_h$ and so, despite the spin contamination, the unrestricted curve gives a far better description of the dissociation of He₂²⁺ than the restricted one. This is a general feature of the dissociation of essentially all diatomic molecules and is well known.

The interesting feature from our point of view is that this simple, one-dimensional PES contains both a (local) minimum *and* a transition state, and both types of stationary point can be examined and discussed in relation to the *same*

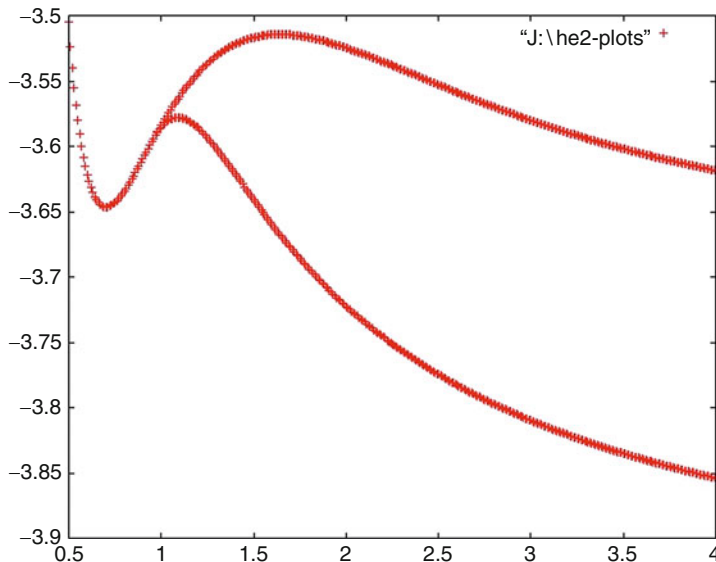


Fig. 1 Energy versus He–He distance (Å) for He_2^+ (RB3LYP/6-31G* – upper curve; UB3LYP/6-31G* – lower curve)

potential energy curve. This is the only one-dimensional potential energy curve that I am aware of that has this feature.

The gradient is zero at both turning points – minimum and transition state – on the potential energy curve. The key difference between them is of course the sign of the energy second derivative (the rate of change of the gradient) which is positive for a minimum and negative for a transition state.

In this example, we have the complete energy curve between He–He distances of 0.5–4.0 Å available to us, but in general of course (in particular for anything more complicated than a diatomic molecule) this would not be the case, and if we wanted to search for, say, a local minimum we would have to provide an initial approximation for the geometry, i.e., a guess for the minimum structure. We would then calculate the energy, the gradient and either calculate, or provide a reliable estimate for, the second derivative and use this information to predict: (a) how far away we are on the PES from the nearest local minimum, and (b) the step to take (length and direction) in order to get there. Note that only in the one-dimensional case are the first and second derivatives single numbers; typically they would be vectors (with elements dE/dv_i , where v_i is the i -th variable) and matrices (with elements $d^2E/dv_i dv_j$), respectively.

Predicting the next step, i.e., how to change the current geometry to get to a new geometry that is closer to the desired stationary point than the old geometry was, is the job of the geometry optimization algorithm. For ab initio methods these algorithms can be very sophisticated as both the energy and the gradient are relatively expensive to compute (the second derivative is usually even more

expensive), and so additional manipulations in the optimization algorithm in order to reduce the number of optimization cycles (i.e., the number of costly energy and gradient calculations) is more than justified. Not so in molecular mechanics where the energy and gradient cost very little and a highly efficient – in terms of CPU time – optimization algorithm is required.

It is obvious that one of the factors that determines how quickly a given geometry optimization will converge is the starting geometry. In our He_2^{2+} example, in order to locate the metastable minimum, which occurs at a He–He distance of around 0.7 Å (shorter than in H_2 , making this the shortest bond distance known), any initial starting bond length greater than ca. 1.2 Å – the distance in the transition state on the lower, unrestricted, curve – will almost inevitably lead to dissociation (see Fig. 1). If you want to find the transition state then the closer you start to a He–He distance of around 1.2 Å the better off you will be.

The Newton–Raphson Step

Consider a starting He–He distance of about 1.0 Å, i.e., midway between the bound minimum and the transition state. This could probably go either way, toward the minimum or the transition state. Which way it does go is guided by the curvature of the surface around the local region, i.e., by the nature of the second derivative. If we do a standard Taylor series expansion about the current point, then we have

$$E(x+h) = E(x) + h \frac{dE}{dx} + \frac{h^2}{2} \frac{d^2E}{dx^2} + \frac{h^3}{6} \frac{d^3E}{dx^3} + \dots \quad (1)$$

In the region about a local minimum (or maximum) the potential energy curve is more-or-less parabolic, so assuming we are in or close to this region we can ignore all terms in the Taylor series expansion (Eq. 1) beyond the quadratic. Taking the derivative with respect to the displacement h gives

$$\frac{dE(x+h)}{dh} = \frac{dE}{dx} + h \frac{d^2E}{dx^2}. \quad (2)$$

At a stationary point $\frac{dE(x+h)}{dh} = 0$ and so the displacement h that achieves this is

$$h = -\frac{\frac{dE}{dx}}{\frac{d^2E}{dx^2}} \quad \text{or} \quad h = -\frac{g}{H}, \quad (3)$$

where g is of course the gradient and H is the energy second derivative.

Equation 3 is the well-known Newton–Raphson step in one dimension. The multidimensional equivalent of this is

$$\mathbf{h} = -\mathbf{H}^{-1} \mathbf{g}, \quad (4)$$

where \mathbf{h} is now a vector (a set of displacements for all variables), \mathbf{g} is the gradient vector, and \mathbf{H} is the second derivative matrix, commonly known as the Hessian matrix.

Going back to our He_2^{2+} example we see that the Newton–Raphson step is normally in the direction of minus the gradient (just like a steepest descent step) with a step length determined by the magnitude of the second derivative. If the second derivative is positive, the step will be downhill (opposite to the gradient) whereas if the second derivative is negative, the step will be uphill (in the same direction as the gradient). Thus at a He–He distance of 1.0 Å (see Fig. 1) the step taken will be uphill toward the transition state if the surface curvature is negative and downhill toward the minimum if it is positive.

In the multidimensional case, we can always find a unitary transformation that diagonalizes the Hessian matrix (which is symmetric about its diagonal), i.e.,

$$\mathbf{D} = \mathbf{U}\mathbf{H}\mathbf{U}^T, \quad (5)$$

where \mathbf{D} is a diagonal matrix and the superscript \mathbf{T} represents a matrix transpose. The vectors in \mathbf{U} are known as eigenvectors and the diagonal elements of \mathbf{D} are the corresponding eigen values. Both the displacement vector \mathbf{h} and the gradient \mathbf{g} can be transformed into this new (diagonal) representation as

$$\mathbf{g}' = \mathbf{U}^T\mathbf{g}; \quad \mathbf{h}' = \mathbf{U}^T\mathbf{h}. \quad (6)$$

Substituting Eq. 5 into Eq. 4 gives

$$\mathbf{h} = -(\mathbf{U}\mathbf{D}^{-1}\mathbf{U}^T)\mathbf{g}. \quad (7)$$

Multiplying both sides by \mathbf{U}^T and rearranging gives

$$\mathbf{U}^T\mathbf{h} = (\mathbf{U}^T\mathbf{U})\mathbf{D}^{-1}(\mathbf{U}^T\mathbf{g}), \quad (8)$$

$$\mathbf{h}' = \mathbf{D}^{-1}\mathbf{g}'. \quad (9)$$

Diagonalizing the Hessian matrix results in a set of mutually orthogonal directions on the PES (the vectors in \mathbf{U}) with the gradient along each direction given by the corresponding element in the vector \mathbf{g}' and the second derivative given by the corresponding diagonal matrix element in \mathbf{D} . Thus we see that the N -dimensional Newton–Raphson step is equivalent to N separate one-dimensional Newton–Raphson steps along the orthogonal directions given by the vectors in \mathbf{U} . If all the eigen values of the Hessian matrix (all the diagonal entries in \mathbf{D}) are positive, then the step taken will attempt to minimize along all directions on the PES (all degrees of freedom) and will head downhill to the nearest minimum. If one (and only one) of the Hessian eigen values is negative, then the Newton–Raphson step will attempt to maximize (head uphill in energy) along that direction, while simultaneously minimizing along all the other directions, i.e., it will move toward

a transition state. If the Hessian has two or more negative eigen values, then, left to its own devices, the optimization will head toward a second (or higher) order stationary point. As these are usually of much less interest than minima or transition states, the problem in these cases is normally how to take a decent optimization step when the Hessian signature (number of negative eigen values) is inappropriate for the stationary point being sought.

I have spent a lot of time on the Newton–Raphson step as it, or a modification thereof, forms the heart of virtually all advanced geometry optimization algorithms. It is usually a good step to take, particularly in the immediate region around the stationary point, and from a practical point of view it utilizes all the information that is likely to be available at the current point on the PES, i.e., first and second derivatives; there are very few codes that routinely compute derivatives higher than the second, and they would be prohibitively expensive in any case.

The Hessian Matrix and Hessian Updates

As I hope the above discussion has shown, the Hessian matrix plays a very important role in determining the step direction, which depends on the surface curvature in the local region of the PES. However, care must be taken with the step length. Hessian information is only local and only applies in the region immediately surrounding the current point. If the predicted step length is small, then you are probably near to a stationary point, in which case the step is best taken “as is”; however too large steps are usually not wise and you can easily get lost. Consequently it is a good idea to impose a limit on the step size – either an overall limit on the total step length or limits on the size of individual step components (or indeed both). Some algorithms attempt to define a “trust radius” which limits the step size dynamically depending on the local nature of the PES (usually determined by comparing the actual energy change from that predicted from the Hessian assuming quadratic behavior). In *PQS* a simple user-defined maximum step length is imposed (default 0.3 au).

Calculation of the exact Hessian for *ab initio* wave functions is expensive and in many cases it can be replaced by an approximation with little effect on the optimization itself, i.e., on the number of optimization cycles required to reach convergence. In some cases having the exact Hessian is a definite disadvantage. If you are fairly close on the PES to the stationary point being sought then, although an exact Hessian will likely get you there faster, the reduction in the number of optimization cycles will probably not be enough to offset the extra computational time required to compute the Hessian in the first place. (The exception is if the surface is fairly flat – long, weak bonds, or, especially, floppy systems with flat torsional potentials – in which case an exact Hessian will be a tremendous help.) A good quality Hessian will usually help the most when you are not really close to a stationary point, but within the region where the surface curvature is still appropriate, i.e., the Hessian matrix still has the correct signature (all positive eigen values for a minimum and one, and only one, negative eigen value for a transition state). Computing an exact Hessian will almost certainly be a waste of time if you

are a long way from where you finally end up; under these circumstances an exact Hessian will almost inevitably have more negative eigen values than required and the Newton–Raphson step alone will clearly be inappropriate. If you are searching for a minimum you will be far better off using an alternative for the Hessian matrix which has all positive eigen values (such a matrix is known as positive-definite) which should keep the energy going down until you get closer to the stationary point. If the gradient is large, then a few steepest descent steps, possibly with a line search, might be in order, simply to lower the energy and get to a more appropriate region of the PES.

In fact, other than in exceptional circumstances, when searching for a minimum you do not normally need to compute a full Hessian matrix at all. In most cases reasonably good starting geometries can be derived simply from “chemical intuition,” for example, C–H bond lengths (except in the rare cases of bridging hydrogens) are likely to be of the order of 1 Å, so there is no point in using a starting geometry with C–H distances differing much from this value. Very reasonable estimates of stretching, bending and also torsional force constants (diagonal Hessian elements) are readily available – for example, stretching force constants via Badgers rule (Badger 1934, 1935) (see also Schlegel 1984) – which are perfectly adequate in the vast majority of cases. Consequently you can normally come up with an approximate estimate for at least the diagonal elements of the Hessian matrix at the starting geometry using some simple empirical rules.

Once you have taken a step from one point on the PES to another (hopefully closer to the final stationary point), then you can compute an energy and a gradient at the new geometry. This information can be utilized, together with the energy, gradient, and anything else that might be to hand from the old point (or points), not only to work out a new step (the job of the geometry optimization algorithm) but also to improve (update) the original approximation for the Hessian matrix. You can think of this as similar to determining the surface curvature by finite difference on the gradient; you know the gradient where you started, you take a small step, calculate a new gradient, and determine the curvature by forward differences. Of course, updating a Hessian matrix on a multidimensional energy surface is a lot more approximate than this as typically the step taken is fairly large and you are attempting to get an update for the whole Hessian from a step in just one direction.

There are a number of different Hessian updating formulae, e.g., Murtagh and Sargent (1970) and Powell (1971), Broyden–Fletcher–Goldfarb–Shannon (BFGS) (Broyden 1970; Fletcher 1970, 1980; Goldfarb 1970; Shanno 1970), which can be used where appropriate depending on the nature of the stationary point being sought. (The exact details are not important.) For minimization the BFGS update is a popular one as, if you start out with a positive-definite Hessian matrix (with all positive eigen values), as needed for a minimum search, you are more likely to retain this desirable feature with BFGS than with the other updates. (You may see statements to the effect that the BFGS update *guarantees* retention of a positive-definite Hessian. This is not true. However, on those occasions – not too common, but they do occur – when the update fails in this regard, you can tell in advance that it is going to fail and either skip the update on that step or take a different one.)

Transition State Searches

For a minimum search then, computing an exact Hessian matrix is rarely required. Things are unfortunately not so straightforward for transition state searches. First of all transition state geometries are not as obvious as are the geometries of stable minima and so it is often quite difficult to come up with good starting structures. This can be a major problem in itself. Secondly, the correct Hessian eigen value structure is much more important for a transition state than it is for a minimum. For a transition state search, the Hessian must have one, and only one, negative eigen value, and furthermore the direction of negative curvature must correspond to that for the structure being sought; to the reaction coordinate if you will. Difficult transition state searches can be plagued with additional small negative Hessian eigen values continually reappearing which correspond to rotation of, e.g., methyl groups, way off the reaction path, i.e., on some side chain or other nowhere near the reaction center. These additional negative eigen values often interfere with the search algorithm making the final structure difficult to locate. And thirdly, there are no simple empirical rules generally available for estimating a reliable starting Hessian for transition state searches as there are for minima. All these factors make transition state searches more difficult than minimizations and in all but the simplest cases a good estimate of the Hessian for at least the region round the active center is essentially mandatory.

As is the case with minimization, once you are “on the right track,” it is not necessary to calculate fresh second derivative data at every optimization cycle, the Hessian matrix can be updated. Currently the best update for a transition state search is probably that proposed by Bofill (1994) which is a linear combination of the Powell (1971) and Murtagh and Sargent (1970) updates. This is a more flexible update than BFGS and allows the eigen value structure of the Hessian to change, which is obviously important if you get too close to a minimum. Unfortunately, unlike the case with minimization, there is no update that preserves the negative eigen value once you have it, and it is all too easy to find that the updated Hessian has either too few (zero) or too many (two or more) negative eigen values.

There are transition state search algorithms that do not require an initial starting structure; instead they take as input the two minima that you wish to connect and attempt to derive some kind of “reaction path” connecting the two. Depending on the nature of the algorithm, the maximum energy structure on this path is either the transition state itself, or a pretty good estimate for it which is subsequently refined. The original algorithm of this type was the Linear Synchronous Transit (LST) approach of Halgren and Lipscomb (1977) which mapped each atom in the reactant to a corresponding atom in the product and derived the best linear path between them in a least-squares sense. The maximum energy structure along this path provided an estimate (unfortunately often not a particularly good one) of the transition state geometry which was subsequently used as input into a standard search algorithm.

An improvement over this was Quadratic Synchronous Transit (QST) (Bell and Crighton 1984; Bell et al. 1981), which attempted to maximize along the line joining reactant and product while minimizing in all directions “orthogonal” (or conjugate) to this. One can easily see the rationale behind this approach: if the direction to be maximized could be reliably isolated from all other directions (imagine a Hessian matrix with what will become the direction of negative curvature having all zero off-diagonal matrix elements, i.e., H_{ij} ; all $H_{ij} = 0$) then it would be straightforward to maximize along the one direction while minimizing in the subspace of all the other variables. Unfortunately maintaining the orthogonality between the isolated direction and the rest of the space (i.e., keeping $H_{ij} = 0$, $j \neq i$) is rather difficult, and so the algorithm can often get lost. Nonetheless, there are several more modern algorithms based on these approaches, for example from Peng and Schlegel (1993) and Ionova and Carter (1993).

If the “reaction coordinate” can be reliably confined to just a couple of (ideally one) variables, then a very good starting geometry for a transition state search can be obtained from an optimized potential scan. Here the variable in question is scanned over a range of values (hopefully including the value the variable has in the transition state) and at each scanned value all other degrees of freedom are minimized. In effect this involves a series of constrained optimizations over the scanned variable. Once again, the transition state is taken to be the geometry at the highest energy found during the scan. (If the highest energy occurs at either of the scan endpoints then either the scan range needs to be extended or there is no transition state in the region of the PES that is being explored.) An optimized potential scan is easy to set up with *PQS* and is normally a very reliable method, albeit potentially time consuming, of obtaining an excellent transition state starting geometry provided that the principal reaction coordinate can be successfully identified.

Choice of Coordinates

In the discussion so far the choice of coordinates in which to carry out minimizations or transition state searches has not been specifically mentioned, although in certain cases – for example, when considering a potential scan – it is implicitly assumed that variables familiar to chemists, such as bond lengths and angles, have been employed. Virtually all quantities of interest that are actually computed in *ab initio* quantum chemistry, such as first (gradient) and second (Hessian) derivatives are calculated in terms of Cartesian coordinates and the molecular geometry, regardless of how it may be input, is stored and manipulated internally as a set of Cartesians. However, Cartesian coordinates are usually not the best choice when it comes to actually carrying out a geometry optimization, because they are generally too coupled. The only case where they are regularly used in the actual geometry optimization itself is in molecular mechanics.

In the late 1960s and early 1970s the so-called *Z*-matrix was widely used in quantum chemistry, initially simply as a means of inputting the molecular geometry,

but later as a means of defining a set of internal coordinates in which to carry out the optimization. With a *Z*-matrix, the geometry can be defined in terms familiar to a chemist using individual (primitive) bond lengths, bond angles, and torsions. The first atom defined in the *Z*-matrix is placed at the origin of a standard Cartesian axis system (i.e., at 0.0, 0.0, 0.0), the second is placed along the *Z* axis at a distance *R*₁, say, and the third in the *XZ* plane at a distance, *R*₂, from the second atom and making an angle, *A*₁, with the first atom (or vice versa). Thereafter the position of all subsequent atoms is given in terms of a bond distance, a bond angle and, say, a torsion, relative to three previously defined atoms. In the days before graphical user interfaces (GUI) and model builders, the *Z*-matrix became a regular way of defining and reading a molecular geometry into the program doing the actual calculation. Dummy atoms could be included in the *Z*-matrix (and were needed to help define linear arrangements of three or more atoms for which the torsion was undefined) and, once geometry optimizations were regularly performed in *Z*-matrix coordinates, the writing of a successful *Z*-matrix became almost an art form, with the aim of defining the *Z*-matrix variables so that the value of any one of them could be changed without changing the values of any of the others, the idea being to reduce the coupling between the variables to a minimum.

A well-defined set of internal coordinates usually performs better in a geometry optimization, i.e., converges in less optimization cycles, than the same optimization carried out in Cartesian coordinates. The playing field can be leveled, particularly for standard organic molecules, if the geometry is preoptimized in advance using a molecular mechanics force field (ensuring a reliable starting geometry) and if second derivative information – computed very cheaply during the mechanics minimization – is transferred to the *ab initio* optimization (Baker and Hehre 1991). However, as the system to be optimized gets larger and/or more complicated, Cartesian coordinates become less and less competitive with respect to a good set of internal coordinates.

The *Z*-matrix is rarely used today. With the increasing popularity of GUIs and model builders it is no longer needed as a means of geometry input, and *Z*-matrix coordinates have been superseded as a coordinate set for optimization, first by natural internal coordinates (Fogarasi et al. 1992; Pulay et al. 1979) and then by delocalized internal coordinates (Baker et al. 1996). The latter, which are linear combinations of stretches, bends, and torsions (as opposed to the individual primitive internals in a *Z*-matrix), are readily generated from an underlying set of Cartesian coordinates and decouple the coordinate set (in a linearized sense) to the maximum extent possible (Baker et al. 1996). In large systems, reductions in the number of optimization cycles to reach convergence by an order of magnitude or more compared to Cartesians can be readily achieved. There is an additional computational cost in using delocalized internal coordinates however, as the transformations required to convert the Cartesian gradient into the gradient over delocalized internals, and – in particular – to convert the new geometry in internals back into Cartesians for the next optimization step (the back-transformation), are expensive and increasingly so with increasing system size. They scale formally as $O(N^3)$, where *N* is a rough measure of the system size, although this can readily

be reduced to $O(N^2)$ (Baker et al. 1999). There have been attempts to reduce the scaling to $O(N)$ (Farkas and Schlegel 1998; Paizs et al. 2000), but I think it is fair to say that these have not been entirely successful. It is principally because of the unfavorable scaling in the coordinate transformations that internal coordinates have not been adopted in molecular mechanics.

The cost of the coordinate transformations and the overall complexity of the algorithm is not really a factor in ab initio methods and the majority of geometry optimizations in modern quantum chemistry packages are carried out in delocalized internals (Baker et al. 1996) or in some other set of redundant internal coordinates (Peng et al. 1996; Pulay and Fogarasi 1992).

The Modified Newton–Raphson Step

We have already seen that the Newton–Raphson step is a pretty good one provided that the eigen value structure of the Hessian matrix is appropriate to the stationary point being sought. What should you do if it is not? An early effort to take remedial action under these circumstances was due to Poppinger (1975) who proposed that, when searching for a transition state, if the negative eigen value was lost then take a step uphill in the direction of the lowest positive eigenvector, while if an additional (unwanted) negative eigen value appeared, then take a step downhill in the direction of the least negative eigenvector. While this approach should lead you back into the right region of the PES, it has at least two drawbacks. Firstly, unlike the Newton–Raphson step which, when appropriate, minimizes/maximizes along all directions on the PES (all Hessian modes) simultaneously, here you are minimizing/maximizing essentially along one direction only. Secondly, successive such steps following a particular Hessian mode tend to become linearly dependent, which degrades many of the standard Hessian updating procedures.

A much better step was subsequently proposed by Cerjan and Miller (1981) which involved modifying the standard Newton–Raphson step with a so-called shift parameter, λ . In terms of a diagonal Hessian representation, the Newton–Raphson step given by Eq. 4 can be written as

$$\mathbf{h} = \sum_i -\frac{F_i \mathbf{U}_i}{d_i}, \quad (10)$$

where \mathbf{U}_i is the i -th eigenvector, d_i is the corresponding eigen value, and $F_i = \mathbf{U}_i^T \mathbf{g}$ is the component of the gradient along the local eigen mode \mathbf{U}_i .

The modified Newton–Raphson step is given by

$$\mathbf{h} = \sum_i -\frac{F_i \mathbf{U}_i}{d_i - \lambda}, \quad (11)$$

where λ can be regarded as a shift parameter on the Hessian eigen values. A suitable choice of λ can guide the step in a particular direction; for example if λ is close to a

particular eigen value d_i , then the step taken will be predominantly in the direction of the corresponding eigenvector. In the early to mid-1980s, several groups were working on this topic, and several alternative transition state search algorithms were proposed differing only in the recipe used to determine the value of λ (Banerjee et al. 1985; Cerjan and Miller 1981; Simons et al. 1983).

The principal optimization algorithm in *PQS*, both for minima and transition states, is the eigenvector following (EF) algorithm developed in 1986 (Baker 1986). It is based on the Rational Functional Optimization (RFO) approach of Simons and coworkers (Banerjee et al. 1985) which was found at the time to provide the best recipe for determining the shift parameter, λ . In this approach, the energy on the PES following a step \mathbf{h} is written as (see Eq. 1):

$$E(x+h) = E(x) + \frac{\mathbf{g}^T \mathbf{h} + \frac{1}{2} \mathbf{h}^T \mathbf{H} \mathbf{h}}{1 + \mathbf{h}^T \mathbf{S} \mathbf{h}}. \quad (12)$$

In the original formalism, the matrix \mathbf{S} was a diagonal scaling matrix but in most practical implementations (certainly in the EF algorithm) it is taken as a unit matrix.

Applying the stationary condition $dE/dh = 0$ to Eq. 12 gives the RFO eigen value equation (Banerjee et al. 1985):

$$\begin{pmatrix} \mathbf{H} & \mathbf{g} \\ \mathbf{g}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{h} \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{h} \\ 1 \end{pmatrix}. \quad (13)$$

The dimensionality of Eq. 13 is one more than the number of variables (n). As it is an eigen value equation, there are $(n+1)$ eigen values, λ ; normally the lowest one, λ_1 , is the one you need. Expanding out Eq. 13 gives

$$(\mathbf{H} - \lambda \mathbf{S}) \mathbf{h} + \mathbf{g} = \mathbf{0}, \quad (14)$$

$$\mathbf{g}^T \mathbf{h} = \lambda. \quad (15)$$

In terms of a diagonal Hessian representation, and with \mathbf{S} a unit matrix, Eq. 14 rearranges to Eq. 11, and substitution of Eq. 4 into the diagonal form of Eq. 15 gives

$$\Sigma_i - \frac{F_i^2}{\lambda - b_i} = \lambda, \quad (16)$$

which can be used to iteratively evaluate λ .

The eigen values, λ , of Eq. 13 have the following properties (Banerjee et al. 1985):

1. The $(n+1)$ eigen values of Eq. 13 bracket the n eigen values of the Hessian:
 $\lambda_i \leq d_i \leq \lambda_{i+1}$.

- At convergence to a minimum both \mathbf{g} and \mathbf{h} are $\mathbf{0}$. The lowest eigen value, λ_1 , of Eq. 13 is also zero and the other n eigen values are those of the Hessian at the minimum point, i.e., $\lambda_{i+1} = d_i$ for all i . Thus, near a local minimum, λ_1 is negative and approaches zero at convergence.
- For a saddle-point of order μ (i.e., a stationary point with μ negative eigen values) the zero eigen value of Eq. 13 at convergence separates the μ negative and $(n - \mu)$ positive Hessian eigen values.

The separability of the positive and negative Hessian eigen values led Simons and coworkers to suggest that the problem be reformulated into μ principal modes to be maximized and $(n - \mu)$ modes to be minimized. There would be two shift parameters, λ_p and λ_n , one for modes along which the energy is to be maximized and the other for which it is minimized. The RFO eigen value equation (Eq. 11) could thus be partitioned into two smaller P-RFO equations, and each one solved separately.

For a transition state, in terms of a diagonal Hessian representation, this gives rise to the two matrix equations:

$$\begin{pmatrix} d_1 & F_1 \\ F_1 & 0 \end{pmatrix} \begin{pmatrix} h_1 \\ 1 \end{pmatrix} = \lambda_p \begin{pmatrix} h_1 \\ 1 \end{pmatrix}, \quad (17)$$

$$\begin{pmatrix} d_2 & 0 & & F_1 \\ 0 & \ddots & & \vdots \\ & & d_n & F_n \\ F_2 & \dots & F_n & 0 \end{pmatrix} \begin{pmatrix} h_2 \\ \vdots \\ h_n \\ 1 \end{pmatrix} = \lambda_n \begin{pmatrix} h_2 \\ \vdots \\ h_n \\ 1 \end{pmatrix}. \quad (18)$$

λ_p is the highest eigen value of Eq. 17; it is always positive and approaches zero at convergence. λ_n is the lowest eigen value of Eq. 18; it is always negative and again approaches zero at convergence.

The above discussion assumes that, for a transition state search, the energy is to be maximized along the lowest Hessian mode, d_1 , and minimized along all the higher modes. However, as first pointed out by Cerjan and Miller (1981), it is possible to maximize along modes other than the lowest and in this way obtain transition states for alternative reactions and rearrangements from the same starting point. The ability of the algorithm to “walk” on the PES, maximizing along different Hessian eigen modes, is why the algorithm is known as Eigenvector Following. For maximization along the k -th mode, instead of the lowest, Eq. 17 would be replaced by

$$\begin{pmatrix} d_k & F_k \\ F_k & 0 \end{pmatrix} \begin{pmatrix} h_k \\ 1 \end{pmatrix} = \lambda_p \begin{pmatrix} h_k \\ 1 \end{pmatrix}, \quad (19)$$

and Eq. 18 would now exclude the k -th mode but include the lowest. Since what was originally the k -th mode is the mode along which the negative eigen value is required, if the search is to be successful this mode must eventually become the lowest mode. The mode that is actually followed during each optimization cycle is the mode which has the largest overlap with the mode followed on the previous cycle. In this way the same mode should hopefully be followed from cycle to cycle as it makes its way through the ranks from k -th position to first (lowest). Once the mode being followed has become the lowest, mode-following is switched off and it is always the lowest mode that is followed.

Although it was developed primarily for transition states, the EF algorithm is also a very powerful minimizer. For minimization the standard RFO-step is taken (requiring a single shift parameter) using Eq. 13. Mode following for transition state searches is useful if you want to explore the PES in detail, but it is likely to be expensive and requires reliable Hessian information to be effective. There will almost certainly be more variables, and hence more Hessian modes, for a given system than there are different transition structures associated with that system; furthermore it is not always obvious which mode to follow at any given point in order to locate a transition state of a desired type. Beware of low energy modes which if followed will lead to rotational barriers for torsions in side chains rather than to the transition state for the main reaction being sought.

GDIIIS

As well as the EF algorithm, the *PQS* package has another algorithm for structure minimization: GDIIIS (Csaszar and Pulay 1984). This is an extension of Pulay's ubiquitous DIIS (Direct Inversion in the Iterative Subspace) procedure for accelerating SCF convergence (Pulay 1980, 1982), only applied instead to geometry optimization.

The essence of the DIIS approach is that parameter vectors (in this case the geometry) generated in previous iterations $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ are linearly combined to find the best geometry in the current iteration. We can express each parameter vector in terms of its deviation from the final solution, \mathbf{x}_f : $\mathbf{x}_i = \mathbf{x}_f + \mathbf{e}_i$. If the conditions $\sum_i c_i \mathbf{e}_i = 0$ and $\sum_i c_i = 1$ are satisfied, then it is also the case that $\sum_i c_i \mathbf{x}_i = \mathbf{x}_f$. The true error vectors, \mathbf{e}_i , are of course unknown. However, they can be approximated by $\mathbf{e}_i = -\mathbf{H}^{-1} \mathbf{g}_i$ (i.e., by the Newton–Raphson step). Minimization of the norm of the residuum vector $\sum_i c_i \mathbf{e}_i$ taking into account the constraint equation $\sum_i c_i = 1$ leads to a system of m equations (Csaszar and Pulay 1984):

$$\begin{pmatrix} B_{11} & B_{12} & \dots & B_{1m} & 1 \\ B_{21} & B_{22} & \dots & B_{2m} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ B_{m1} & B_{m2} & \dots & B_{mm} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \\ -\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \quad (20)$$

Where $B_{ij} = \langle \mathbf{e}_i | \mathbf{e}_j \rangle$, the scalar product of the error vectors \mathbf{e}_i and \mathbf{e}_j , and λ is a Lagrange multiplier.

The coefficients, c_i , obtained from Eq. 20 are used to calculate an intermediate interpolated geometry:

$$\mathbf{x}'_{m+1} = \sum_i c_i \mathbf{x}_i, \quad (21)$$

and a corresponding interpolated gradient vector

$$\mathbf{g}'_{m+1} = \sum_i c_i \mathbf{g}_i. \quad (22)$$

A new, independent geometry is generated from the interpolated geometry by acting on the interpolated gradient with the (approximate) Hessian,

$$\mathbf{x}_{m+1} = \mathbf{x}'_{m+1} - \mathbf{H}^{-1} \mathbf{g}'_{m+1}. \quad (23)$$

Equations 20, 21, 22, and 23 constitute a complete cycle of the GDIIS method. Convergence can be tested for in the usual way (based on the gradient and the displacement from the previous step) or on the norm of the GDIIS residuum vector.

The size of the GDIIS subspace, m , increases by one after each optimization cycle. Experience with the method suggests that m should not be allowed to grow too large; once a predefined maximum size has been reached (say 20, the current default in *PQS*), the earliest vectors are dropped out as new vectors are added. Intermediate quantities (geometries and error vectors) within the GDIIS subspace need to be stored, so limiting the subspace size also limits the storage requirements of the method. (This is not really an issue with GDIIS, but is more important with DIIS, where potentially large-dimension Fock and possibly density matrices are stored.)

GDIIS was originally used in what is termed “static” mode, in which the original estimate for the Hessian matrix was used throughout the optimization procedure. In fact GDIIS performs very well with a simple unit matrix approximation for the Hessian. However, it certainly performs better if the Hessian matrix is updated. It is particularly important that the Hessian approximation remain positive definite, and so the BFGS update is strongly recommended with the update skipped if it is determined that this property is in danger of being lost.

Geometry Optimization and Symmetry

Many molecules have symmetry and it is definitely advantageous to utilize as much symmetry as possible during the calculation of the energy, the gradient and, if it is being computed, the Hessian. Various quantities, particularly the geometry and the gradient, can be symmetrized during the optimization step itself, thus removing numerical round-off error and resulting in a cleaner, and occasionally faster, optimization. Most programs can utilize at least Abelian point-group symmetry,

resulting in savings in computational time by up to a factor of 8, depending on the molecular symmetry.

If symmetry is utilized during a geometry optimization it obviously restricts the search on the PES to structures that retain the molecular symmetry, i.e., it excludes symmetry-breaking geometries. This can be particularly useful during a transition state search as, if it is known that the motion in the transition state breaks symmetry, then the transition state search can be replaced by a minimization within a symmetry-restricted subspace, usually a much simpler prospect. For example, consider ethane. The minimum, lowest energy structure has staggered C–H bonds and D_{3d} symmetry. The transition state for rotation about the C–C bond has eclipsed C–H bonds and D_{3h} symmetry. The latter can in fact be found by a simple minimization utilizing D_{3h} point-group symmetry as the path to the D_{3d} minimum is unavailable under the symmetry constraint. The minimum under D_{3h} symmetry is actually a transition state once the symmetry constraint is removed.

This illustrates another pitfall that occasionally catches out the novice: A structure that has been minimized under a symmetry constraint is not necessarily a true minimum and its identity should be confirmed by a full vibrational analysis. As a corollary to this, a transition state located under a symmetry constraint may not be a true transition state either; it could have one (or more) symmetry breaking modes with negative curvature.

Performance for Minimization

We have had a lot of theory. How well, for example, does the standard eigenvector following (EF) algorithm perform in practice? This is illustrated in Table 1 for a selection of ten organic molecules, selected at random from the fragments library in the graphical user interface *PQSMol* available for use with PQS (2010). These systems were all optimized using a molecular mechanics force field (it is not important which one) and symmetrized prior to being stored.

There have of course been a number of previous comparisons of the performance of various optimization algorithms. I mention two in particular, as I was involved in them myself. In 1991 Baker and Hehre compared the performance of the EF and GDIIS minimization algorithms, as implemented in the version of *Spartan* available at that time, on a test set of 20 organic molecules with a variety of structural motifs (Baker and Hehre 1991). The two algorithms performed similarly, but the main focus of this paper was to demonstrate that, provided reliable second derivative information was available, i.e., a good starting Hessian, then Cartesian coordinates were just as efficient as Z-matrix coordinates for routine geometry optimization. A later paper Baker (1993) did the same thing for Cartesians versus natural internal coordinates on a larger test set of 30 mainly organic molecules. This latter test set became a kind of “standard” for checking improved optimization algorithms and several papers showing minor enhancements in performance (as would be expected with the passage of time) for these 30 molecules were published subsequently (Bakken and Helgaker 2002; Eckert et al. 1997; Lindh et al. 1995; Swart and

Table 1 Number of optimization cycles required to reach convergence for ten average-sized organic molecules under a range of starting conditions^a

Molecule	Formula	Atoms	Symm.	Cartesians		Delocalized internals		
				Unit ^b	PM3 ^c	Unit ^b	Default ^d	PM3 ^e
Alanine	C ₃ H ₇ NO ₂	13	C ₁	28	11	20	10	8
Indole	C ₈ H ₇ N	16	C _s	22	5	11	5	5
Cycloheptane	C ₇ H ₁₄	21	C ₂	27	6	23	5	6
Bicyclo-2,2,2-octane	C ₈ H ₁₄	22	D ₃	12 ^e	10	29	8	8
Adamantane	C ₁₀ H ₁₆	26	T _d	12	3	11	4	4
<i>cis</i> -decalin	C ₁₀ H ₁₈	28	C ₂	27	4	26	4	4
amp-5	C ₁₀ H ₁₄ N ₅ O ₇ P	37	C ₁	193 ^e	68	85	29	35
Porphine	C ₂₀ H ₁₄ N ₄	38	D _{2h}	21	7	11	6	7
Sucrose	C ₁₂ H ₂₂ O ₁₁	45	C ₁	163	31	67	17	21
Cyclohexadecane	C ₁₆ H ₃₂	48	S ₄	32	10	27	6	9

^aStandard EF algorithm used with a maximum allowed step size in the optimization space of 0.3 au and a BFGS Hessian update; convergence criteria were the standard *PQS* defaults – maximum component of the gradient vector less than 0.0003 au and either a maximum predicted displacement of 0.0003 au or an energy change from the previous cycle of less than 10⁻⁶ Eh

^bStarting Hessian was a unit matrix

^cStarting Hessian was computed using PM3 by central-differences on the Cartesian gradient

^dDiagonal elements of the starting Hessian were estimated using simple empirical rules

^eConverged to higher energy local minimum

Bickelhaupt 2006). A similar test set is also available for transition structures (Baker and Chan 1996).

The conclusion in Baker (1993) was that with a reasonable starting geometry and a good initial Hessian, optimizations in Cartesian coordinates were as efficient as those employing natural internal coordinates. However, there were certainly signs that, as the system size increased, the use of natural internals became more and more advantageous even if a reliable starting Hessian was available, and it was further concluded that with no initial Hessian information, natural internals (for which read also delocalized internals which have more-or-less replaced them) were in general superior to both Cartesian and Z-matrix coordinates, and for unconstrained optimization were therefore the coordinates of first choice.

One of the issues with the published studies (Baker 1993; Baker and Hehre 1991) is that the size of the systems in both test suites (all molecules had less than 30 atoms, and most much less), although typical for the time (the early 1990s), is pretty small by today's standards. The molecules in Table 1 are about 50 % larger on the average, four of them have more than 30 atoms, with the largest having 48. The level of theory used, B3LYP/6-31G*, is also more typical of today. They are still only average-sized molecules, but nonetheless are representative as far as performance is concerned.

All the conclusions reached in the previous studies are fully supported by the results shown in Table 1. The use of a decent Hessian, even just an approximate one at the starting geometry (in this case computed by central differences using

the semiempirical PM3 method (Stewart 1989)), improves the performance of minimizations in Cartesian coordinates substantially, by up to a factor of almost 7 (cis-decalin) and by an average over all 10 molecules of around 3.5. Despite this, Cartesian coordinates are not really competitive with delocalized internals under the same conditions, taking about 45 % more cycles on the average to reach convergence. The performance of Cartesians is notably worse for large, floppy molecules with no symmetry (amp-5 and sucrose). Perhaps surprisingly, the performance of delocalized internals using the standard default Hessian (estimating diagonal matrix elements only using simple empirical rules) is even better, with the number of optimization cycles being the same or less than with the PM3 starting Hessian for all systems except one (alanine).

Of some surprise, at least to me, is the relatively poor performance of internal coordinates when no Hessian information (a unit matrix) is used. They are certainly better than Cartesian coordinates under the same conditions, which require about 75 % more cycles on the average to converge, but perform over three times worse than when the default diagonal Hessian is used, showing the advantages of even rudimentary information about the surface curvature.

One of the molecules in Table 1, bicyclo-2,2,2-octane, took significantly less optimization cycles to converge in Cartesian coordinates with a unit Hessian (12) than with the same optimization in delocalized internal coordinates (29); the reverse is the case for all of the other systems. This apparent anomaly is due to the fact that the Cartesian unit Hessian optimization converged to a higher energy local minimum than the other optimizations. This was not the only occasion when this occurred: higher energy local minima were also found for amp-5 and to a lesser extent for sucrose and alanine (although in the latter case the difference was only 40 μ H and may well be simply due to premature convergence, a fairly common occurrence for unit Hessian Cartesian optimizations (Baker et al. 1996)).

These results support another assertion made in Baker (1993), namely, that Cartesian optimizations tend to converge to the nearest local minimum whereas optimizations in natural (or delocalized) internals tend to converge more globally. This global tendency, and that is really all it is, with the internal coordinate optimizations is principally due to the fact that relatively small displacements in internal coordinates can lead to large changes when the geometry is converted back to Cartesian coordinates. For example a small change in a central bond angle can lead to a large Cartesian displacement at the end of a lengthy attached side chain; such large displacements enable internal coordinate optimizations to “jump over” local minima. (This phenomenon was first observed in the days of the Z-matrix.)

This is an appropriate point to emphasize that all the optimization algorithms discussed so far in this chapter are local methods, aimed at locating a single stationary point. Whether this is the lowest energy structure on the entire PES, i.e., the global minimum, is quite another matter. The number of (mainly conformational) minima a molecule has increases dramatically with system size and the problem of locating the one with the lowest energy quite quickly becomes almost insurmountable. Additionally, of course, unless the global minimum is noticeably lower in energy than other low-lying local minima, there will be a Boltzmann distribution at room

temperature with several different minima having appreciable population; under these circumstances to talk about a well-defined structure does not make a lot of sense.

Problems in locating global minima have only recently become an issue in *ab initio* calculations as the size of the systems that can be handled using these techniques has increased significantly in recent years. A number of global search algorithms are available, although none of them are bulletproof (i.e., absolutely guaranteed to locate the global minimum in a finite time). The problem of global minimization is outside the scope of this chapter and the interested reader is referred to the existing literature for more detail (Pardalos et al. 1995).

Minimization: An Example

Figure 2 shows the optimization output at each cycle in the standard default EF minimization of indole. There is more printout than normal for the first and last cycles as I have used a higher print flag in order to show more detail. The output is of course specific to *PQS* but would be similar – at least in terms of the quantities printed – in other programs.

At the beginning of each optimization cycle the current geometry is printed in Cartesian coordinates together with the point group (C_s) and the number of degrees of freedom (29). Indole is planar and is oriented in the *XY* plane; consequently all *Z*-coordinates are zero. Next comes the SCF energy at that geometry and the Cartesian forces (the gradient – not normally printed). The latter should fully reflect the symmetry of the system, which it does in this case, and so all *Z*-gradient components are zero.

This is followed by a symmetry analysis of the set of delocalized internal coordinates generated for indole which results in the elimination of all primitive torsions, all of which are fixed by symmetry. This analysis is typically done once only as thereafter the delocalized internal coordinates formed and used on the first cycle are retained unaltered throughout the optimization. One does not have to do this; several algorithms (e.g., Schlegel and coworkers (Peng et al. 1996)) effectively regenerate a set of active delocalized internals on each optimization cycle although in most cases this is not necessary and does not improve the overall performance.

There is then a statement as to how many Hessian modes (eigenvectors) are used during the construction of the optimization step – this does not have to be all of them, but is usually most, and cannot be more than the total number of degrees of freedom – followed by a listing of the corresponding Hessian eigenvalues (the curvature along each of the eigenvector directions). In this particular optimization, a diagonal Hessian in the underlying space of primitive internals (individual stretches, bends and torsions) is estimated using empirical rules, and is then transformed into the space of selected delocalized internal coordinates, \mathbf{U} , via $\mathbf{H} = \mathbf{U}^T \mathbf{H}_{\text{prim}} \mathbf{U}$ (Baker et al. 1996). The simple RFO step is then computed and taken, appropriate for minimization, resulting in a single shift parameter, λ (see

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **
Searching for a Minimum

Optimization Cycle: 1

		Coordinates (Angstroms)		
ATOM		X	Y	Z
1	c	-1.128517	1.256320	0.000000
2	c	-2.175159	0.350932	0.000000
3	c	-1.944638	-1.056645	0.000000
4	c	-0.661806	-1.575795	0.000000
5	c	0.418576	-0.654407	0.000000
6	c	0.193566	0.736550	0.000000
7	n	1.801059	-0.848856	0.000000
8	c	1.464739	1.385216	0.000000
9	c	2.431380	0.400066	0.000000
10	h	-1.307547	2.344176	0.000000
11	h	-3.212371	0.726200	0.000000
12	h	-2.806900	-1.744633	0.000000
13	h	-0.476805	-2.662311	0.000000
14	h	2.242590	-1.695842	0.000000
15	h	1.641099	2.472979	0.000000
16	h	3.520733	0.566050	0.000000

Point Group: Cs Number of degrees of freedom: 29

Energy is -363.810295355

		Cartesian Forces (au)		
ATOM		X	Y	Z
1	c	0.0055736	-0.0029170	0.0000000
2	c	-0.0116662	-0.0021539	0.0000000
3	c	-0.0102295	0.0028293	0.0000000
4	c	0.0158848	-0.0032863	0.0000000
5	c	-0.0012122	-0.0121078	0.0000000
6	c	-0.0098733	0.0169732	0.0000000
7	n	-0.0277428	0.0598278	0.0000000
8	c	0.0127970	0.0051731	0.0000000
9	c	-0.0032541	0.0095309	0.0000000
10	h	0.0020479	-0.0106104	0.0000000
11	h	0.0092568	-0.0048346	0.0000000
12	h	0.0083209	0.0070569	0.0000000
13	h	-0.0025071	0.0095819	0.0000000
14	h	0.0262691	-0.0500248	0.0000000
15	h	-0.0010053	-0.0136614	0.0000000
16	h	-0.0126474	-0.0113918	0.0000000

Fig. 2 (continued)

Iterative generation of Internal Gradient
 Eliminated 13 Coordinates due to Symmetry
 Setting Up Default Hessian

Eliminating Redundant Primitive Internals from Space
 Removed 40 Torsions
 There are now 44 Primitive Internals

29 Hessian modes will be used to form the next step

Hessian Eigenvalues:

0.160000	0.160000	0.160000	0.160000	0.160000	0.160000
0.160000	0.220000	0.221923	0.226566	0.240305	0.246612
0.333419	0.333513	0.334083	0.334451	0.334655	0.334700
0.371362	0.386468	0.401879	0.411214	0.422264	0.441836
0.450459	0.468353	0.478270	0.480256	0.565271	

Minimum Search -- Taking Simple RFO Step
 Searching for Lambda that Minimizes Along All modes
 Value Taken Lambda = -0.01172058
 Step Taken. Stepsize is 0.174879

Parameter Values and Displacements in Internal Coordinates

Coordinate	Current Value	Gradient	Displacement	New Value
1	0.101253	-0.014349	0.053801	0.155053
2	5.456291	0.008615	-0.016785	5.439506
3	0.217497	-0.007632	0.010903	0.228400
4	-3.211789	-0.009303	0.024794	-3.186995
5	-0.012152	-0.000858	0.006348	-0.005803
6	1.909226	0.003986	-0.024340	1.884886
7	-3.105834	-0.009981	0.008840	-3.096993
8	2.478733	0.002437	0.009541	2.488274
9	-6.023465	-0.014279	0.061052	-5.962413
10	1.567775	-0.024103	0.052225	1.620000
11	0.120409	0.004716	-0.000988	0.119421
12	0.470803	-0.006865	0.010435	0.481238
13	0.108849	0.029674	-0.072756	0.036093
14	-0.084456	-0.000889	-0.005425	-0.089881
15	-0.019703	-0.018787	0.045992	0.026288
16	-0.655038	-0.006993	0.018148	-0.636890
17	0.053151	0.028830	-0.077571	-0.024421
18	-0.542248	-0.008997	0.031837	-0.510411
19	0.418788	-0.009691	0.003553	0.422341
20	0.027192	-0.012413	0.020235	0.047427
21	-0.521699	-0.020982	0.035801	-0.485898
22	0.026544	0.014739	-0.036257	-0.009713

Fig. 2 (continued)

23	0.025617	0.012415	-0.020628	0.004989
24	0.106965	-0.014751	0.019131	0.126096
25	0.010931	0.000862	0.003784	0.014715
26	0.078250	0.014521	-0.003008	0.075242
27	-0.017986	0.002536	0.000400	-0.017586
28	-0.036740	0.005844	-0.010000	-0.046740
29	0.059767	-0.010038	0.023970	0.083736

	Maximum	Tolerance	Cnvgd?
Gradient	0.029674	0.000300	NO
Displacement	0.077571	0.000300	NO
Energy change	*****	0.000001	NO

New Cartesian Coordinates Obtained by Inverse Iteration

Cycle: 1 Maximum deviation: 0.07757147 RMS deviation: 0.03247425
 Cycle: 2 Maximum deviation: 0.00251062 RMS deviation: 0.00103769
 Cycle: 3 Maximum deviation: 0.00000221 RMS deviation: 0.00000090
 Cycle: 4 Maximum deviation: 0.00000000 RMS deviation: 0.00000000
 Full backtransformation converged in 4 cycles

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **

Searching for a Minimum

Optimization Cycle: 2

Energy is -363.816456941

Hessian Updated using BFGS Update

24 Hessian modes will be used to form the next step

Hessian Eigenvalues:

0.151802	0.160279	0.220081	0.222573	0.225149	0.238697
0.243289	0.333210	0.333512	0.334201	0.334467	0.334607
0.343195	0.364439	0.384434	0.402067	0.410203	0.415618
0.442013	0.451793	0.468281	0.475225	0.483089	0.592272

Minimum Search -- Taking Simple RFO Step

Searching for Lambda that Minimizes Along All modes

Value Taken Lambda = -0.00036212

Step Taken. Stepsize is 0.041360

	Maximum	Tolerance	Cnvgd?
Gradient	0.004782	0.000300	NO
Displacement	0.016892	0.000300	NO
Energy change	-0.006162	0.000001	NO

Fig. 2 (continued)

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **

Searching for a Minimum

Optimization Cycle: 3

Energy is -363.816671150

Hessian Updated using BFGS Update

25 Hessian modes will be used to form the next step

Hessian Eigenvalues:

0.130519	0.160040	0.163016	0.220115	0.222415	0.225439
0.237352	0.243655	0.333380	0.333585	0.334086	0.334518
0.334959	0.343578	0.347729	0.392411	0.400153	0.404240
0.415041	0.441905	0.448922	0.467657	0.472002	0.481316
0.592359					

Minimum Search -- Taking Simple RFO Step

Searching for Lambda that Minimizes Along All modes

Value Taken Lambda = -0.00002978

Step Taken. Step size is 0.012738

	Maximum	Tolerance	Cnvgd?
Gradient	0.001411	0.000300	NO
Displacement	0.005681	0.000300	NO
Energy change	-0.000214	0.000001	NO

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **

Searching for a Minimum

Optimization Cycle: 4

Energy is -363.816687350

Hessian Updated using BFGS Update

26 Hessian modes will be used to form the next step

Hessian Eigenvalues:

0.125429	0.160015	0.160339	0.163176	0.219075	0.222406
0.225145	0.240033	0.248037	0.333477	0.333628	0.334164
0.334523	0.335030	0.340061	0.343940	0.388531	0.394562
0.402733	0.413258	0.443109	0.446921	0.466923	0.472407
0.484096	0.575065				

Fig. 2 (continued)

Minimum Search -- Taking Simple RFO Step
 Searching for Lambda that Minimizes Along All modes
 Value Taken Lambda = -0.00000167
 Step Taken. Stepsize is 0.002748

	Maximum	Tolerance	Cnvgd?
Gradient	0.000329	0.000300	NO
Displacement	0.001530	0.000300	NO
Energy change	-0.000016	0.000001	NO

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **

Searching for a Minimum

Optimization Cycle: 5

Coordinates (Angstroms)			
ATOM	X	Y	Z
1 c	-1.139936	1.251805	0.000000
2 c	-2.184123	0.336460	0.000000
3 c	-1.936378	-1.052260	0.000000
4 c	-0.639460	-1.552003	0.000000
5 c	0.409411	-0.626076	0.000000
6 c	0.184879	0.779966	0.000000
7 n	1.776536	-0.819763	0.000000
8 c	1.477261	1.408786	0.000000
9 c	2.412786	0.408513	0.000000
10 h	-1.341534	2.320123	0.000000
11 h	-3.211059	0.691818	0.000000
12 h	-2.774412	-1.743982	0.000000
13 h	-0.449126	-2.622400	0.000000
14 h	2.239283	-1.715056	0.000000
15 h	1.683185	2.470391	0.000000
16 h	3.492686	0.463678	0.000000

Point Group: Cs Number of degrees of freedom: 29

Energy is -363.816688224

Cartesian Forces (au)			
ATOM	X	Y	Z
1 c	0.0000584	0.0000769	0.0000000
2 c	-0.0000324	-0.0000754	0.0000000
3 c	-0.0000236	0.0000544	0.0000000
4 c	0.0000253	0.0000010	0.0000000
5 c	0.0000447	0.0000047	0.0000000
6 c	-0.0000126	-0.0000224	0.0000000

Fig. 2 (continued)

7	n	0.0000101	-0.0000947	0.0000000
8	c	-0.0000175	0.0000305	0.0000000
9	c	-0.0000117	-0.0000575	0.0000000
10	h	-0.0000167	-0.0000124	0.0000000
11	h	0.0000098	0.0000222	0.0000000
12	h	0.0000205	-0.0000148	0.0000000
13	h	-0.0000036	-0.0000004	0.0000000
14	h	-0.0000063	0.0000234	0.0000000
15	h	-0.0000125	0.0000115	0.0000000
16	h	-0.0000212	0.0000385	0.0000000

Iterative generation of Internal Gradient

Cycle: 1 Maximum deviation: 0.00000045 RMS deviation: 0.00000017

Cycle: 2 Maximum deviation: 0.00000002 RMS deviation: 0.00000000

Cycle: 3 Maximum deviation: 0.00000000 RMS deviation: 0.00000000

Gradient converged in 3 cycles

Hessian Updated using BFGS Update

27 Hessian modes will be used to form the next step

Hessian Eigenvalues:

0.123423	0.145668	0.160015	0.160222	0.160354	0.219732
0.223230	0.234867	0.238794	0.266437	0.329579	0.333566
0.333688	0.334223	0.334658	0.336381	0.344207	0.373478
0.394783	0.402322	0.411472	0.445584	0.447885	0.466973
0.472563	0.490846	0.574440			

Minimum Search -- Taking Simple RFO Step

Searching for Lambda that Minimizes Along All modes

Value Taken Lambda = -0.00000011

Step Taken. Stepsize is 0.000794

Parameter Values and Displacements in Internal Coordinates

Coordinate	Current Value	Gradient	Displacement	New Value
1	0.174852	0.000041	-0.000220	0.174632
2	5.446165	0.000048	-0.000143	5.446022
3	0.220440	-0.000048	0.000192	0.220632
4	-3.186512	-0.000013	0.000045	-3.186467
5	0.005499	0.000019	-0.000053	0.005447
6	1.881035	-0.000030	0.000240	1.881275
7	-3.115511	-0.000001	-0.000056	-3.115567
8	2.497738	0.000031	-0.000321	2.497417
9	-5.949299	0.000024	-0.000100	-5.949400
10	1.621591	-0.000050	0.000147	1.621738
11	0.121225	0.000010	-0.000072	0.121153
12	0.483100	0.000032	-0.000143	0.482958
13	0.028700	-0.000023	0.000133	0.028833

Fig. 2 (continued)

14	-0.091955	-0.000001	0.000091	-0.091864
15	0.036701	0.000029	-0.000142	0.036559
16	-0.640073	0.000000	0.000031	-0.640042
17	-0.040582	-0.000043	0.000238	-0.040343
18	-0.513581	0.000025	-0.000128	-0.513709
19	0.416043	0.000048	-0.000159	0.415884
20	0.049659	0.000005	0.000034	0.049693
21	-0.490411	0.000020	-0.000080	-0.490491
22	-0.031165	0.000042	-0.000214	-0.031380
23	0.006415	-0.000016	0.000033	0.006449
24	0.139909	-0.000005	0.000087	0.139996
25	0.020428	0.000004	-0.000045	0.020383
26	0.091251	0.000013	-0.000090	0.091161
27	-0.014788	0.000006	-0.000093	-0.014881
28	-0.049225	-0.000002	-0.000010	-0.049236
29	0.094109	0.000053	-0.000269	0.093840
		Maximum	Tolerance	Cnvgd?
Gradient		0.000053	0.000300	YES
Displacement		0.000321	0.000300	NO
Energy change		-0.000001	0.000001	YES

Fig. 2 *PQS* output for the minimization of indole

Eq. 16), with a value of -0.01172058 . The resulting step length of 0.174879 is below the default maximum of 0.3 and so no scaling of the step is required.

The parameter values, and their gradients and displacements are then printed out. As with the Cartesian forces, these are not normally printed, and are not particularly informative given that each coordinate is potentially a linear combination of all underlying primitive internals (minus the torsions in the case of indole). A persistently large gradient for a particular coordinate would suggest some problems with the overall coordinate set, suggesting that either the delocalized internals, the underlying primitives or likely both, should be regenerated.

The convergence criteria – current value, tolerance, and whether or not convergence is satisfied – are then printed, together with details of the back-transformation (conversion of the current geometry in internal coordinates back into Cartesians, again not normally printed). This constitutes a complete optimization cycle.

On subsequent optimization cycles, assuming that the system is converging, the energy, the gradient components (both Cartesian and internal) and the step size should all decrease, with the latter two tending to zero. This is exactly what occurs for indole (see Fig. 2) which converges smoothly to the minimum energy structure. Another quantity that should tend to zero is the shift parameter, λ , which

again clearly does so; its magnitude on the last cycle is 0.00000011. Note that, as previously stated (see the comments under Eq. 16) λ is always negative and so approaches zero from below. The closer λ is to zero, the closer is the step taken to the simple Newton–Raphson step.

After the initial estimate, the Hessian matrix is updated at the beginning of each new optimization step. This is done using the BFGS update, as appropriate for a minimization. Indole is a planar, rigid molecule and so, not surprisingly, convergence is achieved quickly taking just five optimization cycles. At final convergence the displacement criteria is not fully satisfied; this is not an error as the default convergence criteria in *PQS* allow for just one of the displacement and energy change criteria to be satisfied (see the footnotes under Table 1); this is to prevent wasteful cycles near the end of the optimization for excessively floppy molecules for which relatively large displacements result in only small changes in energy. If it is desired that the displacement criterion be rigorously satisfied, then the energy criterion should be reduced so that it is almost impossible to satisfy (for example to 10^{-8} Eh); Convergence can then only be achieved by satisfying the displacement criterion.

Transition State Searches: An Example

Figure 3 shows the optimization output at each cycle (truncated for the intermediate cycles) in the standard default EF transition state search for the reaction $\text{HCN} \leftrightarrow \text{HNC}$. This is a very simple reaction but is still instructive. Both reactant and product are linear. Reaction can be considered to occur essentially by motion of the hydrogen atom in a semicircle centered about the midpoint of the C–N bond, with the $\angle\text{HCN}$ angle going from 180° in HCN to 0° in HNC (CNH).

Consider the starting geometry. A first approximation to the transition state (TS) geometry might be to position the hydrogen atom directly vertically above the midpoint of the C–N bond. (The H–C distance of course has to be decided.) This corresponds to $\angle\text{HCN}$ a bit less than 90° . However, knowing the relative energetics (HNC is less stable than HCN, the reaction as written is endothermic), the Hammond postulate (Hammond 1955) suggests that the TS is likely to resemble HNC more than HCN, in which case $\angle\text{HCN}$ will very likely be noticeably less than 90° , i.e., relative to HCN we have a late transition state. Now this is only a simple consideration, but the use of “chemical intuition” in this way can potentially save several optimization cycles. In this particular example, it does not save much because the system as a whole only has three atoms and the transition state is fairly easy to locate, but starting with $\angle\text{HCN} = 85^\circ$ instead of 70° costs an extra optimization cycle, and starting with $\angle\text{HCN} = 130^\circ$ takes three more cycles. (The other starting parameters were $r\text{C} - \text{N} = 1.21 \text{ \AA}$ and $r\text{H} - \text{C} = 1.5 \text{ \AA}$.) As has already been mentioned, a good initial Hessian is almost mandatory for a TS search and a full Hessian at the starting geometry was computed at the semiempirical PM3 level (Stewart 1989) for the $\text{HCN} \leftrightarrow \text{HNC}$ reaction.

The output for a transition state search is of course similar to that for a minimization. Quantities such as the Cartesian geometry, the SCF energy, the forces, parameter values, and displacements are common to both TS searches and minimizations and are printed out in the same way. The first difference occurs in the treatment of the starting Hessian: for the indole minimization (Fig. 2) a default Hessian, diagonal in the underlying space of primitive internals, is used whereas the TS search started with a full, albeit approximate, Cartesian Hessian which needs to be transformed into internal coordinates. This transformation is done once only, on the first optimization cycle, as thereafter the Hessian will be updated.

Note the comment printed on cycle 1 that the Hessian transformation does not include the derivative of the \mathbf{B} -matrix. The \mathbf{B} -matrix (Wilson et al. 1955) formally relates a displacement in internal coordinates (q) to one in Cartesian coordinates (X), $\Delta q = \mathbf{B}\Delta X$, and it turns out that a full transformation of the Hessian to internal coordinates requires the derivative of the \mathbf{B} -matrix; the precise equation is Pulay (1977):

$$\mathbf{H}^{\text{int}} = \mathbf{B}^{\text{inv}} \left(\mathbf{H}^{\text{cart}} - \mathbf{g}^{\text{int}} \frac{d\mathbf{B}}{d\mathbf{X}} \right) (\mathbf{B}^{\text{inv}})^{\text{T}}, \quad (24)$$

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **
 Searching for a Transition State

Optimization Cycle: 1

		Coordinates (Angstroms)		
ATOM		X	Y	Z
1	c	-0.469846	0.571010	0.000000
2	n	-0.469846	-0.628990	0.000000
3	h	0.939693	0.057980	0.000000

Point Group: Cs Number of degrees of freedom: 3

Energy is -93.319741527

		Cartesian Forces (au)		
ATOM		X	Y	Z
1	c	0.0404791	-0.0073674	0.0000000
2	n	0.0272344	0.0058884	0.0000000
3	h	-0.0677118	0.0014596	0.0000000

Iterative generation of Internal Gradient
 Transforming Cartesian Hessian to Internal Coordinates

Fig. 3 (continued)

Hessian Transformation does not Include Derivative of B-matrix

3 Hessian modes will be used to form the next step

Hessian Eigenvalues:

-0.187665 0.238971 1.173533

Transition State Search -- Taking P-RFO Step

Searching for Lambda that Maximizes Along the Lowest mode

Value Taken Lambda = 0.00189009

Searching for Lambda that Minimizes Along All other modes

Value Taken Lambda = -0.02788347

Calculated Step too Large. Step scaled by 0.888690

Step Taken. Stepsize is 0.300000

Parameter Values and Displacements in Internal Coordinates

Coordinate	Current Value	Gradient	Displacement	New Value
1	2.134802	0.074274	-0.131560	2.003242
2	-0.247567	-0.037317	0.203494	-0.044073
3	-3.170360	-0.032424	0.176868	-2.993492

	Maximum	Tolerance	Cnvgd?
Gradient	0.074274	0.000300	NO
Displacement	0.203494	0.000300	NO
Energy change	*****	0.000001	NO

New Cartesian Coordinates Obtained by Inverse Iteration

Cycle: 1 Maximum deviation: 0.20349400 RMS deviation: 0.17320508

Cycle: 2 Maximum deviation: 0.00532215 RMS deviation: 0.00463403

Cycle: 3 Maximum deviation: 0.00001577 RMS deviation: 0.00001192

Cycle: 4 Maximum deviation: 0.00000000 RMS deviation: 0.00000000

Full backtransformation converged in 4 cycles

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **

Searching for a Transition State

Optimization Cycle: 2

Energy is -93.339112477

Hessian Updated using Powell/Murtagh-Sargent Update

Mixing factors: 0.645582 Powell

 0.354418 Murtagh-Sargent

Fig. 3 (continued)

```

3 Hessian modes will be used to form the next step
Hessian Eigenvalues:
  -0.196173    0.099238    1.154056

Transition State Search -- Taking P-RFO Step
Searching for Lambda that Maximizes Along the Lowest mode
Value Taken    Lambda =    0.00827163
Searching for Lambda that Minimizes Along All other modes
Value Taken    Lambda =   -0.02009120
Calculated Step too Large.  Step scaled by  0.661233
Step Taken.  Stepsize is  0.300000

                Maximum    Tolerance    Cnvgd?
Gradient        0.062199    0.000300    NO
Displacement    0.233225    0.000300    NO
Energy change   -0.019371    0.000001    NO

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **
  Searching for a Transition State

Optimization Cycle:   3

Energy is    -93.342916439

Hessian Updated using Powell/Murtagh-Sargent Update
Mixing factors:    0.556632 Powell
                  0.443368 Murtagh-Sargent

3 Hessian modes will be used to form the next step
Hessian Eigenvalues:

  -0.214546    0.216669    1.155709

Transition State Search -- Taking P-RFO Step
Searching for Lambda that Maximizes Along the Lowest mode
Value Taken    Lambda =    0.00009941
Searching for Lambda that Minimizes Along All other modes
Value Taken    Lambda =   -0.00011765
Step Taken.  Stepsize is  0.030919

                Maximum    Tolerance    Cnvgd?
Gradient        0.006976    0.000300    NO
Displacement    0.028637    0.000300    NO
Energy change   -0.003804    0.000001    NO

```

Fig. 3 (continued)

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **
Searching for a Transition State

Optimization Cycle: 4

Energy is -93.342897201

Hessian Updated using Powell/Murtagh-Sargent Update

Mixing factors: 0.111685 Powell
0.888315 Murtagh-Sargent

3 Hessian modes will be used to form the next step

Hessian Eigenvalues:
-0.203277 0.258806 1.161986

Transition State Search -- Taking P-RFO Step

Searching for Lambda that Maximizes Along the Lowest mode

Value Taken Lambda = 0.00000265

Searching for Lambda that Minimizes Along All other modes

Value Taken Lambda = -0.00000882

Step Taken. Stepsize is 0.006795

	Maximum	Tolerance	Cnvgd?
Gradient	0.001321	0.000300	NO
Displacement	0.004864	0.000300	NO
Energy change	0.000019	0.000001	NO

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **
Searching for a Transition State

Optimization Cycle: 5

Energy is -93.342899327

Hessian Updated using Powell/Murtagh-Sargent Update

Mixing factors: 0.377168 Powell
0.622832 Murtagh-Sargent

3 Hessian modes will be used to form the next step

Hessian Eigenvalues:
-0.177469 0.279946 1.171011

Transition State Search -- Taking P-RFO Step

Searching for Lambda that Maximizes Along the Lowest mode

Fig. 3 (continued)

Value Taken Lambda = 0.00000027
 Searching for Lambda that Minimizes Along All other modes
 Value Taken Lambda = -0.00000018
 Step Taken. Stepsize is 0.001445

	Maximum	Tolerance	Cnvgd?
Gradient	0.000343	0.000300	NO
Displacement	0.001440	0.000300	NO
Energy change	-0.000002	0.000001	NO

** GEOMETRY OPTIMIZATION IN DELOCALIZED INTERNAL COORDINATES **

Searching for a Transition State

Optimization Cycle: 6

Coordinates (Angstroms)			
ATOM	X	Y	Z
1 c	-0.341598	0.551098	0.000000
2 n	-0.417301	-0.638152	0.000000
3 h	0.758898	0.087054	0.000000

Point Group: Cs Number of degrees of freedom: 3

Energy is -93.342899294

Cartesian Forces (au)			
ATOM	X	Y	Z
1 c	0.0000015	0.0000201	0.0000000
2 n	-0.0000112	-0.0000468	0.0000000
3 h	0.0000097	0.0000043	0.0000000

Iterative generation of Internal Gradient

Cycle: 1 Maximum deviation: 0.00000053 RMS deviation: 0.00000036
 Cycle: 2 Maximum deviation: 0.00000000 RMS deviation: 0.00000000

Gradient converged in 2 cycles

Hessian Updated using Powell/Murtagh-Sargent Update

Mixing factors: 0.600156 Powell
 0.399844 Murtagh-Sargent

3 Hessian modes will be used to form the next step

Hessian Eigenvalues:

-0.188567 0.275416 1.153882

Fig. 3 (continued)

```

Transition State Search -- Taking P-RFO Step
Searching for Lambda that Maximizes Along the Lowest mode
Value Taken   Lambda =   0.00000000
Searching for Lambda that Minimizes Along All other modes
Value Taken   Lambda =   0.00000000
Step Taken.   Stepsize is 0.000133

      Parameter Values and Displacements in Internal Coordinates
Coordinate   Current Value   Gradient   Displacement   New Value
    1             2.021093   -0.000032   -0.000069     2.021024
    2             0.137423   -0.000025   -0.000037     0.137386
    3            -2.754498    0.000023   -0.000108    -2.754606

                                Maximum   Tolerance   Cnvgd?
Gradient                0.000032    0.000300    YES
Displacement            0.000108    0.000300    YES
Energy change           0.000000    0.000001    YES

```

Fig. 3 PQS output for the HCN \leftrightarrow HNC Transition State Search

where

$$\mathbf{B}^{\text{inv}} = (\mathbf{B}\mathbf{B}^{\text{T}})^{-1}\mathbf{B}, \quad (25)$$

and a superscript ^T represents a matrix transpose.

If you require an *exact* transformation of all quantities from Cartesian to internal coordinates, then the $d\mathbf{B}/d\mathbf{X}$ term must be included; however it is usually advantageous to leave it out as its inclusion often changes the Hessian eigen value structure. In general, Cartesian Hessian matrices with a known eigen value structure will retain that structure if the $d\mathbf{B}/d\mathbf{X}$ term is omitted; this is far from guaranteed when it is included. When starting off a geometry optimization (either a minimization or a TS search) using an optimized geometry and starting Hessian obtained from a previous calculation at a lower level of theory – a common occurrence – then the $d\mathbf{B}/d\mathbf{X}$ term should almost certainly be omitted. Note that for a Hessian transformation at a stationary point, the gradient is zero in any case and so the $d\mathbf{B}/d\mathbf{X}$ term will also be zero. For the HCN \leftrightarrow HNC TS search the starting Hessian has the correct eigen value structure, i.e., one negative eigen value, so our initial geometry was at least reasonable.

The other major difference from the indole minimization is with the shift parameter(s). A TS search uses the P-RFO step, as opposed to the simple RFO step used for minimization (see Eqs. 17, 18, and 19), and so there are *two* shift parameters instead of just one. As explained previously, one shift parameter maximizes along the TS mode (the eigenvector with the negative eigen value), while the other minimizes along all the other Hessian modes. The first shift parameter is positive

while the other is negative. On the first cycle the computed step length is initially greater than the default maximum of 0.3 and so the step is scaled down (by 0.888690).

As with a minimization, on subsequent optimization cycles the gradient components and the step size should both decrease, tending to zero if the structure is converging. The two shift parameters, λ , also both tend to zero, one from above and the other from below. On the last cycle they are both zero to all eight decimal places printed out. The energy may increase or decrease, depending on whether the increase in energy from maximization along the TS mode is greater or smaller than the decrease in energy from minimization along all the other modes; the change in energy however should of course again tend to zero. (Overall a decrease in energy is the more likely, especially as the system gets larger, as the number of modes which are minimized becomes far greater than the single mode along which the energy is maximized.)

The Hessian update, at the beginning of each new optimization step, is done using the Bofill formula (Bofill 1994), which is the default for a TS search. This is a linear combination of the Powell (1971) and Murtagh and Sargent (1970) updates, and the two mixing coefficients are printed out. They vary from cycle to cycle, as can be seen (Fig. 3). The final TS geometry is well converged (in six cycles), with the three convergence criteria all well below their respective tolerances.

Using Optimized Potential Scans in Transition State Searches

We turn now to a somewhat more complicated example, namely, a search for the transition state for the parent Diels–Alder reaction between *cis*-1,3-dibutadiene and ethylene to give cyclohexene. Here we use an optimized potential scan to obtain an initial approximation to the TS geometry. The starting geometry for the potential scan is shown in Fig. 4. This was obtained by building cyclohexene using *PQSMol*, the graphical user interface for PQS (2010), and then stretching the appropriate C–C bonds ($C_{10} - C_{13}$ and $C_7 - C_{16}$ as shown in Fig. 4) to separate off an ethylene fragment. The overall system symmetry (C_s) was maintained using the symmetrization tools available in *PQSMol*. The $\angle HCH$ angles in the ethylene fragment are likely to be too acute, but this will be rectified during the potential scan.

The $C_{10} - C_{13}$ and $C_7 - C_{16}$ distances in the starting geometry (see Fig. 4; the two distances are equivalent by symmetry) are 1.720 Å. What we are going to do is scan this distance over a range of (increasing) values, at each scanned value optimizing all the remaining degrees of freedom. If the scanned variable(s) is a good approximation to the reaction coordinate, and if the value of the variable at the transition state geometry lies within the scanned range, then what we should see is a maximum in the energy somewhere along the scan. The (optimized) geometry at this energy maximum should be a very good approximation to the transition state.

An optimized potential scan is effectively a series of constrained geometry optimizations, with the constraint being the scanned variable. *PQS* has powerful

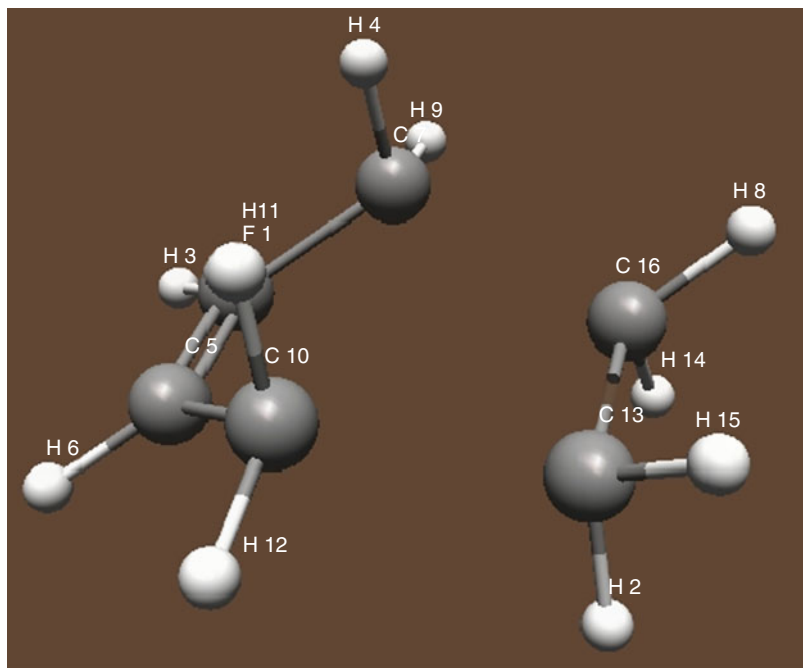


Fig. 4 Starting structure for potential scan for Diels–Alder TS search

algorithms for constrained optimization using the method of Lagrange multipliers (Baker 1992, 1997; Baker and Bergeron 1993). Because of the symmetry of the system, the scanned variable is not a single distance but rather a linear combination (with equal weights) of the $C_{10} - C_{13}$ and $C_7 - C_{16}$ distances. The requested linear combination is normalized by the program and so the actual parameter scanned is $\frac{1}{\sqrt{2}} [(r_{C_{10} - C_{13}}) + (r_{C_7 - C_{16}})]$. The potential energy scan module in *PQS* is capable of scanning any linear combination of primitive internal coordinates selected by the user.

The optimized energies at $C_{10} - C_{13}$ and $C_7 - C_{16}$ distances from 1.720 Å to about 2.65 Å are shown in Table 2. Note that, although the final level of theory used in this chapter for all examples given is B3LYP/6-31G*, the scan was actually carried out using the smaller 3-21G (Binkley et al. 1980) basis set. This is simply to save computer time. I expect the B3LYP/3-21G energy surface to be very similar to the B3LYP/6-31G* one in this system, and since we are using the optimized scan to obtain a good, but still approximate, starting geometry for the proper TS search, there is no point using the larger basis set for the scan.

As can be seen in Table 2, there is a clear energy maximum at a parameter value somewhere around 8.75. The parameter printed out is the sum of the two scanned C–C distances (unnormalized) given in atomic units (bohr); the maximum is equivalent to a C–C distance of 2.315 Å. During the scan, the first energy (from

Table 2 Optimized B3LYP/3-21G energies (E_h) for the Diels–Alder Potential Scan at combined scan distances (in bohr) from 6.50 to 10.00 in steps of 0.25

Current value: 6.5000	Energy is -233.352945649
Current value: 6.7500	Energy is -233.341556134
Current value: 7.0000	Energy is -233.328803153
Current value: 7.2500	Energy is -233.315693812
Current value: 7.5000	Energy is -233.303165136
Current value: 7.7500	Energy is -233.292127679
Current value: 8.0000	Energy is -233.283459373
Current value: 8.2500	Energy is -233.277798713
Current value: 8.5000	Energy is -233.275148592
Current value: 8.7500	Energy is -233.274846534
Current value: 9.0000	Energy is -233.276044729
Current value: 9.2500	Energy is -233.278051315
Current value: 9.5000	Energy is -233.280392073
Current value: 9.7500	Energy is -233.282767273
Current value: 10.0000	Energy is -233.285004883

the starting geometry) took the most optimization cycles (12) to converge; thereafter each subsequent point typically converged in five to seven steps. Note that if the scanned range is insufficient to locate the maximum, it can simply be extended. For example, if the initial scan only went as far as 8.00 bohr (for the combined C–C distances), then it can be repeated for distances greater than this starting from the last optimized geometry found. That the scan needs to go to greater C–C distances is clear from the energy, which is increasing over the initial scan range; if it had been decreasing then one would need to go to shorter C–C distances.

The full B3LYP/6-31G* TS search commenced using the optimized geometry at the scan maximum with a full Hessian matrix computed at this geometry using the 3-21G basis (the same basis as used in the scan). Not surprisingly this starting Hessian had the desired one negative eigen value, and the search converged in just four optimization cycles, indicating just how good the geometry resulting from the scan was. A full vibrational analysis at the B3LYP/6-31G* level on the converged transition state structure indicated that it was indeed a genuine TS. The C₁₀–C₁₃ and C₇–C₁₆ distances in the final TS were 2.272 Å, very close to those in the structure at the scan maximum.

Comparison of Experimental and Theoretical Geometries

Before turning to the second part of this chapter, the computational simulation of vibrational spectra, I want to make a few comments on the accuracy of theoretically computed geometries. Direct comparison with reliable experimental data shows that theoretical geometries are now really very accurate indeed, particularly for “normal” organic molecules. The level of theory used here, B3LYP/6-31G*, is typically accurate for, say, bond angles to within a few degrees and bond lengths to within 0.02 Å at worst, and if any experimental bond distance differs by much more

than this from the theoretical value, it is more likely in my opinion for experiment to be in error rather than theory. One needs to bear in mind that none of the experimental methods commonly used in molecular structure determination, X-ray crystallography, microwave spectroscopy, nuclear magnetic resonance (NMR) etc., measure a bond distance directly; rather some kind of spectrum (or pattern) is obtained which has to be interpreted, and bond distances are extracted indirectly from this. It was fairly typical in the early 1970s (and even into the 1980s) for X-ray diffraction structures to be published, as a set of Cartesian coordinates, which resulted in quite ridiculous bond distances involving hydrogen. This was due principally of course to the difficulty in locating hydrogen atoms because of the much lower electron density surrounding hydrogen nuclei compared to heavier elements, and the fact that this density could easily be displaced in bonds involving more electronegative atoms. From the mid-1970s onward it became increasingly more common for theoretically computed bond distances involving hydrogen to be used in the fitting of X-ray data; this also resulted in much better agreement with heavy atom bond distances as well (Schäfer 1983).

Another factor to be considered is that theory and experiment do not quite measure the same thing. Within the Born–Oppenheimer approximation (Born and Oppenheimer 1927; Wikipedia 2010), theory ideally determines the geometry at the bottom of the well, at the energy minimum, whereas this geometry is inaccessible to experiment due to the existence of zero-point vibrational energy. Such differences, however, are usually minor.

Table 3 provides a comparison between optimized B3LYP/6-31G* and experimental geometrical parameters for a number of small, organic molecules. These results were taken from two papers published in the mid-1990s; a DFT study of some organic reactions (Baker et al. 1985) and a systematic DFT study of fluorination in methane, ethane, and ethylene (Muir and Baker 1996). The focus was mainly on the energetics, but many computed and experimental geometrical parameters were provided. The theoretical approaches did not include B3LYP, using instead a related functional (Becke's original ACM (Becke 1993; see also Hertwig and Koch 1997), better known today as B3PW91), but I have reoptimized all the geometries at the B3LYP/6-31G* level for compatibility with this chapter.

As can be seen from Table 3, the agreement between the computed and experimental geometrical parameters is excellent. The mean average deviation between the B3LYP/6-31G* and experimental bond lengths is just 0.006 Å, with a maximum error (for the C–N triple bond in HCN) of 0.022 Å. The mean average deviation for bond angles is only 0.8°. Such results are typical.

Included in Table 3 is the series of fluorine-substituted methanes: mono-, di-, tri-, and tetra-fluoromethane. These compounds exhibit a well-known substituent effect (the geminal effect) of decreasing C–F bond length (and increasing bond strength) with increasing fluorine substitution on the same carbon atom. This effect of fluorine (not seen with any of the other halogens) has been known for a long time (although there are still arguments about the explanation for it) and is very accurately reproduced by the theoretical calculations.

Table 3 Optimized B3LYP/6-31G* and experimental geometrical parameters for a number of small organic molecules

Parameter	B3LYP	Experiment	Parameter	B3LYP	Experiment
Vinyl alcohol (CH ₂ = CHOH)			trans-1,1,2,2-tetrafluoroethane (CHF ₂ CHF ₂)		
rC – O	1.363	1.369	rC – C	1.524	1.518
rC = C	1.334	1.335	rC – F	1.360	1.350
rO – H	0.972	0.962	rC – H	1.095	1.098
rC – H	1.086	1.080	∠CCF	108.8	108.2
∠CCO	127.2	126.0	∠CCH	111.8	110.3
∠COH	108.7	108.5	∠FCF	108.8	107.3
Acetaldehyde (CH ₂ CHO)			Carbonyl fluoride (CF ₂ O)		
rC = O	1.211	1.210	rC = O	1.180	1.17
rC – C	1.508	1.515	rC – F	1.322	1.32
rC – H	1.114	1.128	∠FCF	107.8	112.5
∠CCO	124.7	124.1	Hydrogen fluoride (HF)		
trans-butadiene (CH ₂ = CHCH = CH ₂)			rH – F	0.934	0.917
rC = C	1.341	1.349	Hydrogen (H ₂)		
rC – C	1.458	1.467	rH – H	0.743	0.741
∠C = CH	119.4	120.9	Water (H ₂ O)		
∠CCC	124.3	124.4	rO – H	0.969	0.958
Ethylene (CH ₂ = CH ₂)			∠HOH	103.6	104.5
rC = C	1.331	1.339	Formaldehyde (CH ₂ O)		
rC – H	1.087	1.087	rC = O	1.206	1.208
∠CCH	121.9	121.3	rC – H	1.110	1.116
Cyclohexene (C ₆ H ₁₀)			∠HCH	115.2	116.5
rC = C	1.337	1.334	Acetylene (C ₂ H ₂)		
rC – C	1.510	1.50	rC – C	1.205	1.203
rC – C	1.537	1.52	rC – H	1.067	1.060
C – C	1.535	1.54	Ethane (CH ₃ CH ₃)		
∠CC = C	123.5	123.4	rC – C	1.531	1.535
∠CCC	112.0	112.0	rC – H	1.096	1.094
∠CCC	110.9	110.9	∠CCH	111.4	111.2
Tetrazine (C ₂ N ₄ H ₂)			Fluoromethane (CH ₃ F)		
rN – N	1.326	1.321	rC – F	1.383	1.382
rC – N	1.339	1.334	rC – H	1.096	1.095
∠NCH	126.6	127.4	∠HCH	109.3	110.4
Hydrogen cyanide (HCN)			Difluoromethane (CH ₂ F ₂)		
rC – N	1.157	1.135	rC – F	1.361	1.357
rC – H	1.070	1.066	rC – H	1.096	1.093
Nitrogen (N ₂)			∠HCH	112.2	113.7
rN – N	1.105	1.098	∠FCF	109.2	108.3
rC = C	1.340	1.342	rC – F	1.342	1.333
rC – C	1.519	1.517	rC – H	1.093	1.098
rC – C	1.573	1.566	∠FCF	108.6	108.8
∠CCC	86.5	85.5	Tetrafluoromethane (CF ₄)		
∠C = CC	94.4	94.2	rC – F	1.329	1.323

Geometry Optimization of Molecular Clusters

The situation at the time of writing is that molecular structures for stable or metastable minima are readily located theoretically using powerful optimization algorithms usually carried out in natural, delocalized or redundant internal coordinates. The resulting structural parameters (computed using B3LYP or better with a decent basis set) are typically in excellent agreement with experimental results where these are available. The location of transition states can be more difficult, but perseverance usually brings its reward. Theoretical calculations are essentially the only means for providing structural data on transition states as they are extremely difficult to study experimentally.

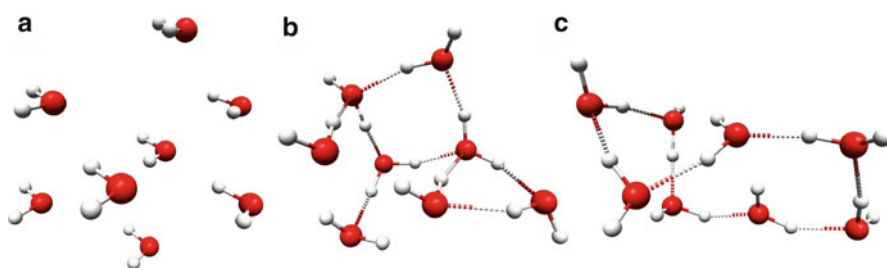
One area where standard natural or delocalized internal coordinates are not the best choice is for optimizing the geometry of molecular clusters. For larger clusters the problem of global versus local minimization becomes an issue, but even for relatively small clusters (containing half-a-dozen or so molecules) deriving a set of coordinates suitable for an efficient optimization from stretches, bends, and torsions alone usually results in poor optimization performance. One reason for this is that there are really two different types of forces involved in determining the structure of a molecular cluster: the normal, fairly strong intramolecular forces that determine the geometry of each individual molecule, and the weak intermolecular forces that determine the arrangement of each molecule relative to the others. The latter are best described in terms of scaled inverse distance coordinates (*not* standard distances) and an efficient algorithm which uses normal primitive internal coordinates to describe the geometry of each molecule in the cluster, and scaled inverse distances (with a cutoff) to describe the relative position of the cluster molecules with respect to each other (Baker and Pulay 2000) is available in *PQS*. The resulting set of coordinates, which are linear combinations of standard primitive internals and scaled inverse distances, are known as cluster coordinates.

Examples of geometry optimization of some small molecular clusters are shown in Table 4. Optimization performance using cluster coordinates ($1/R$ with a 5 Å cutoff) is compared with Cartesian coordinates. A modest level of theory was used (Hartree–Fock with the 3-21G basis set) and all optimizations were started with a unit Hessian matrix. In most cases the starting geometry was poor, and very far from the final converged structure. No comparisons with natural or delocalized internal coordinates are included in Table 4; with such poor starting geometries this is usually a waste of time as the large changes that occur in the relative positions of molecules in the cluster often result in individual bond angles, as defined, exceeding 180° causing problems with the various transformations, particularly the Cartesian backtransformation. This is yet another reason why standard internal coordinates are inappropriate for optimizing the geometry of molecular clusters. This is not a problem with inverse distance (or simple distance) coordinates as these are defined essentially from zero to infinity.

As shown in Table 4, cluster coordinates typically converge much faster than Cartesians (on average for the seven clusters shown by a factor of almost 5) and usually – but not always – to lower energy structures. (Five out of the seven

Table 4 Comparison of the performance of cluster coordinates and Cartesian coordinates for the optimization of some small molecular clusters

Clusters	Cluster coordinates		Cartesian coordinates	
	Cycles	Energy	Cycles	Energy
4 × benzene	52	−917.687877	273	−917.680357
6 × DMSO	124	−3,292.233057	660	−3,292.192089
formamide + 6 × water	75	−621.674713	222	−621.687307
6 × methanol	39	−686.489506	301	−686.492821
5 × pyridine	108	−1,226.595255	479	−1,226.587376
3 × THF	65	−689.117483	348	−689.113449
8 × water	69	−604.913202	331	−604.911548
Average cycles	76		373	

**Fig. 5** Starting and final structures for optimization of the $(\text{H}_2\text{O})_8$ cluster from Table 4: (a) starting geometry, (b) final geometry cluster coords, (c) final geometry Cartesian

converged to lower energies than the corresponding Cartesian optimizations, two to higher.) The starting and final geometries are shown for the eight water molecule cluster in Fig. 5.

Geometry Optimization in the Presence of External Forces

I want to end this section by briefly mentioning a simple but relatively new technique which has proven to be very useful in deriving potentially new molecular structures. It was originally developed to study the geometrical changes inferred in atomic force microscopy and similar experiments (known collectively as mechano-chemistry) (Beyer and Clausen-Schaumann 2005), but can be applied as a general optimization tool which also has potential for the location of transition states. This is geometry optimization in the presence of external forces (Wolinski and Baker 2009) known as enforced geometry optimization (EGO). The force is applied along a straight line joining any two atoms in the molecule under study, either to push the two atoms together or to pull them apart. Whatever force is applied to one atom of the pair, an equal and opposite force is applied to the other. This ensures that there is no tendency for the molecule to translate or rotate; in essence it remains in place

under simulated pressure. Forces can be applied between as many pairs of atoms as desired, although calculations to date have typically involved only a single pair of atoms.

If you are applying an external force to, say, push two atoms together in a stable molecule there will be a tendency for a bond to form between the two atoms. If the applied force is insufficient to form a bond (or if a bond won't form anyway), there will be some minor geometrical changes as the system endeavors to compensate for the excess force. If the applied force is sufficient, then what you should see is the energy rise as the two atoms are pushed together, then fall as the bond forms and additional molecular rearrangement occurs. The final "optimized" geometry will still involve a structural arrangement under considerable strain, but if this structure is allowed to relax (by reoptimizing with the additional force removed) it could well converge to a new, stable geometry, different from the initial structure. If it does so, then the maximum energy structure found as the two atoms were pushed together is often a reasonable estimate for the transition state joining the original and the new structure.

The method as described above has been used to find seven previously unknown geometrical isomers of $C_{10}N_2H_8$, starting from the ground-state geometry of *cis*-azobenzene, at the B3LYP/6-31G* level by pushing together various pairs of atoms, one from each of the two benzene rings (Wolinski and Baker 2009). All of the new structures were classified by vibrational analysis, and all were found to be minimal. Transition states between each of the new structures and *cis*-azobenzene were also located and characterized. Additional calculations were carried out at MP2/6-311G** to ensure that the energetics remained stable. The technique overall shows considerable promise for the discovery of new chemistry and new chemical structures.

Molecular Vibrations

A molecular system containing N atoms nominally has $3N$ degrees of freedom. In Cartesian coordinates, one can consider each degree of freedom as separate X , Y , or Z displacements of each atom about its equilibrium position. This atomic motion can be induced for example by heat or by interaction with light or other electromagnetic radiation. Because the atoms in a molecule are connected by chemical bonds, the various atomic motions are coupled. In a classical treatment, it turns out that three of these motions involve overall translation of the system as a whole (along the X , Y , and Z directions of a standard Cartesian axis system), and three of them involve overall rotation of the system as a whole (about the same axes). These are known as rigid-body motions. The Cartesian axes pass through (they have their origin at) the center of mass of the system and the X , Y , and Z axes coincide with the principal axes of inertia of the undistorted molecule. Under these rigid-body motions the positions of the atoms in the molecule relative to each other remain unchanged, and so the potential energy of the system in a vacuum is also unchanged. The remaining $3N - 6$ coupled motions, which do affect the relative atomic positions and hence the

potential energy, are classified as molecular vibrations. (Note that a linear molecule has $3N - 5$ vibrations, not $3N - 6$, as one of the rotations – that about its own axis – cannot be observed, and is effectively ignored.)

To a reasonable approximation, at least for small displacements, the vibrations in a polyatomic molecule can be described as a kind of simple harmonic motion. This is essentially equivalent to considering a chemical bond between two atoms as a weightless spring that obeys Hooke's law (i.e., the force is proportional to the displacement). The simplest vibration is that in a diatomic molecule, and we can refer back to the potential energy curve for He_2^{2+} (Fig. 1) as an example. At the minimum, near the bottom of the well, the potential energy curve is indeed parabolic to a very good approximation.

The force/displacement equation for a diatomic molecule obeying Hooke's law is

$$F = -kq, \quad (26)$$

where F is the applied force, k is the so-called force constant, and q is the displacement from equilibrium. The negative sign indicates that the force exerted by the spring is in a direction opposite to that of the displacement. The potential energy stored in the spring (the integral of force over distance) is

$$E = \frac{1}{2}kq^2. \quad (27)$$

By Newton's second law of motion, the force can also be given by mass times acceleration

$$F = m \frac{d^2q}{dt^2}. \quad (28)$$

The two forces in Eqs. 26 and 28 are one and the same, and so the differential equation relating displacement, q , with time, t , for the harmonic oscillator is

$$m \frac{d^2q}{dt^2} + kq = 0. \quad (29)$$

The solution to Eq. 29 is

$$q(t) = A \cos(2\pi\nu t); \quad \nu = \frac{1}{2\pi} \sqrt{\frac{k}{m}}, \quad (30)$$

where A is the amplitude of the vibrational coordinate q and ν is the fundamental frequency of the vibration. The "mass," m , is actually the so-called reduced mass, μ , which for a diatomic molecule A–B, with atomic masses m_A and m_B , respectively, is given by

$$\frac{1}{\mu} = \frac{1}{m_A} + \frac{1}{m_B}. \quad (31)$$

Using the reduced mass ensures that the center of mass of A–B remains unchanged during the vibrational motion.

From Eq. 27 the force constant, k , is given by

$$k = \frac{d^2 E}{dq^2}. \quad (32)$$

The force constant in a one-dimensional system (a system with a single degree of vibrational freedom) is thus directly related to the curvature of the potential energy surface, i.e., the shape of the well. From the force constant, one can directly determine the fundamental frequency (via Eq. 30). In a polyatomic molecule the single $d^2 E/dq^2$ term is replaced by the Hessian matrix, \mathbf{H} (often called the force constant matrix). In Cartesian coordinates this is the same Hessian matrix introduced in Eq. 4 when discussing the Newton–Raphson step during a geometry optimization, and a classical treatment of molecular vibrations and the Newton–Raphson step both assume the same thing about the potential energy surface, namely that it is quadratic. The only major difference is that a vibrational analysis involves the atomic masses, whereas the potential energy surface is independent of mass. The fundamental vibrational frequencies for a polyatomic molecule can be extracted from the mass-weighted Hessian matrix, and the normal modes (the coupled motion of the atoms during the vibration) can be obtained by a suitable treatment of the Hessian eigenvectors. The precise steps involved are as follows:

1. Mass-weight the Cartesian Hessian matrix.

$$\mathbf{H}_{ij}^{\text{MW}} = \frac{\mathbf{H}_{ij}}{\sqrt{m_i m_j}}. \quad (33)$$

2. Project out the translations and rotations from the mass-weighted Hessian.

As discussed briefly above, the translational and rotational degrees of freedom can be separated out from the vibrations for a rigid body. This requires shifting the origin of the coordinate system to the center of mass, determining the principal moments and the rotation generators, setting up orthogonal coordinate vectors for translation and rotation about the principal axes of inertia, setting up a projection matrix \mathbf{P} from these vectors and using it to “remove” the translations and rotations from the Hessian matrix, leaving only the vibrational modes.

If the orthogonal set of vectors representing translation and rotation (there are six of them, each of length $3N$, where N is the number of atoms) are given by \mathbf{R} , then the projection matrix \mathbf{P} is formed by

$$\mathbf{P}_{ij} \ (i \neq j) = -\sum_k \mathbf{R}_{ik} \mathbf{R}_{jk}; \quad \mathbf{P}_{ii} = 1 - \sum_k \mathbf{R}_{ik} \mathbf{R}_{ik}, \quad (34)$$

and the projected Hessian is given by

$$\mathbf{H}^{\text{P}} = \mathbf{P} \mathbf{H}^{\text{MW}} \mathbf{P}^{\text{T}}. \quad (35)$$

3. Diagonalize the projected, mass-weighted Hessian matrix.

The effect of the projection leaves the eigenvectors of the Hessian corresponding to translations and rotations with exactly zero eigen values (or at least zero within numerical round-off error), allowing the $3N - 6$ nonzero vibrations to be separated out. These are converted to cm^{-1} prior to print out, and the corresponding eigenvectors are mass-weighted and normalized to become the normal modes. The mass-weighting uses isotopically averaged atomic masses based on the average isotopic abundance for each atom.

At first sight it might be surprising that the Hessian matrix, which after all in *ab initio* molecular orbital theory is inherently quantum mechanical, is amenable to a purely classical treatment. This is because the Born–Oppenheimer approximation (Born and Oppenheimer 1927; Wikipedia 2010) allows for a pretty good separation of the electronic and nuclear motion, allowing the latter to be treated classically. A quantum mechanical description of the simple harmonic oscillator leads to quantized energy levels given by

$$E_n = h\nu \left(n + \frac{1}{2} \right); \quad \nu = \frac{1}{2\pi} \sqrt{\frac{k}{m}}, \quad (36)$$

where the fundamental vibrational frequency, ν , is exactly the same as in a classical description (Eq. 30). The quantum number, n , takes integer values $0, 1, 2, \dots$ and so even in the vibrational ground-state ($n = 0$) there is a residual energy, $\frac{1}{2}h\nu$, known as the zero-point energy. This is a fundamental quantum mechanical property of all oscillators at the absolute zero of temperature (Einstein and Stern 1913).

A molecular vibration is excited when the system absorbs a quantum of energy, $h\nu$, corresponding to the vibrational frequency, ν , of one of its normal modes. This is commonly accomplished under standard conditions by interaction with infrared (IR) radiation. An IR spectrum can be recorded by passing a beam of infrared light covering a range of frequencies through the sample. This can be done by using a monochromatic beam which changes in wavelength over time, or by using a Fourier Transform instrument which effectively measures all wavelengths at once. From this a transmission (or absorbance) spectrum can be produced showing at which frequencies/wavelengths the sample absorbs.

Not all normal modes are IR active, i.e., result in the absorption of energy at the normal mode frequency. In order for a normal mode to be IR active, its motion must result in a change in the dipole moment of the vibrating molecule. For example, the totally symmetric C–H stretch in methane (in which all four C–H bonds are vibrating in the same way simultaneously) does not change the molecular dipole moment (which is zero in any case), and so this mode is IR inactive. Change one of the hydrogen atoms for a fluorine, however, and it is a different picture. In general, molecules with a permanent dipole moment in their ground-state geometries give rise to more intense signals in the IR spectrum (i.e., greater absorbance) than molecules in which a dipole moment is induced by the vibration itself.

Vibrational excitation can occur, and usually does, in conjunction with rotational excitations, resulting in vibration-rotation spectra. For straightforward IR vibrational spectra, the simultaneous rotational excitations result in broadening of the vibrational bands. Bands are additionally broadened if two (or more) normal modes have frequencies close to one another, as well as by instrument resolution and other experimental artifacts. Additional bands, other than those caused by vibrational fundamentals, can arise in the spectrum by the absorption of two or more quanta of energy simultaneously, giving rise to combination and overtone bands, and by further excitations from already vibrationally excited states which give rise to what are called hot bands.

It is worth noting at this point that although our analysis of vibrational motion derives from consideration of the simple harmonic oscillator, many common features observed in vibrational spectra are present precisely because the potential energy surface is not harmonic. For example, transitions between vibrational energy levels in the quantum harmonic oscillator formally allow the quantum number, n , to change only by 1 ($\Delta n = \pm 1$). As the vibrational energy levels are equally spaced (see Eq. 36), this means that light of only one frequency can be absorbed, that of the vibrational fundamental for each normal mode, and there are no overtone bands. The $\Delta n = \pm 1$ selection rule does not apply to an anharmonic oscillator and the observation of overtones is only possible because vibrations are anharmonic. A further consequence of anharmonicity is that the spacing between the vibrational energy levels decreases with increasing quantum number (eventually reaching a continuum allowing for vibrational dissociation) which results in the frequency of the first overtone being slightly less than twice the frequency of the vibrational fundamental.

IR spectroscopy is a very important method in molecular structure determination. Even though the normal modes depend on the whole molecule, certain structural motifs have very similar motions regardless of how many atoms the molecule contains. For example, virtually all molecules containing a carbonyl group give rise to an intense (because of the bond polarity) signal in the IR spectrum around $1,800\text{ cm}^{-1}$ due to excitation of the C = O stretch, thus allowing for the identification of this group. Theory can not only determine the fundamental frequencies and normal modes of a molecule, but by computing the dipole moment derivatives, it can reliably estimate the resulting band intensities as well, enabling the IR spectrum for the vibrational fundamentals to be predicted a priori.

Another technique used in vibrational spectroscopy is Raman scattering (Raman and Krishnan 1928). This normally uses visible light, which is typically scattered by the molecule. Most of the light is scattered elastically at the same frequency as the initial radiation (Rayleigh scattering), but a small amount interacts with the system, excites a vibrational mode, and is scattered inelastically at a lower frequency. If the system is already in a vibrationally excited state, then the scattered radiation can have a higher frequency than the initial radiation. These are known as Stokes and anti-Stokes Raman scattering, respectively. Raman scattering constitutes only a very small component of the total scattered radiation (only about one part in 10^7) and a major difficulty in Raman spectroscopy is to separate out the weak, inelastic

scattering from the intense Rayleigh scattered light. The frequency difference corresponds to, e.g., the frequency of one of the molecules normal modes.

As with IR spectroscopy, not all of a molecules vibrations are active in the Raman spectrum. In order for a vibrational mode to be Raman active, the motion must result in a change in the polarizability of the electron charge cloud surrounding the molecule. The degree of change determines the intensity of the band. Whether or not a band is Raman, or indeed IR, active can be determined in advance by a symmetry analysis of the normal modes. In molecules with no symmetry, all bands are potentially both IR and Raman active, but if the molecule has symmetry then vibrations of certain symmetry types may be inactive in either the Raman or the IR spectrum, or both. For example, in a molecule with a center of symmetry (such as trans-1,2-dichloroethylene; point group C_{2h}), the symmetric modes (a_g and b_g for C_{2h}) are Raman active but IR inactive, and vice versa for the unsymmetric modes. The intensity of a Raman band can be determined theoretically by calculating the molecule's polarizability derivatives.

The polarization of Raman scattered light also contains useful structural information. This property can be measured by using plane polarized light and a polarization analyzer. Spectra recorded with the analyzer set both parallel and perpendicular to the excitation plane can be used to calculate the depolarization ratio of each vibrational mode. This provides insight into molecular orientation and the symmetry of the vibrational modes, as well as information about molecular shape. It is often used to determine macromolecular orientation in crystal lattices, liquid crystals, or polymer samples.

Yet another spectroscopic method that can be used to help determine molecular structure is vibrational circular dichroism (VCD). This technique detects differences in attenuation of left and right circularly polarized light passing through the sample. VCD is sensitive to the mutual orientation of groups of atoms in a molecule and provides three-dimensional structural information. It is especially important in the study of chirality and molecular conformation. Only chiral molecules have a VCD spectrum. In particular, molecules that have either a plane of symmetry or a center of symmetry are VCD inactive.

The differential absorption of polarized light in vibrational circular dichroism is proportional to the rotational strength, a quantity that depends on both the electric and magnetic dipole transition moments. For a vibrational transition from the ground (g) to an excited vibrational state (e), this is given by

$$R(g \rightarrow e) = \text{Im}(\langle g | \mu_{\text{el}} | e \rangle \langle e | \mu_{\text{mag}} | g \rangle), \quad (37)$$

where μ_{el} and μ_{mag} are the electric and magnetic dipole operators, respectively. These quantities can be calculated theoretically, enabling VCD intensities to be computed. For details see the review by Stephens and Lowe (1985). Modern computational approaches use so-called gauge-invariant atomic orbitals (GIAOs) for the magnetic component (Bak et al. 1995) (as does *PQS*) which are particularly efficient, and the computation of VCD rotational strengths as part of a vibrational analysis is now almost routine.

IR, Raman, and VCD spectroscopy all excite the same vibrational fundamentals. The respective vibrational spectra are different because the mechanism by which light is absorbed is different in each case, the amount of absorption depending on changes in the dipole moment, the polarizability and the rotational strength, respectively. All these quantities are amenable to computation, and modern *ab initio* theory can reliably predict the frequencies of a molecule's vibrational fundamentals as well as the intensity of the signal in IR, Raman, and VCD spectra. The latter in particular are greatly assisted by electronic structure calculations; comparison of experimental and theoretical VCD spectra enable absolute molecular conformations to be determined.

1,2-Dichloroethane: An Illustrative Example

Figure 6 shows the *PQS* output for a full vibrational and thermodynamic analysis of *trans (anti)* 1,2-dichloroethane following a geometry optimization under C_{2h} symmetry at the B3LYP/6-31G* level of theory. It includes IR and Raman intensities obtained from the dipole moment and polarizability derivatives, respectively. (Because of its symmetry the system is not VCD active.) In *PQS*, dipole moment derivatives are computed analytically along with the Hessian matrix and atomic axial tensors (for VCD rotational strengths) (Bak et al. 1995) are computed in the same module that calculates NMR chemical shifts. Polarizability derivatives are obtained numerically by finite difference on the force (gradient) in the presence of a series of external electric fields. The number of applied fields varies with the molecular symmetry but requires a maximum of 12 separate single-point energy plus gradient calculations in the worst case (when the system is C_1).

```
Vibrational Frequency & Thermodynamics Module

Hessian read from file C2H4Cl2.hess

Translations and Rotations Projected Out of Hessian

Vibrational Frequencies in atomic units
0.000000    0.000000    0.000000    0.000000    0.000000    0.000000
0.023226    0.041718    0.057471    0.137739    0.146736    0.153218
0.199410    0.207432    0.225626    0.248960    0.255115    0.264838
0.295409    0.295559    0.606973    0.608510    0.617940    0.622129

There are 4 real representations:  ag  bg  au  bu
```

Fig. 6 (continued)

** VIBRATIONAL FREQUENCIES (CM**-1) AND NORMAL MODES **

Label:	1			2			3		
Symmetry:	au			bu			ag		
Frequency:	119.40			214.45			295.43		
IR Active:	YES			YES			NO		
IR Inten:	7.534			10.811			0.000		
dmux/dQ:	0.00000			-0.10225			0.00000		
dmuy/dQ:	0.00000			0.02513			0.00000		
dmuz/dQ:	0.08790			0.00000			0.00000		
Raman Active:	NO			NO			YES		
Raman Inten:	0.000			0.000			5.558		
Depolar:	0.000			0.000			0.384		
Rot.Strength:	0.654			0.635			0.000		
	X	Y	Z	X	Y	Z	X	Y	Z
c	0.000	0.000	0.301	-0.357	-0.019	0.000	0.023	0.223	0.000
cl	0.000	0.000	-0.125	0.144	0.001	0.000	-0.194	0.426	0.000
h	0.043	0.191	0.399	-0.409	0.090	0.011	0.050	-0.336	0.004
h	0.043	0.191	0.399	-0.409	0.090	0.011	-0.050	0.336	-0.004
h	-0.043	-0.191	0.399	-0.409	0.090	-0.011	-0.050	0.336	0.004
c	0.000	0.000	0.301	-0.357	-0.019	0.000	-0.023	-0.223	0.000
h	-0.043	-0.191	0.399	-0.409	0.090	-0.011	0.050	-0.336	-0.004
cl	0.000	0.000	-0.125	0.144	0.001	0.000	0.194	-0.426	0.000
Label:	4			5			6		
Symmetry:	bu			ag			au		
Frequency:	708.05			754.30			787.62		
IR Active:	YES			NO			YES		
IR Inten:	108.720			0.000			3.168		
dmux/dQ:	-0.06593			0.00000			0.00000		
dmuy/dQ:	0.32733			0.00000			0.00000		
dmuz/dQ:	0.00000			0.00000			0.05700		
Raman Active:	NO			YES			NO		
Raman Inten:	0.000			51.557			0.000		
Depolar:	0.000			0.305			0.000		
Rot.Strength:	-4.191			0.005			-0.313		
	X	Y	Z	X	Y	Z	X	Y	Z
c	-0.017	0.495	0.000	0.086	0.390	0.000	0.000	0.000	-0.065
cl	0.004	-0.187	0.000	-0.027	-0.101	0.000	0.000	0.000	0.009
h	0.031	0.329	-0.031	0.016	-0.403	0.050	0.422	0.123	0.234
h	0.031	0.328	-0.031	-0.017	0.403	-0.050	0.422	0.123	0.234
h	0.031	0.328	0.031	-0.016	0.403	0.050	-0.422	-0.123	0.234
c	-0.017	0.496	0.000	-0.086	-0.390	0.000	0.000	0.000	-0.065
h	0.031	0.328	0.031	0.017	-0.402	-0.050	-0.422	-0.123	0.234
cl	0.004	-0.187	0.000	0.027	0.101	0.000	0.000	0.000	0.009
Label:	7			8			9		

Fig. 6 (continued)

Symmetry:	bg			ag			au		
Frequency:	1025.07			1066.31			1159.83		
IR Active:	NO			NO			YES		
IR Inten:	0.000			0.000			1.739		
dmux/dQ:	0.00000			0.00000			0.00000		
dmuy/dQ:	0.00000			0.00000			0.00000		
dmuz/dQ:	0.00000			0.00000			-0.04223		
Raman Active:	YES			YES			NO		
Raman Inten:	10.492			6.766			0.000		
Depolar:	0.750			0.746			0.000		
Rot.Strength:	0.002			-0.001			2.024		
	X	Y	Z	X	Y	Z	X	Y	Z
c	0.000	0.000	-0.124	-0.371	0.060	0.000	0.000	0.000	0.035
cl	0.000	0.000	0.016	0.006	-0.008	0.000	0.000	0.000	-0.011
h	-0.170	-0.442	-0.135	0.415	-0.087	0.002	0.181	-0.465	-0.017
h	0.170	0.442	0.135	-0.415	0.087	-0.002	0.181	-0.465	-0.017
h	-0.170	-0.442	0.135	-0.414	0.087	0.002	-0.181	0.465	-0.017
c	0.000	0.000	0.124	0.371	-0.060	0.000	0.000	0.000	0.035
h	0.170	0.442	-0.135	0.414	-0.087	-0.002	-0.181	0.465	-0.017
cl	0.000	0.000	-0.016	-0.006	0.008	0.000	0.000	0.000	-0.011
Label:	10			11			12		
Symmetry:	bu			bg			ag		
Frequency:	1279.78			1311.42			1361.40		
IR Active:	YES			NO			NO		
IR Inten:	43.117			0.000			0.000		
dmux/dQ:	-0.01423			0.00000			0.00000		
dmuy/dQ:	-0.20980			0.00000			0.00000		
dmuz/dQ:	0.00000			0.00000			0.00000		
Raman Active:	NO			YES			YES		
Raman Inten:	0.000			11.740			11.310		
Depolar:	0.000			0.750			0.678		
Rot.Strength:	-13.415			-0.004			-0.002		
	X	Y	Z	X	Y	Z	X	Y	Z
c	0.056	-0.053	0.000	0.000	0.000	-0.082	-0.040	0.094	0.000
cl	-0.006	-0.007	0.000	0.000	0.000	-0.004	0.004	0.007	0.000
h	-0.222	0.444	0.008	-0.393	0.300	-0.045	-0.198	0.451	0.046
h	-0.222	0.444	0.008	0.393	-0.300	0.045	0.198	-0.451	-0.046
h	-0.222	0.444	-0.008	-0.393	0.300	0.045	0.198	-0.451	0.046
c	0.056	-0.053	0.000	0.000	0.000	0.082	0.040	-0.094	0.000
h	-0.222	0.444	-0.008	0.393	-0.300	-0.045	-0.198	0.451	-0.046
cl	-0.006	-0.007	0.000	0.000	0.000	0.004	-0.004	-0.007	0.000
Label:	13			14			15		
Symmetry:	ag			bu			ag		

Fig. 6 (continued)

Frequency:	1518.55			1519.32			3120.15		
IR Active:	NO			YES			NO		
IR Inten:	0.000			6.293			0.000		
dmux/dQ:	0.00000			-0.07728			0.00000		
dmuy/dQ:	0.00000			0.02192			0.00000		
dmuz/dQ:	0.00000			0.00000			0.00000		
Raman Active:	YES			NO			YES		
Raman Inten:	20.486			0.000			160.189		
Depolar:	0.743			0.000			0.085		
Rot.Strength:	-0.006			3.532			0.000		
	X	Y	Z	X	Y	Z	X	Y	Z
c	0.064	0.006	0.000	0.056	0.026	0.000	0.042	0.022	0.000
cl	0.000	0.002	0.000	0.001	0.002	0.000	0.000	0.000	0.000
h	0.365	0.155	0.300	-0.346	-0.191	-0.305	0.236	0.140	-0.416
h	-0.366	-0.156	-0.301	-0.345	-0.190	-0.303	-0.236	-0.140	0.416
h	-0.366	-0.156	0.301	-0.345	-0.190	0.303	-0.236	-0.140	-0.416
c	-0.063	-0.006	0.000	0.056	0.026	0.000	-0.042	-0.022	0.000
h	0.365	0.155	-0.300	-0.346	-0.191	0.305	0.236	0.140	0.416
cl	0.000	-0.002	0.000	0.001	0.002	0.000	0.000	0.000	0.000
Label:	16			17			18		
Symmetry:	bu			bg			au		
Frequency:	3128.05			3176.52			3198.06		
IR Active:	YES			NO			YES		
IR Inten:	16.489			0.000			5.976		
dmux/dQ:	0.08192			0.00000			0.00000		
dmuy/dQ:	0.10099			0.00000			0.00000		
dmuz/dQ:	0.00000			0.00000			-0.07828		
Raman Active:	NO			YES			NO		
Raman Inten:	0.000			106.923			0.000		
Depolar:	0.000			0.750			0.000		
Rot.Strength:	0.708			0.000			0.266		
	X	Y	Z	X	Y	Z	X	Y	Z
c	0.040	0.024	0.000	0.000	0.000	0.069	0.000	0.000	-0.069
cl	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
h	-0.238	-0.140	0.415	-0.245	-0.146	0.408	-0.248	-0.143	0.407
h	-0.238	-0.140	0.415	0.245	0.146	-0.408	-0.248	-0.143	0.407
h	-0.238	-0.140	-0.415	-0.245	-0.146	-0.408	0.248	0.143	0.407
c	0.040	0.024	0.000	0.000	0.000	-0.069	0.000	0.000	-0.069
h	-0.238	-0.140	-0.415	0.245	0.146	0.408	0.248	0.143	0.407
cl	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

STANDARD THERMODYNAMIC QUANTITIES AT 298.180 K AND 1.000 ATM

This Molecule has 0 Imaginary Frequencies
 Zero point vibrational energy: 36.802 kcal/mol

Fig. 6 (continued)

```

Atom   1 Element c Has Mass 12.01115
Atom   2 Element cl Has Mass 35.45270
Atom   3 Element h Has Mass 1.00794
Atom   4 Element h Has Mass 1.00794
Atom   5 Element h Has Mass 1.00794
Atom   6 Element c Has Mass 12.01115
Atom   7 Element h Has Mass 1.00794
Atom   8 Element cl Has Mass 35.45270
Molecular Mass: 98.959460 amu
Principal axes and moments of inertia in atomic units:
              1          2          3
Eigenvalues -- 62.16074 1242.14322 1281.43626
X              0.36607 -0.93059 0.00000
Y             -0.93059 -0.36607 0.00000
Z              0.00000 0.00000 1.00000
Rotational Symmetry Number is 2
The Molecule is an Asymmetric Top
Translational Enthalpy: 0.889 kcal/mol
Rotational Enthalpy: 0.889 kcal/mol
Vibrational Enthalpy: 38.101 kcal/mol
gas constant (RT): 0.593 kcal/mol
Translational Entropy: 39.687 cal/mol.K
Rotational Entropy: 24.718 cal/mol.K
Vibrational Entropy: 7.625 cal/mol.K

Total Enthalpy: 40.471 kcal/mol
Total Entropy: 72.030 cal/mol.K

```

Fig. 6 PQS vibrational frequency output for trans-1,2-dichloroethane

The first thing printed out is the name of the file containing the Hessian matrix. In *PQS* a vibrational and thermodynamic analysis is done independently of the Hessian computation enabling a full vibrational analysis of isotopomers to be carried out from the same Hessian matrix simply by changing the atomic masses. Analysis is done on any matrix read from the Hessian file, including partial and/or approximate Hessians (e.g., an updated Hessian left over from a geometry optimization), so care must be taken that the source of the Hessian matrix is known. A vibrational analysis of an approximate Hessian will of course, at best, be only approximate; in particular an analysis based on a geometry optimized at one level of theory and a Hessian matrix computed at another (lower) level of theory is, strictly speaking, invalid.

Then comes a list of the vibrational frequencies in atomic units after the translations and rotations have been projected out. This is simply a printout of the eigen values of the projected mass-weighted Hessian (see Eq. 35); as can be seen, the six eigen values corresponding to the translations and rotations are all zero. The remaining eigen values are all positive, showing that the optimized structure of trans-1,2-dichloroethane is a genuine minimum. As any decent chemist should know, 1,2-dichloroethane has two stable conformations, trans and gauche; the former is the more stable by around $1.5 \text{ kcal mol}^{-1}$ (El Youssoufi et al. 1998a, b). This energy difference is sufficiently small that both conformers will be present, certainly at room temperature, and experimental vibrational spectra of 1,2-dichloroethane, unless recorded at very low temperature, will contain bands from the gauche as well as the more abundant trans conformer. This will be discussed in more detail later in this chapter.

This is followed by a list of all symmetry types possible for the vibrational modes. This is simply a list of the irreducible representations present in the character table for the molecule's symmetry point group; for trans-1,2-dichloroethane, having C_{2h} symmetry, this is a_g, b_g (symmetric motions) and a_u, b_u (antisymmetric).

All $3N - 6$ normal modes (Cartesian displacements for each atom) are then listed, together with as much information for each mode as is available, depending on which, if any, of the dipole derivatives, the polarizability derivatives or the rotational strengths have been computed for the molecule in question. For trans-1,2-dichloroethane all these quantities are available except the latter, so listed for each mode is its symmetry type, its frequency in wavenumbers (cm^{-1}), whether or not the mode is IR active, the corresponding IR intensity, the X, Y and Z dipole derivatives with respect to the motion in each normal mode, whether the mode is Raman active, its corresponding Raman intensity and its Raman depolarization ratio. If the molecule were VCD active, its rotational strength (relative intensity in a VCD spectrum) would also be printed.

As already noted, whether a mode is IR or Raman active can be determined in advance (although its intensity cannot) by group theoretical considerations and depends on its symmetry type. For trans-1,2-dichloroethane (because of its point group symmetry) each mode is *either* IR active *or* Raman active, but not both, and so IR and Raman spectroscopy excite completely different vibrational fundamentals. In the Raman spectrum, only bands of a_g and b_g symmetry will appear. These can be distinguished experimentally by measuring the depolarization ratio; according to theory (Wilson et al. 1955), the depolarization ratio of totally symmetric modes (a_g under C_{2h}) is less than 0.75 (these are called polarized bands) whereas that of all other modes is 0.75 (depolarized bands). (See the depolarization ratios calculated for trans-1,2-dichloroethane in Fig. 5.)

The vibrational analysis is followed by a standard, classical statistical thermodynamic analysis at 298.18 K (25 °C) and 1 atm pressure. (For details, see McQuarrie (2000)). Computed quantities include the principal axes and moments of inertia, the rotational symmetry number and symmetry classification, and the translational, rotational, vibrational, and total enthalpy and entropy, respectively.

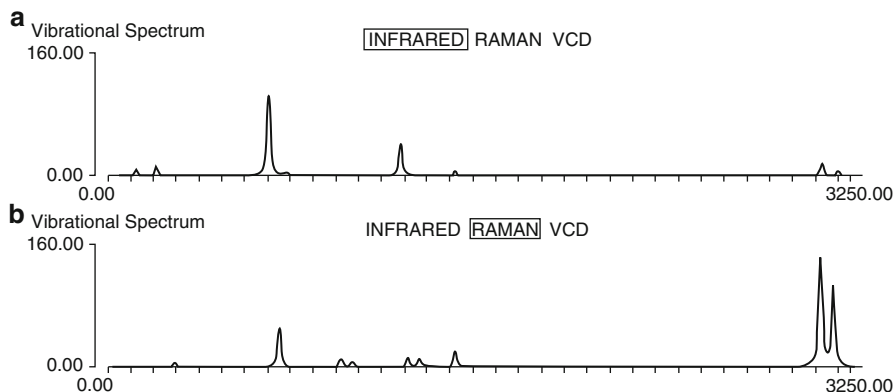


Fig. 7 Simulated IR and Raman spectra using data derived from Fig. 6: (a) IR Spectrum, (b) Raman Spectrum

Both the temperature and pressure can be altered from standard conditions and/or scanned across a requested range of values. The total zero-point energy at 0 K is given by $\sum_v \frac{1}{2} h\nu$, summed over all real frequencies (converted to kcal mol⁻¹; see Eq. 36).

The predicted IR and Raman spectra, based on the data printed out from the vibrational analysis in Fig. 6 are shown in Fig. 7. These are direct “screen captures” from *PQSVIEW*, which is the job output and visualization component of the *PQSMol* GUI (PQS 2010). For each spectrum, the frequency range (horizontal axis), the intensity range (vertical axis), and the band half-width (fitted using a Lorentzian band profile) can be varied; it is also possible to “zoom in” on selected regions of the spectrum.

So how good are the results? First of all it should be noted that we would not expect theory to be able to exactly reproduce experimental vibrational spectra as factors such as intensity and bandwidth depend strongly on experimental conditions (as to some extent do the frequencies themselves). Also, most importantly, the theoretical analysis assumes harmonic behavior, whereas the real spectrum is of course anharmonic. So we would expect the calculated vibrational fundamentals to deviate from experiment simply because they are harmonic, without even considering other sources of error (such as incomplete basis set and approximate treatment of electron correlation). For the level of theory used here, B3LYP/6-31G*, the average (mean) error between computed (harmonic) vibrational fundamentals and experimentally observed (anharmonic) values is around 50 cm⁻¹. For example, in a study of 900 individual vibrational fundamentals from 111 different small molecules using B3LYP and a polarized-valence triple-zeta basis, Schlegel and coworkers (Halls et al. 2001) found an average absolute deviation between theory and experiment of 44.69 cm⁻¹.

Agreement with experiment can be much improved by scaling. This has been done since the early days of Hartree–Fock theory, when all computed HF/6-31G* frequencies were regularly scaled by a factor of around 0.89. Optimized Hartree–Fock bond lengths are typically too short compared with experiment; hence computed frequencies are almost uniformly too high. Single scale factors, applied directly to computed vibrational fundamentals, have since been derived for a range of other wave functions, including post-HF and DFT. A recent paper from Radom and coworkers derived least-squares fitted scale factors for more than 100 different levels of theory (varying both the methodology and the basis set) (Merrick et al. 2007). Scale factors for DFT-based methods are much closer to unity than scale factors derived using Hartree–Fock (which completely neglects electron correlation) showing that the raw (unscaled) results are simply better. The scale factor for B3LYP/6-31G* is 0.9614 (Merrick et al. 2007).

Scaled Quantum Mechanical Force Fields

Although scaling of the computed vibrational frequencies using a single, global, scale factor is common, the first scaling methods applied to ab initio force constants used several different scale factors to correct for systematic errors in different types of molecular deformations, e.g., stretches, bends, or torsions. This procedure requires the transformation of the molecular force field (the Hessian matrix) into chemically meaningful internal coordinates and scales the resulting Hessian elements themselves. It cannot be applied directly to the calculated frequencies. It is thus less convenient than global scaling using a single scaling factor. Largely because of its simplicity, global scaling became popular, but using multiple scale factors yields much better results as was convincingly demonstrated by Blom and Altona in a series of papers starting in the mid-1970s (Blom and Altona 1976, 1977a, b; Blom et al. 1976). Their method forms the basis of the scaled quantum mechanical (SQM) procedure which has been in widespread use for over 25 years (Pulay et al. 1983).

In the original SQM procedure, the molecular geometry was expressed in terms of a full set of nonredundant natural internal coordinates (Fogarasi et al. 1992; Pulay et al. 1979). On the basis of chemical intuition, the natural internal coordinates of all molecules under consideration (there can be more than one) are sorted into groups sharing a common scale factor, and factors for each group are determined by a least-squares fit to experimental vibrational frequencies. Force constants, originally calculated in Cartesian coordinates, are transformed into the internal coordinate representation, and scaling is applied to the elements of the internal force constant matrix (not to the individual vibrational frequencies) according to

$$\mathbf{F}_{ij} \text{ (scaled)} = \sqrt{s_i s_j} \mathbf{F}_{ij}, \quad (38)$$

where s_i and s_j are scaling factors for internal coordinates i and j , respectively.

The accuracy obtained by selective scaling in this way is naturally greater than if just a single global scaling factor were used. Additionally, scaling the force constant matrix also affects the resultant normal modes, and hence the calculated intensities (which are unaffected if only the frequencies are scaled), leading to better agreement with experimental intensities. Furthermore the vibrational frequencies derive directly from the (scaled) Hessian matrix, and are thus fully consistent with it; not the case at all if the frequencies are scaled directly.

The SQM procedure has been widely used in the interpretation of vibrational spectra. A further important role is the development of transferable scale factors which can be used to modify calculated force constants and so predict the vibrational spectrum *a priori*.

The SQM module in *PQS* uses a modified scaling procedure involving the scaling of individual valence coordinates (Baker et al. 1998) (not the linear combinations present in natural internal coordinates). This has immediate advantages in terms of ease of use, as no natural internals need to be generated (a procedure which may fail for complicated molecular topologies), and it simplifies the classification and presorting of the coordinates. In addition, the extra flexibility involved in the scaling of individual primitive internals generally leads to an increase in accuracy and to more transferable scale factors. On typical organic molecules one can expect average differences between predicted SQM and experimental fundamentals of around 10 cm^{-1} with maximum errors of about 30 cm^{-1} or so (Baker and Pulay 1998). In my experience, if differences greatly exceed this maximum, it is more likely that the observed experimental peaks have been misassigned than the predicted SQM frequency is wrong.

1,2-Dichloroethane: A Further Analysis

We are going to apply the SQM method to the existing Hessian matrix for trans-1,2-dichloroethane in order to improve the agreement with experiment; furthermore, we will also need to consider the gauche conformer which will almost certainly contribute to the experimental vibrational spectra at room temperature.

The results of a vibrational analysis of gauche-1,2-dichloroethane, optimized at the B3LYP/6-31G* level, are shown in Fig. 8. This is direct printout from the corresponding *PQS* log file (not the output file). In addition to the full output, *PQS* produces a summary output – the log file – which contains only data that is of direct interest to the user and omits all intermediate printout (such as integral and timing data, intermediate steps in an optimization etc.). As can be seen, the log file reproduces the thermodynamic analysis that is printed in the full output file, but provides only a summary of the vibrational analysis and completely omits the normal modes.

Gauche-1,2-dichloroethane has C_2 symmetry and formally all of its vibrational fundamentals are both IR and Raman active. (Additionally, unlike the trans conformer, it is also potentially VCD active.) However, several modes have only relatively low intensity and may therefore be difficult to see in the experimental

Vibrational frequencies and intensities

No.	Symm.	Freq.	IR	Raman	IR int.	Raman int.	Depol.	Rot.Str.
18	b	3172.85	YES	YES	6.495	21.936	0.7500	1.306
17	a	3159.99	YES	YES	0.688	107.324	0.7427	-1.162
16	a	3103.37	YES	YES	22.911	190.623	0.0225	-3.727
15	b	3094.75	YES	YES	3.074	48.934	0.7500	-2.061
14	a	1500.18	YES	YES	0.172	6.079	0.6749	-0.632
13	b	1495.96	YES	YES	10.565	19.427	0.7500	0.851
12	a	1364.60	YES	YES	25.618	6.217	0.7390	-18.289
11	b	1342.07	YES	YES	49.651	0.508	0.7500	-4.220
10	a	1246.88	YES	YES	1.274	19.900	0.7301	5.497
9	b	1178.97	YES	YES	1.164	7.360	0.7500	-4.560
8	a	1054.50	YES	YES	0.629	3.929	0.7451	2.522
7	a	960.62	YES	YES	13.324	6.693	0.4295	-5.248
6	b	904.47	YES	YES	19.675	1.655	0.7500	16.416
5	b	679.03	YES	YES	30.638	10.525	0.7500	-29.476
4	a	654.31	YES	YES	21.755	13.678	0.1022	19.772
3	b	409.09	YES	YES	9.479	2.353	0.7500	7.781
2	a	261.37	YES	YES	1.042	1.127	0.3706	-4.698
1	a	112.97	YES	YES	1.019	1.793	0.7446	-0.892

STANDARD THERMODYNAMIC QUANTITIES AT 298.180 K AND 1.000 ATM

This Molecule has 0 Imaginary Frequencies

Zero point vibrational energy: 36.734 kcal/mol

Atom	1	Element c	Has Mass	12.01115
Atom	2	Element cl	Has Mass	35.45270
Atom	3	Element h	Has Mass	1.00794
Atom	4	Element h	Has Mass	1.00794
Atom	5	Element h	Has Mass	1.00794
Atom	6	Element c	Has Mass	12.01115
Atom	7	Element h	Has Mass	1.00794
Atom	8	Element cl	Has Mass	35.45270

Molecular Mass: 98.959460 amu

Principal axes and moments of inertia in atomic units:

		1	2	3
Eigenvalues	--	181.49676	849.35170	971.53423
	X	-0.75466	0.00000	0.65612
	Y	0.65612	0.00000	0.75466
	Z	0.00000	1.00000	0.00000

Fig. 8 (continued)

```

Rotational Symmetry Number is 2
The Molecule is an Asymmetric Top
Translational Enthalpy:      0.889 kcal/mol
Rotational Enthalpy:       0.889 kcal/mol
Vibrational Enthalpy:     37.940 kcal/mol
gas constant (RT):        0.593 kcal/mol
Translational Entropy:    39.687 cal/mol.K
Rotational Entropy:      25.130 cal/mol.K
Vibrational Entropy:     6.966 cal/mol.K

Total Enthalpy:          40.310 kcal/mol
Total Entropy:          71.783 cal/mol.K

```

Fig. 8 PQS vibrational frequency summary from the log file for *gauche*-1,2-dichloroethane

Table 5 Standard B3LYP/6-31G* SQM scale factors relevant to 1,2-dichloroethane

Type		Scale factor
Stretch	rC – C	0.9207
Stretch	rC – H	0.9164
Stretch	rC – Cl	1.0438
Bend	∠CCCl	1.0144
Bend	∠CCH; ∠ClCH	0.9431
Bend	∠HCH	0.9016
Torsion	all	0.9523

spectrum, especially given that the *gauche* is the higher energy conformer and is only expected to comprise about 13–25 % of the total at room temperature (El Youssefi et al. 1998a, b). The computed B3LYP/6-31G* energy difference between the *gauche* and *trans* conformers is 1.5 kcal mol⁻¹, including (minor) zero-point energy effects.

In the paper that introduced the scaling of individual primitive internals into the SQM method (Baker et al. 1998), a set of 11 scale factors were derived from a test set of 30 molecules containing C, H, O, N and Cl at the B3LYP/6-31G* level. These scale factors will be used “as is” to scale the raw Hessian data for both *trans* and *gauche*-1,2-dichloroethane. The seven scale factors appropriate for use with 1,2-dichloroethane as shown in Table 5.

There have been several experimental studies of the vibrational spectra of 1,2-dichloroethane. Of note are the 1975 IR and Raman measurements on the solid, liquid, and gaseous *trans* and *gauche* conformers from Mizushima and coworkers (1975), and the combined theoretical and FT-IR study of El Youssefi et al. already referred to (El Youssefi et al. 1998a, b). These results are summarized, together with the unscaled, scaled using a single scale factor (Merrick et al. 2007), and the SQM predicted fundamentals in Table 6.

Table 6 Comparison of experimental and theoretical (B3LYP/6-31G*) vibrational fundamentals for gas-phase 1,2-dichloroethane and their assignments

Experiment		Scaled (SQM)	Scaled (0.9614)	Unscaled	Assignment
A ^a	B ^b				
transconformer					
122 (<i>a_u</i>)	123	117 (<i>a_u</i>)	114 (<i>a_u</i>)	119 (<i>a_u</i>)	Torsion
221 (<i>b_u</i>)	220	214 (<i>b_u</i>)	206 (<i>b_u</i>)	214 (<i>b_u</i>)	CCCl def.
305 (<i>a_g</i>)		296 (<i>a_g</i>)	284 (<i>a_g</i>)	295 (<i>a_g</i>)	CCCl def.
727 (<i>b_u</i>)	728	724 (<i>b_u</i>)	681 (<i>b_u</i>)	708 (<i>b_u</i>)	C–Cl str.
772 (<i>a_g</i>)		766 (<i>a_g</i>)	725 (<i>a_g</i>)	754 (<i>a_g</i>)	C–Cl str.
772 (<i>a_u</i>)	773	767 (<i>a_u</i>)	758 (<i>a_u</i>)	788 (<i>a_u</i>)	CH ₂ rock
996 (<i>b_g</i>)		999 (<i>b_g</i>)	985 (<i>b_g</i>)	1,025 (<i>b_g</i>)	CH ₂ rock
1,057 (<i>a_g</i>)		1,027 (<i>a_g</i>)	1,025 (<i>a_g</i>)	1,066 (<i>a_g</i>)	C–C str.
1,124 (<i>a_u</i>)	1,122	1,128 (<i>a_u</i>)	1,115 (<i>a_u</i>)	1,160 (<i>a_u</i>)	CH ₂ twist
1,233 (<i>b_u</i>)	1,232	1,244 (<i>b_u</i>)	1,231 (<i>b_u</i>)	1,280 (<i>b_u</i>)	CH ₂ wag.
1,267 (<i>b_g</i>)		1,273 (<i>b_g</i>)	1,260 (<i>b_g</i>)	1,311 (<i>b_g</i>)	CH ₂ twist
1,310 (<i>a_g</i>)		1,322 (<i>a_g</i>)	1,308 (<i>a_g</i>)	1,361 (<i>a_g</i>)	CH ₂ wag.
1,431 (<i>a_g</i>)		1,459 (<i>a_g</i>)	1,460 (<i>a_g</i>)	1,519 (<i>a_g</i>)	CH ₂ scissor
1,461 (<i>b_u</i>)	1,460	1,460 (<i>b_u</i>)	1,460 (<i>b_u</i>)	1,519 (<i>b_u</i>)	CH ₂ scissor
2,979 (<i>a_g</i>)		2,987 (<i>a_g</i>)	3,000 (<i>a_g</i>)	3,120 (<i>a_g</i>)	CH ₂ sym. str.
2,983 (<i>b_u</i>)	2,983	2,994 (<i>b_u</i>)	3,007 (<i>b_u</i>)	3,128 (<i>b_u</i>)	CH ₂ sym. str.
3,001 (<i>b_g</i>)		3,041 (<i>b_g</i>)	3,054 (<i>b_g</i>)	3,177 (<i>b_g</i>)	CH ₂ antisym. str.
3,009 (<i>a_u</i>) ^c	3,011	3,062 (<i>a_u</i>)	3,075 (<i>a_u</i>)	3,198 (<i>a_u</i>)	CH ₂ antisym. str.
Mean error^d(cm⁻¹)		9.3 (30)	17.9 (47)	38.8 (145)	
gaucheconformer					
125 (<i>a</i>)		111 (<i>a</i>)	109 (<i>a</i>)	113 (<i>a</i>)	torsion
263 (<i>a</i>)		260 (<i>a</i>)	251 (<i>a</i>)	261 (<i>a</i>)	CCCl def.
411 (<i>b</i>)		408 (<i>b</i>)	393 (<i>b</i>)	409 (<i>b</i>)	CCCl def.
669 (<i>a</i>)	668	663 (<i>a</i>)	629 (<i>a</i>)	654 (<i>a</i>)	C–Cl str.
694 (<i>b</i>)	694	689 (<i>b</i>)	653 (<i>b</i>)	679 (<i>b</i>)	C–Cl str.
891 (<i>b</i>)	891	889 (<i>b</i>)	869 (<i>b</i>)	904 (<i>b</i>)	CH ₂ rock
947 (<i>a</i>)	948	938 (<i>a</i>)	924 (<i>a</i>)	961 (<i>a</i>)	CH ₂ rock
1,028 (<i>a</i>)	1,027	1,021 (<i>a</i>)	1,014 (<i>a</i>)	1,055 (<i>a</i>)	C–C str.
1,146 (<i>b</i>)	1,146	1,146 (<i>b</i>)	1,133 (<i>b</i>)	1,179 (<i>b</i>)	CH ₂ twist
1,214 (<i>a</i>)	1,214	1,211 (<i>a</i>)	1,199 (<i>a</i>)	1,247 (<i>a</i>)	CH ₂ twist
1,292 (<i>b</i>)	1,292	1,303 (<i>b</i>)	1,290 (<i>b</i>)	1,342 (<i>b</i>)	CH ₂ wag.
1,315 (<i>a</i>)	1,313	1,325 (<i>a</i>)	1,312 (<i>a</i>)	1,365 (<i>a</i>)	CH ₂ wag.
1,436 (<i>b</i>)	1,437	1,438 (<i>b</i>)	1,438 (<i>b</i>)	1,496 (<i>b</i>)	CH ₂ scissor
1,446 (<i>a</i>)		1,443 (<i>a</i>)	1,442 (<i>a</i>)	1,500 (<i>a</i>)	CH ₂ scissor
2,957 (<i>b</i>)	2,957	2,971 (<i>b</i>)	2,976 (<i>b</i>)	3,095 (<i>b</i>)	CH ₂ sym. str.
2,966 (<i>a</i>)	2,963	2,971 (<i>a</i>)	2,983 (<i>a</i>)	3,103 (<i>a</i>)	CH ₂ sym. str.
3,005 (<i>a</i>)		3,025 (<i>a</i>)	3,038 (<i>a</i>)	3,160 (<i>a</i>)	CH ₂ antisym. str.
3,005 (<i>b</i>)	3,040	3,037 (<i>b</i>)	3,051 (<i>b</i>)	3,173 (<i>b</i>)	CH ₂ antisym. str.
Mean error^d(cm⁻¹)	5.4 (14)	16.1 (41)	40.9 (138)		

^aRef. Mizushima et al. (1975)^bRef. El Youssofi et al. (1998a, b)^cReported as 3,040 cm⁻¹ in the solid^dOmitting the two highest fundamentals, maximum deviation in parentheses (see text)

The first thing to note from Table 6 is that the order and the symmetry assignments of the fundamentals is predicted to be exactly the same theoretically as found experimentally. The actual assignments as to the type of motion for each vibrational mode can be readily checked by animating the mode and observing it visually. In this regard, theory is far easier to interpret than experiment. When you determine a vibrational fundamental computationally you automatically get its normal mode, its symmetry type, its IR and Raman intensities, its depolarization ratio, indeed the whole vibrational analysis, so there is absolutely no dispute as to the bands assignment. Additionally, it is possible to do an energy distribution analysis (Pulay and Torok 1966) of the normal modes which determines which primitive internals have more than a given weight (typically 3 %) to each mode, often enabling the principal motion in the mode (assuming there is one) to be found without having to animate it.

The two sets of experimental fundamentals are for the most part in excellent agreement with each other and typically differ by 1–2 cm^{-1} only. However, there are more serious discrepancies for the two highest frequencies. These are C–H stretches which, primarily due to tunneling, often show large differences from harmonic behavior, and both the position and the intensity of these bands in the experimental spectrum can deviate significantly from theoretical predictions. They are also often difficult to determine experimentally. In this case the highest frequency for the *gauche* conformer differs by 35 cm^{-1} between the two sets of experimental values: 3,005 and 3,040 cm^{-1} . Additionally, although the agreement in this region for the *trans* conformer looks fine, Mizushima and coworkers (1975) report a value for the a_u C–H antisymmetric stretch in the solid of 3,040 cm^{-1} , compared to 3,009 cm^{-1} in the gas phase, a substantial difference. Because of these experimental differences, we have omitted the two highest frequency C–H stretches from our error analysis, and the mean average (unsigned) and maximum errors between theory and experiment is reported over 16 vibrational fundamentals (not 18).

The largest errors are of course found for the raw, unscaled frequencies, which show average differences from experiment for the two conformers of around 40 cm^{-1} (and maximum errors of about 140 cm^{-1}). These are in line with the average errors reported by Schlegel (Halls et al. 2001) previously mentioned. Despite the large average differences, even the unscaled B3LYP/6-31G* frequencies are helpful in assigning the experimental spectrum as the order is the same as that found experimentally, even if the positions are off.

Scaling even using just a single scaling factor significantly improves the comparison with experiment, reducing the average error by well over 50 %. However, the SQM results are by far the best and are genuinely predictive. Results for the *trans* conformer are typical for this approach (average error 9.3 cm^{-1} , maximum error 30 cm^{-1}) and those for the *gauche* are even better (average error 5.4 cm^{-1} , maximum error 14 cm^{-1}). In the latter case, they are three times better than using a single scaling factor, both on the average and for the maximum deviation.

A recently published SQM reappraisal of the vibrational spectrum of toluene (Baker 2008) showed a similar average deviation from the 2005 experimental study of Keefe and coworkers (Bertie et al. 2005) of 5.3 cm^{-1} (maximum error 17 cm^{-1})

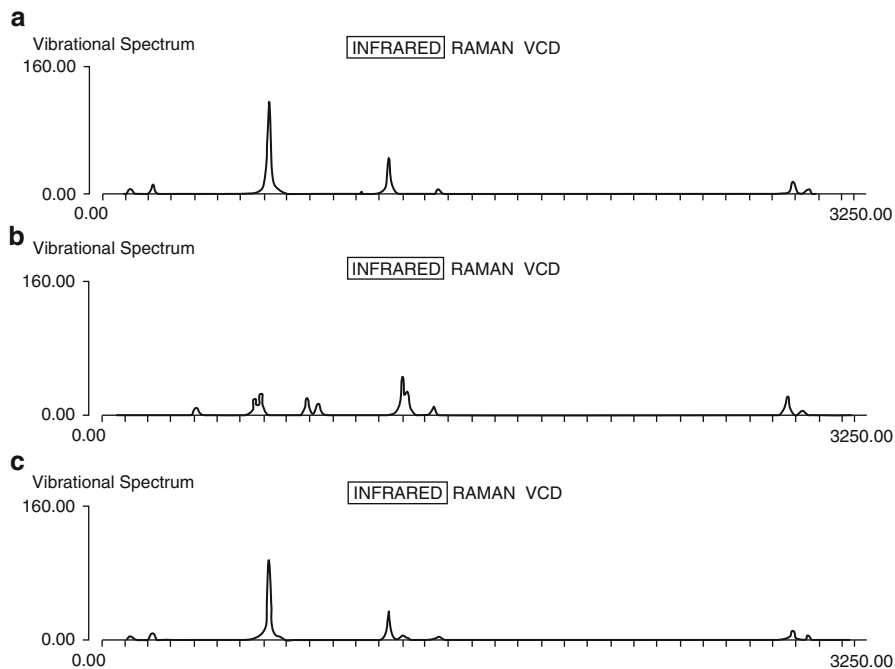


Fig. 9 Predicted SQM IR spectra for 1,2-difluoroethane: (a) IR spectrum for *trans*-1,2-dichloroethane, (b) IR spectrum for *gauche*-1,2-dichloroethane, (c) IR spectrum for 1,2-dichloroethane at room temperature

and convincingly proposed that three observed bands either unassigned or assigned as combination bands were in fact fundamentals. In their 1998 SQM study of the vibrational spectra of several small fluorocarbons, Baker and Pulay (1998) proposed several reassignments to previously published experimental spectra and predicted the fundamental frequencies and IR intensities of nine other fluorocarbons for which there was, at that time, no experimental data. A subsequent jet-cooled FT-IR investigation by McNaughton and coworkers (Jiang et al. 2002) essentially confirmed all of the theoretical predictions.

The final SQM-predicted IR and Raman Spectra for 1,2-dichloroethane are shown in Fig. 9 and 10, respectively. In each case individual spectra for the *trans* and *gauche* conformers are given followed by a combined spectrum assuming a room-temperature distribution of 85 % *trans* and 15 % *gauche*.

Not surprisingly, because of its dominance at room temperature, the observed IR spectrum of 1,2-dichloroethane is predicted to look very much like that for the *trans*-conformer alone. However there are some peaks due to the *gauche* form that should be visible in room-temperature spectra and their intensity should increase at higher temperature. In particular, the CH₂ wag. at $\sim 1,300\text{ cm}^{-1}$ should have sufficient intensity to be seen in the IR spectrum and the CH₂ twist at $1,211\text{ cm}^{-1}$ should be visible in the Raman spectrum.

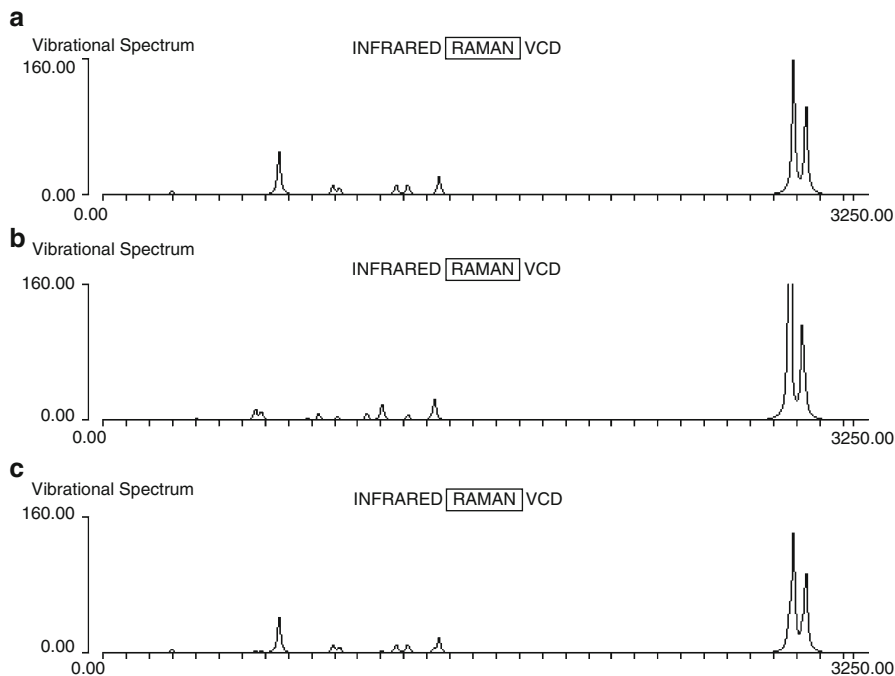


Fig. 10 Predicted SQM Raman spectra for 1,2-dichloroethane: (a) Raman spectrum for *trans*-1,2-dichloroethane, (b) Raman spectrum for *gauche*-1,2-dichloroethane, (c) Raman spectrum for 1,2-dichloroethane at room temperature

In addition to their results for normal 1,2-dichloroethane, Mizushima et al. also reported the experimental vibrational fundamentals for the fully deuterated isotopomer ($\text{CD}_2\text{ClCD}_2\text{Cl}$) (Mizushima et al. 1975). These are compared with the SQM scaled frequencies (using the same scaling factors as for $\text{CH}_2\text{ClCH}_2\text{Cl}$) in Table 7. (As the Hessian matrix is independent of mass, the vibrational fundamentals for all possible isotopomers can be derived from a single Hessian matrix simply by changing the mass weighting.) The agreement is again excellent, with average errors (mean average deviation) of just over 7 cm^{-1} and maximum errors around 20 cm^{-1} for both deuterated conformers. There is one frequency for the *gauche* conformer where there is significant disagreement between theory and experiment (see Table 7); this is a *b* mode observed experimentally at 707 cm^{-1} but predicted theoretically at 752 cm^{-1} . In Table 3 in Mizushima et al. (1975) this frequency is in parentheses (the only one to be thus treated) which suggests that its value is uncertain. The *trans* conformer shows an intense band at precisely this frequency (707 cm^{-1}) which is predicted to be the most intense band in the entire IR spectrum of deuterated 1,2-dichloroethane, and it may well be that a *gauche* mode was assumed to lie hidden under this intense band. The prediction here is that this is not the case, and there is an unassigned feature in the IR spectrum around 752 cm^{-1}

Table 7 Comparison of experimental and theoretical (B3LYP/6-31G*) vibrational fundamentals for fully deuterated 1,2-dichloroethane and their assignments

Experiment ^a	SQM	Assignment	Experiment ^a	SQM	Assignment
<i>trans</i> conformer			<i>gauche</i> conformer		
128 (<i>a_u</i>)	107 (<i>a_u</i>)	torsion	120 (<i>a</i>)	106 (<i>a</i>)	torsion
217 (<i>b_u</i>)	201 (<i>b_u</i>)	CCCl def.	229 (<i>a</i>)	224 (<i>a</i>)	CCCl def.
299 (<i>a_g</i>)	292 (<i>a_g</i>)	CCCl def.	365 (<i>b</i>)	359 (<i>b</i>)	CCCl def.
568 (<i>a_u</i>)	566 (<i>a_u</i>)	CD ₂ rock	624 (<i>a</i>)	616 (<i>a</i>)	C–Cl str.
707 (<i>b_u</i>)	701 (<i>b_u</i>)	C–Cl str.	643 (<i>b</i>)	620 (<i>b</i>)	C–Cl str.
708 (<i>a_g</i>)	709 (<i>a_g</i>)	C–Cl str.	(707) (<i>b</i>)	752 (<i>b</i>)	CD ₂ rock
768 (<i>b_g</i>)	770 (<i>b_g</i>)	CD ₂ rock	759 (<i>a</i>)	766 (<i>a</i>)	CD ₂ rock
817 (<i>a_u</i>)	815 (<i>a_u</i>)	CD ₂ twist	828 (<i>a</i>)	822 (<i>a</i>)	CD ₂ twist
909 (<i>a_g</i>)	895 (<i>a_g</i>)	C–C str.	828 (<i>b</i>)	826 (<i>b</i>)	C–C str.
944 (<i>b_u</i>)	946 (<i>b_u</i>)	CD ₂ wag.	941 (<i>a</i>)	935 (<i>a</i>)	CD ₂ twist
975 (<i>b_g</i>)	977 (<i>b_g</i>)	CD ₂ twist	1,022 (<i>b</i>)	1,022 (<i>b</i>)	CD ₂ wag.
1,036 (<i>a_g</i>)	1,046 (<i>a_g</i>)	CD ₂ wag.	1,052 (<i>a</i>)	1,053 (<i>a</i>)	CD ₂ wag.
1,085 (<i>b_u</i>)	1,078 (<i>b_u</i>)	CD ₂ sissor	1,067 (<i>b</i>)	1,060 (<i>b</i>)	CD ₂ sissor
1,160 (<i>a_g</i>)	1,146 (<i>a_g</i>)	CD ₂ sissor	1,153 (<i>a</i>)	1,135 (<i>a</i>)	CD ₂ sissor
2,179 (<i>a_g</i>)	2,180 (<i>a_g</i>)	CD ₂ sym. str.	2,171 (<i>b</i>)	2,159 (<i>b</i>)	CD ₂ sym. str.
2,185 (<i>b_u</i>)	2,180 (<i>b_u</i>)	CD ₂ sym. str.	2,171 (<i>a</i>)	2,168 (<i>a</i>)	CD ₂ sym. str.
2,264 (<i>b_g</i>)	2,271 (<i>b_g</i>)	CD ₂ antisym. str.	2,265 (<i>a</i>)	2,256 (<i>a</i>)	CD ₂ antisym. str.
2,271 (<i>a_u</i>)	2,282 (<i>a_u</i>)	CD ₂ antisym. str.	2,265 (<i>b</i>)	2,261 (<i>b</i>)	CD ₂ antisym. str.
MAD (cm⁻¹)	7.3 (21)		7.7 (23)^b		

^aFrom Ref. Mizushima et al. (1975); gas-phase where given, otherwise liquid

^bOmitting mode in parentheses (707)

which should be assigned as a *gauche* fundamental. Because of the possible experimental misinterpretation, this mode was omitted from the error analysis.

As noted, *gauche*-1,2-dichloroethane is potentially VCD active. The two chiral forms of this molecule are shown in Fig. 11 together with their simulated VCD spectra. (They are denoted by *gauche*1 and *gauche*2, respectively; the rotational strengths given in Fig. 8 correspond to *gauche*1.) The VCD spectra for the two chiral conformers are precisely the opposite of each other; where one form has a positive rotational strength, the other has exactly the same magnitude but with a negative sign. In any real mixture of *gauche*-1,2-dichloroethane, however, both forms will be present in equal amounts, i.e., it will be a racemic mixture, and so overall there will be no observable VCD spectrum as everything will cancel. However, this clearly shows the potential of VCD spectroscopy to determine the absolute molecular configuration of a pure, chiral conformer.

Finally Table 8 compares optimized geometrical parameters for *trans*-1,2-dichloroethane with experimental electron diffraction data from Kvesethi as reported in El Youssoufi et al. (1998a, b).

I hope that the above has demonstrated just how useful modern theory can be for the prediction and interpretation of molecular vibrational spectra and the

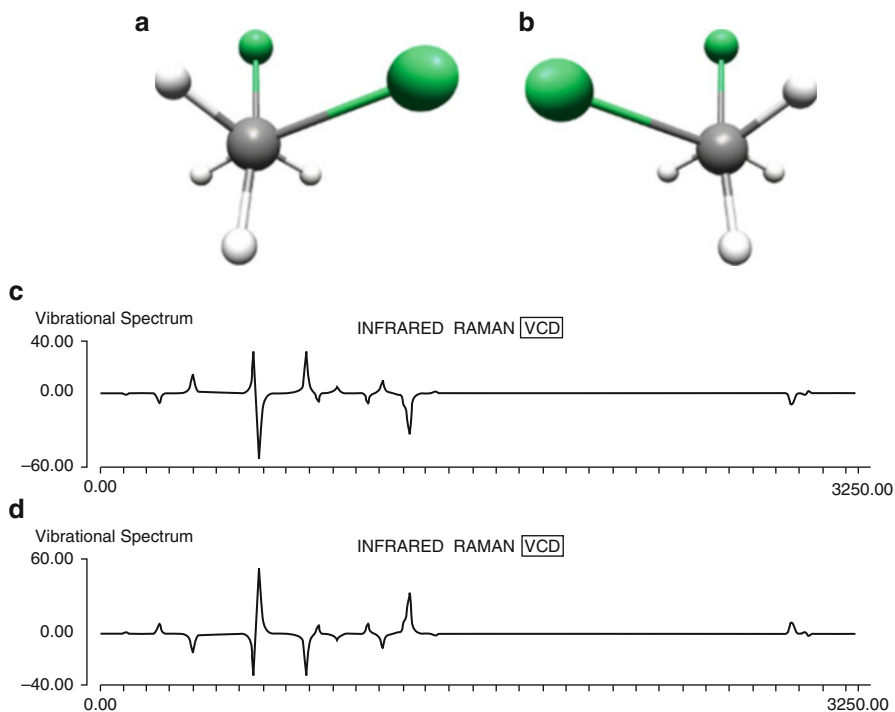


Fig. 11 Predicted VCD spectra for the two chiral forms of *gauche*-1,2-dichloroethane: (a) *gauche*1, (b) *gauche*2, (c) simulated VCD spectrum of *gauche*1, (d) simulated VCD spectrum of *gauche*2

identification and assignment of vibrational fundamentals. I have attempted to show this with a fairly extensive theoretical study (carried out especially for this chapter) on one particular molecule, 1,2-dichloroethane, rather than merely summarizing the results of previous studies.

Density Functional Theory and Weight Derivatives

Before concluding, I would like to comment on the use of weight derivatives in DFT calculations. As will be discussed, this has important ramifications for the reliable computation of DFT vibrational frequencies. Following Pople and coworkers (Johnson et al. 1993), we can write the DFT exchange-correlation energy as

$$E_{XC} = \int f(\rho) dr, \quad (39)$$

Table 8 Comparison of experimental and theoretical (B3LYP/6-31G^{*}) geometrical parameters for *trans*-1,2-dichloroethane

Parameter ^a	B3LYP/6-31G [*]	Experiment ^b
rC – C	1.518	1.528(6)
rC – Cl	1.815	1.796(3)
rC – H	1.090	1.120(10)
∠CCCl	109.2	108.9(0.3)
∠CCH	111.6	113.0(1.3)
τClCCCl	180.0	180.0

^aDistances in Å and angles in degrees

^bFrom Ref. Kveseth (1974), standard deviation in parentheses

where ρ represents all the density dependence, of whatever form, in the density functional $f(\rho)$. This integral is too complicated to be evaluated analytically and numerical integration has to be used. Virtually all DFT programs employ an atomic partitioning scheme pioneered by Becke (1988) to evaluate this integral, which separates the molecular integral into atomic contributions which may then be individually treated using single-center techniques.

If this is done, we can replace Eq. 39 by

$$E_{XC} = \sum_A \sum_i \omega_{Ai} f(\rho; \mathbf{r}_{Ai}), \quad (40)$$

where the first summation is over all atoms, A , and the second is over the numerical quadrature grid points, i , for the current atom. The ω_{Ai} are quadrature weights and the grid points, \mathbf{r}_{Ai} are given by $\mathbf{r}_{Ai} = \mathbf{R}_A + \mathbf{r}_i$, where \mathbf{R}_A is the position of nucleus A and \mathbf{r}_i represents a suitable quadrature grid centered on A .

Because the grid points are atom-centered they are not fixed in space, but “move with” the atom. This means that the derivative of Eq. 40 (needed for the DFT contribution to the gradient) formally has two parts:

$$\nabla_A E_{XC} = \sum_B \sum_i [\omega_{Bi} \nabla_A f(\rho; \mathbf{r}_{Bi}) + (\nabla_A \omega_{Bi}) f(\rho; \mathbf{r}_{Bi})]. \quad (41)$$

The first term in Eq. 41 looks like the numerical integral of the exchange-correlation contribution to the total energy gradient, and is the only term in a fixed grid, but the second term is formally needed and the implementation in Johnson et al. (1993) was the first time that this term – the weight derivative – was properly included.

The error in neglecting the weight-derivative term can be reduced to virtually zero by taking a large enough number of grid points, but for smaller grids its omission is not justified (Johnson et al. 1993). I suspect that the hope was that if the second term in Eq. 41 were properly included, then small grids would no longer give unreliable results, but if this was indeed the hope, it was quickly dashed when it was shown that small grids which included weight derivatives often gave worse results than calculations in which they were omitted (Baker et al. 1994). Basically you have to have a decent quadrature grid to get reliable gradients; if the grid is not

Table 9 Optimized BLYP/6-31G* geometrical parameters for hydrogen peroxide using numerical quadrature grids of decreasing quality with and without weight derivatives^a

Grid quality ^b	Weight derivs	Geometrical parameters ^c			
		rO – O	rO – H	∠HOO	τHOOH
Conv.	Yes	1.494	0.985	98.4	121.9
1	Yes	1.494	0.985	98.4	122.0
1	No	1.494	0.985	98.4	121.9
2	Yes	1.494	0.985	98.4	122.0
2	No	1.494	0.985	98.4	121.9
3	Yes	1.495	0.985	98.4	122.8
3	No	1.493	0.986	98.6	118.3
4	Yes	1.496	0.985	98.3	126.2
4	No	1.492	0.986	98.7	113.9
5	Yes	1.502	0.983	97.7	142.8
5	No	1.493	0.986	98.4	114.3

^aFrom Ref. Baker et al. (1994)^b1 = best grid, 5 = worst; Conv. represents converged results^cIn Å and degrees

good enough using weight derivatives will not save you, if it is good enough then you probably do not need weight derivatives. At least this is typically the case for the gradient.

One clear advantage of including weight derivatives is that the gradient zero coincides essentially exactly (for minimization) with the energy minimum. If the weight derivative term is omitted and grid quality is poor then there is a strong tendency for the energy to rise near the end of a geometry optimization as the gradient supposedly decreases. This was a common occurrence with many early DFT codes almost entirely due to sloppy numerical integration.

The kind of errors in geometries seen when the DFT integration grid gets increasingly worse is shown in Table 9 which gives geometrical parameters for BLYP/6-31G* optimizations of hydrogen peroxide with a range of decreasing quality grids. The table is taken from Baker et al. (1994) for which see for more details. Results are shown with and without inclusion of the weight derivative term. As can be seen from Table 9, the HOOH torsion changes substantially with grid quality, as do some of the other parameters with the poorest quality grid. If anything, results with weight derivatives are worse than those without as grid quality worsens. The actual optimization for the poorest quality grid (grid 5) converged smoothly when weight derivatives were included, with the energy decreasing on every optimization cycle; it was only when examining the final geometry that one realizes that something is badly wrong.

The situation as far as the first energy derivative (the gradient) is concerned is that for reliable results you need a good quality grid, but if you have that then you don't really need to include weight derivatives (although there is no harm in doing so of course). For the second energy derivative (the Hessian matrix and hence vibrational frequencies) things are somewhat different. As convincingly

demonstrated by Malagoli and Baker (2003), weight derivatives are essentially *de rigueur* for the reliable computation of vibrational frequencies, particularly if heavier elements (beyond the first row) are involved. However, the reason for this is not what you might think.

It turns out that there is a potential, major difficulty in integrating derivatives of basis functions that represent core electrons on their own atomic grid. Such basis functions have large exponents and their contribution to the electron density consequently dies off very rapidly with increasing distance from the nucleus. Exponents of the order of a million or more are not at all uncommon, especially if the core – as is often the case – is represented by a contracted function, and such exponents only increase in magnitude with increasing atomic number. In order to accurately integrate a Gaussian function that dies off so rapidly a high density of points are needed near the nucleus. The numerical inaccuracy does not matter very much at all for the exchange-correlation energy, but for the gradient and especially for the second derivative, then a basis function $\exp(-\omega r^2)$ which has a small value at some grid point near the nucleus, has a magnitude $2\omega \exp(-\omega r^2)$ for its gradient and $4\omega^2 \exp(-\omega r^2)$ for its second derivative. With ω of the order of 10^6 , then the contribution of this function to the second derivative has a magnitude four trillion r^2 times its contribution to the energy at that particular grid point, i.e., it has changed from something fairly small to something potentially large, as well as having a greater radial extent. Consequently it requires a far more accurate numerical integration than it previously had.

It is possible to ameliorate this inaccuracy by increasing the grid quality, but eventually with a large enough exponent, you will simply be swamped and it will not be possible to accurately integrate the function on any viable radial grid. The tremendous advantage of including the weight derivative term (as opposed to ignoring it and in effect assuming a fixed grid) is that, when considering any derivative term, because a grid “moves with” its atom there is no need to take the derivative of any basis function centered on that atom. Thus the highly inaccurate integration of basis functions with large exponents on their own atomic center no longer needs to be carried out. In other words, the inclusion of weight derivatives is so important for reliable frequencies with DFT, not because of anything inherent in the weight derivative term itself, but because, by including them, one can completely avoid a highly inaccurate numerical integration.

If you are using DFT code which does *not* have weight derivatives, then consider carrying out any vibrational analysis with a Hessian matrix computed numerically. As noted above, the gradient is normally reliable even without weight derivatives provided the quadrature grid is of sufficient quality. The calculation will almost certainly take longer compared to computing a fully analytical Hessian, but at least it should give reasonable results.

Table 10 shows the effects of not including weight derivatives in analytical Hessian matrices on the computed B3LYP/m6-31G* harmonic vibrational frequencies for Fe(CO)₅. (The m6-31G* basis (Mitin et al. 2003) is a modified version of the 6-31G* basis for first-row transition metals (Rassalov et al. 1998) that corrects deficiencies in the latter due to the lack of a sufficiently diffuse *d*-function in the

Table 10 Computed B3LYP/m6-31G* harmonic vibrational frequencies (cm^{-1}) for $\text{Fe}(\text{CO})_5$ ^a

Vib. mode Symmetry	Standard grid		Better grid		Numerical
	No weight	Weight	No weight	Weight	Hessian
e'	150.09i	47.44	46.55	47.38	46.66
e''	100.44	98.96	98.46	98.81	98.99
e'	52.52	102.00	101.72	101.79	101.73
a_2''	122.60i	109.13	108.75	108.93	109.13
a_2'	349.78	350.44	349.78	350.33	350.20
e''	360.98	361.49	360.89	361.25	361.05
a_1'	419.73	419.44	419.74	419.41	419.35
e'	417.52	428.61	428.40	428.50	428.87
a_1'	441.88	441.57	441.85	441.57	441.46
a_2''	440.85	471.81	471.77	471.81	471.58
e'	463.61	484.05	484.22	484.01	484.18
e''	561.95	562.26	561.97	662.14	562.06
a_2''	579.54	620.21	620.05	620.22	620.49
e'	627.12	658.12	658.07	658.07	658.18
e'	2,089.44	2,089.55	2,089.59	2,089.40	2,090.10
a_1'	2,115.53	2,115.43	2,115.69	2,115.43	2,115.98
a_2''	2,116.03	2,116.09	2,116.20	2,116.20	2,116.65
a_1'	2,187.58	2,187.40	2,187.59	2,187.31	2,187.96

^aFrom Malagoli and Baker (2003)

3 *d*-shell.) This Table, taken from Malagoli and Baker (2003), shows analytical frequencies computed both with and without weight derivatives using the standard grid in *PQS* (which is perfectly adequate for DFT energies and gradients) as well as a much better quality grid; additionally it gives numerical frequencies derived from a Hessian obtained via central differences on the (analytical) gradient using the standard grid.

As can be seen, all computed frequencies are in excellent agreement with one another (all agreeing within 1 cm^{-1}) except for those computed with the standard grid without weight derivatives. These results are terrible. Apart from the high frequency C–O stretches (above $2,000 \text{ cm}^{-1}$), all the e' modes show significant errors, as do the a_2'' . The low frequencies are completely wrong, the order of the fundamentals is incorrect, and there are two imaginary frequencies even though the geometry was minimized. All this is solely due to integration error.

Conclusions

Modern ab initio theory can compute a large number of molecular properties, particularly those for individual molecules – such as geometries and spectroscopic data – that are of direct relevance for experimental studies. This chapter has

concentrated on molecular structure and vibrational spectroscopy where theory can be tremendously helpful to experimentalists. Theoretical data – for example, bond lengths involving hydrogen – are now regularly used as an aid when fitting X-ray and microwave data in order to extract geometrical parameters. IR, Raman and VCD spectra can readily be simulated theoretically, with reliable estimates of both band positions and intensities, especially if the raw data is scaled via an SQM treatment. I have tried to illustrate the latter in particular by concentrating for the most part on an extensive study of a single system using a standard level of theory (B3LYP/6-31G^{*}) rather than attempting to summarize results on a large number of molecules using a variety of theoretical methods.

Thirty years ago I was attending scientific meetings and conferences in which the theoretical chemists in the audience were often attempting to justify their existence to their experimental colleagues. At that time calculations were not especially helpful to the experimentalist. The standard level of theory for most applications was Hartree–Fock, with perhaps some higher level post-HF single-point energies to hopefully improve the energetics. Computed geometries were not particularly good, computed frequencies were typically much too high (indeed, the paper introducing one of the first useable analytical Hartree–Fock second derivative codes (Pople et al. 1979) had only just been published), reaction energetics were fairly poor and barrier heights were simply not reliable. All that has changed. Today the experimentalist who refuses to consider any input from theory is just making things harder for him/herself. This is especially the case in the field covered by this chapter.

Bibliography

- Badger, R. M. (1934). A relation between internuclear distances and bond force constants. *Journal of Chemical Physics*, 2, 128.
- Badger, R. M. (1935). The relation between the internuclear distances and force constants of molecules and its application to polyatomic molecules. *Journal of Chemical Physics*, 3, 710.
- Bak, K. L., Devlin, F. J., Ashvar, C. S., Taylor, P. R., Frisch, M. J., & Stephens, P. J. (1995). Ab initio calculation of vibrational circular dichroism spectra using gauge-invariant atomic orbitals. *Journal of Physical Chemistry*, 99, 14918.
- Baker, J. (1986). An algorithm for the location of transition states. *Journal of Computational Chemistry*, 7, 385.
- Baker, J. (1992). Geometry optimization in Cartesian coordinates: Constrained optimization. *Journal of Computational Chemistry*, 13, 240.
- Baker, J. (1993). Techniques for geometry optimization: A comparison of Cartesian and natural internal coordinates. *Journal of Computational Chemistry*, 14, 1085.
- Baker, J. (1997). Constrained optimization in delocalized internal coordinates. *Journal of Computational Chemistry*, 18, 1079.
- Baker, J. (2008). A scaled quantum mechanical reinvestigation of the vibrational spectrum of toluene. *Journal of Molecular Structure THEOCHEM*, 865, 49.
- Baker, J., & Bergeron, D. (1993). Constrained optimization in Cartesian coordinates. *Journal of Computational Chemistry*, 14, 1339.
- Baker, J., & Chan, F. (1996). The location of transition states: A comparison of Cartesian, Z-matrix, and natural internal coordinates. *Journal of Computational Chemistry*, 17, 888.

- Baker, J., & Hehre, W. J. (1991). Geometry optimization in Cartesian coordinates: The end of the Z-matrix? *Journal of Computational Chemistry*, *12*, 606.
- Baker, J., & Pulay, P. (1998). Predicting the vibrational spectra of some simple fluorocarbons by direct scaling of primitive valence force constants. *Journal of Computational Chemistry*, *19*, 1187.
- Baker, J., & Pulay, P. (2000). Efficient geometry optimization of molecular clusters. *Journal of Computational Chemistry*, *21*, 69.
- Baker, J., Muir, M., & Andzelm, J. (1985). A study of some organic reactions using density functional theory. *Journal of Chemical Physics*, *102*, 2063.
- Baker, J., Andzelm, J., Scheiner, A., & Delley, B. (1994). The effect of grid quality and weight derivatives in density functional calculations. *Journal of Chemical Physics*, *101*, 8894.
- Baker, J., Kessi, A., & Delley, B. (1996). The generation and use of delocalized internal coordinates in geometry optimization. *Journal of Chemical Physics*, *105*, 192.
- Baker, J., Jarzecki, A. A., & Pulay, P. (1998). Direct scaling of primitive valence force constants: An alternative approach to scaled quantum mechanical force fields. *Journal of Physical Chemistry A*, *102*, 1412.
- Baker, J., Kinghorn, D., & Pulay, P. (1999). Geometry optimization in delocalized internal coordinates: An efficient quadratically scaling algorithm for large molecules. *Journal of Chemical Physics*, *110*, 4986.
- Baker, J., Wolinski, K., Malagoli, M., Kinghorn, D., Wolinski, P., Magyarfalvi, G., Saebø, S., Janowski, T., & Pulay, P. (2009). Quantum chemistry in parallel with PQS. *Journal of Computational Chemistry*, *30*, 317.
- Bakken, V., & Helgaker, T. (2002). The efficient optimization of molecular geometries using redundant internal coordinates. *Journal of Chemical Physics*, *117*, 9160.
- Banerjee, A., Adams, N., Simons, J., & Shepard, R. (1985). Search for stationary points on surfaces. *Journal of Physical Chemistry*, *89*, 52.
- Becke, A. D. (1988). A multicenter numerical integration scheme for polyatomic molecules. *Journal of Chemical Physics*, *88*, 2547.
- Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *Journal of Chemical Physics*, *98*, 5648.
- Bell, S., & Crighton, J. S. (1984). Locating transition states. *Journal of Chemical Physics*, *80*, 2464.
- Bell, S., Crighton, J. S., & Fletcher, R. (1981). A new efficient method for locating saddle points. *Chemical Physics Letters*, *82*, 122.
- Bertie, J. E., Apelblat, V., & Keefe, C. D. (2005). Infrared intensities of liquids XXV: Dielectric constants, molar polarizabilities and integrated intensities of liquid toluene at 25 °C between 4800 and 400 cm⁻¹. *Journal of Molecular Structure*, *750*, 78.
- Beyer, M., & Clausen-Schaumann, H. (2005). Mechanochemistry: The mechanical activation of covalent bonds. *Chemical Review*, *105*, 2921.
- Binkley, J. S., Pople, J. A., & Hehre, W. J. (1980). Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *Journal of the American Chemical Society*, *102*, 939.
- Blom, C. E., & Altona, C. (1976). Application of self-consistent-field ab initio calculations to organic molecules II. Scale factor method for the calculation of vibrational frequencies from ab initio force constants: Ethane, propane and cyclopropane. *Molecular Physics*, *31*, 1377.
- Blom, C. E., & Altona, C. (1977a). Application of self-consistent-field ab initio calculations to organic molecules IV. Force constants of propene scaled on experimental frequencies. *Molecular Physics*, *33*, 875.
- Blom, C. E., & Altona, C. (1977b). Application of self-consistent-field ab initio calculations to organic molecules V. Ethene: General valence force field scaled on harmonic and anharmonic data, infra-red and Raman intensities. *Molecular Physics*, *34*, 177.
- Blom, C. E., Otto, L. P., & Altona, C. (1976). Application of self-consistent-field ab initio calculations to organic molecules III. Equilibrium structure of water, methanol and dimethylether,

- general valence force field of water and methanol scaled on experimental frequencies. *Molecular Physics*, 32, 1137.
- Bofill, J. M. (1994). Updated Hessian matrix and the restricted step method for locating transition structures. *Journal of Computational Chemistry*, 15, 1.
- Born, M., & Oppenheimer, R. (1927). Zur Quantentheorie der Molekeln. *Annalen der Physik*, 389, 457.
- Broyden, C. G. (1970). The convergence of a class of double-rank minimization algorithms. I: General considerations. *Journal of the Institute of Mathematics and Its Applications*, 6, 76.
- Čerjan, C. J., & Miller, W. H. (1981). On finding transition states. *Journal of Chemical Physics*, 75, 2800.
- Csaszar, P., & Pulay, P. (1984). Geometry optimization by direct inversion in the iterative subspace. *Journal of Molecular Structure THEOCHEM*, 114, 31.
- Ditchfield, R., Hehre, W. J., & Pople, J. A. (1971). Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *Journal of Chemical Physics*, 54, 724.
- Eckert, F., Pulay, P., & Werner, H.-J. (1997). Ab initio geometry optimization for large molecules. *Journal of Computational Chemistry*, 18, 1473.
- Einstein, A., & Stern, O. (1913). Einige Argumente für die Annahme einer molekularen Agitation beim absoluten Nullpunkt. *Annals of Physics*, 40, 551.
- El Youssefi, Y., Herman, M., & Liévin, J. (1998a). The ground electronic state of 1,2-dichloroethane I. Ab initio investigation of the geometrical, vibrational and torsional structure. *Molecular Physics*, 94, 461.
- El Youssefi, Y., Liévin, J., van der Auwera, J., Herman, M., Federov, A., & Snavelly, D. L. (1998b). The ground electronic state of 1,2-dichloroethane II. Experimental investigation of the fundamental and overtone vibrations. *Molecular Physics*, 94, 473.
- Farkas, O., & Schlegel, H. B. (1998). Methods for geometry optimization of large molecules. I. An $O(N^2)$ algorithm for solving systems of linear equations for the transformation of coordinates and forces. *Journal of Chemical Physics*, 109, 7100.
- Fletcher, R. (1970). A new approach to variable metric algorithms. *The Computer Journal*, 13, 317.
- Fletcher, R. (1980). *Practical methods of optimization* (Vol. 1). New York: Wiley.
- Fogarasi, G., Zhou, X., Taylor, P. W., & Pulay, P. (1992). The calculation of ab initio molecular geometries: Efficient optimization by natural internal coordinates and empirical correction by offset forces. *Journal of the American Chemical Society*, 114, 8191.
- Goldfarb, D. (1970). A family of variable-metric methods derived by variational means. *Mathematics of Computation*, 24, 23.
- Halgren, T. A., & Lipscomb, W. N. (1977). The synchronous-transit method for determining reaction pathways and locating molecular transition states. *Chemical Physics Letters*, 49, 225.
- Halls, M. D., Velkovski, J., & Schlegel, H. B. (2001). Harmonic frequency scaling factors for Hartree-Fock, S-VWN, B-LYP, B3-LYP, B3-PW91 and MP2 with the Sadlej pVTZ electric property basis set. *Theoretical Chemistry Accounts*, 105, 413.
- Hammond, G. S. (1955). A correlation of reaction rates. *Journal of the American Chemical Society*, 77, 334.
- Hertwig, R. H., & Koch, W. (1997). On the parameterization of the local correlation functional. What is Becke-3-LYP? *Chemical Physics Letters*, 268, 345.
- Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. *Physical Review B*, 136, 864.
- Ionova, I. V., & Carter, E. A. (1993). Ridge method for finding saddle points on potential energy surfaces. *Journal of Chemical Physics*, 98, 6377.
- Jiang, H., Appadoo, D., Robertson, E., & McNaughton, D. (2002). A comparison of predicted and experimental vibrational spectra in some small fluorocarbons. *Journal of Computational Chemistry*, 23, 1220.
- Johnson, B. G., Gill, P. M. W., & Pople, J. A. (1993). The performance of a family of density functional methods. *Journal of Chemical Physics*, 98, 5612.
- Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Physical Review A*, 140, 1133.

- Kveseth, K. (1974). Conformational analysis. 1. The temperature effect on the structure and composition of the rotational conformers of 1,2-dichloroethane as studied by gas electron diffraction. *Acta Chemica Scandinavica A*, 28, 482.
- Lindh, R., Bernhardsson, A., Karlström, G., & Malmquist, P.-A. (1995). On the use of a Hessian model function in molecular geometry optimizations. *Chemical Physics Letters*, 241, 423.
- Malagoli, M., & Baker, J. (2003). The effect of grid quality and weight derivatives in density functional calculations of harmonic vibrational frequencies. *Journal of Chemical Physics*, 119, 12763.
- McQuarrie, D. A. (2000). *Statistical mechanics*. California: University Science.
- Merrick, J. P., Moran, D., & Radom, L. (2007). An evaluation of harmonic vibrational frequency scale factors. *Journal of Physical Chemistry A*, 111, 11683.
- Mitin, A. V., Baker, J., & Pulay, P. (2003). An improved 6-31G* basis set for first-row transition metals. *Journal of Chemical Physics*, 118, 7775.
- Mizushima, S., Shimanouchi, T., Harada, I., Abe, Y., & Takeuchi, H. (1975). Infrared and Raman spectra of 1,2-dichloroethane and its deuterium compound in the gaseous, liquid, and solid states. *Canadian Journal of Physics*, 53, 2085.
- Muir, M., & Baker, J. (1996). A systematic density functional study of fluorination in methane, ethane and ethylene. *Molecular Physics*, 89, 211.
- Murtagh, B. A., & Sargent, R. W. H. (1970). Computational experience with quadratically convergent minimisation methods. *Computer Journal*, 13, 185.
- Paizs, B., Baker, J., Suhai, S., & Pulay, P. (2000). Geometry optimization of large biomolecules in redundant internal coordinates. *Journal of Chemical Physics*, 113, 6566.
- Pardalos, P. M., Shalloway, D., & Xue, G. (Eds.). (1995). *Global minimization of nonconvex functions: Molecular conformation and protein folding*. Providence: American Mathematical Society.
- Peng, C., & Schlegel, H. B. (1993). Combining synchronous transit and quasi-Newton methods to find transition states. *Israel Journal of Chemistry*, 33, 449.
- Peng, C., Ayala, P. Y., Schlegel, H. B., & Frisch, M. J. (1996). Using redundant internal coordinates to optimize equilibrium geometries and transition states. *Journal of Computational Chemistry*, 17, 49.
- Pople, J. A., Krishnan, R., Schlegel, H. B., & Binkley, J. S. (1979). Derivative studies in Hartree-Fock and Møller-Plesset theories. *International Journal of Quantum Chemistry: Symposium*, 13, 225.
- Poppinger, D. (1975). On the calculation of transition states. *Chemical Physics Letters*, 35, 550.
- Powell, M. J. D. (1971). Recent advances in unconstrained optimization. *Mathematical Programming*, 1, 26.
- PQS. (2010). *PQS version 4.0*. Parallel Quantum Solutions 2013 Green Acres Road, Suite A, Fayetteville, AR 72703. Email: sales@pqs-chem.com. URL: <http://www.pqs-chem.com>
- Pulay, P. (1977). *Direct use of the gradient for investigating molecular energy surfaces*. New York: Plenum.
- Pulay, P. (1980). Convergence acceleration of iterative sequences. The case of SCF iteration. *Chemical Physics Letters*, 73, 393.
- Pulay, P. (1982). Improved SCF convergence acceleration. *Journal of Computational Chemistry*, 3, 556.
- Pulay, P., & Fogarasi, G. (1992). Geometry optimization in redundant internal coordinates. *Journal of Chemical Physics*, 96, 2856.
- Pulay, P., & Torok, F. (1966). On the parameter form of the force constant matrix II. Investigation of the assignment with the aid of the parameter form. *Acta Chimica Academiae Scientiarum Hungaricae*, 47, 273.
- Pulay, P., Fogarasi, G., Pang, F., & Boggs, J. E. (1979). Systematic ab initio gradient calculation of molecular geometries, force constants, and dipole moment derivatives. *Journal of the American Chemical Society*, 101, 2550.
- Pulay, P., Fogarasi, G., Pongor, G., Boggs, J. E., & Vargha, A. (1983). Combination of theoretical ab initio and experimental information to obtain reliable harmonic force constants. Scaled

- quantum mechanical (QM) force fields for glyoxal, acrolein, butadiene, formaldehyde, and ethylene. *Journal of the American Chemical Society*, *105*, 7037.
- Raman, C. V., & Krishnan, K. S. (1928). A new type of secondary radiation. *Nature*, *121*, 501.
- Rassalov, V. A., Pople, J. A., Ratner, M. A., & Windus, T. L. (1998). 6-31G* basis set for atoms K through Zn. *Journal of Chemical Physics*, *109*, 1223.
- Schäfer, L. (1983). The ab initio gradient revolution in structural chemistry: The importance of local molecular geometries and the efficacy of joint quantum mechanical and experimental procedures. *Journal of Molecular Structure*, *100*, 51.
- Schlegel, H. B. (1984). Estimating the hessian for gradient-type geometry optimizations. *Theoretica Chimica Acta*, *66*, 333.
- Shanno, D. F. (1970). Conditioning of quasi-Newton methods for function minimization. *Mathematics of Computation*, *24*, 647.
- Simons, J., Jorgensen, P., Taylor, H., & Ozment, J. (1983). Walking on potential energy surfaces. *Journal of Physical Chemistry*, *87*, 2745.
- Stephens, P. J., & Lowe, M. A. (1985). Vibrational circular dichroism. *Annual Review of Physical Chemistry*, *36*, 213.
- Stewart, J. J. P. (1989). Optimization of parameters for semiempirical methods I. Method. *Journal of Computational Chemistry*, *10*, 209.
- Swart, M., & Bickelhaupt, F. M. (2006). Optimization of strong and weak coordinates. *International Journal of Quantum Chemistry*, *106*, 2536.
- Wikipedia (2010). The article "Born–Oppenheimer approximation."
- Wilson, E. B., Decius, J. C., & Cross, P. C. (1955). *Molecular vibrations*. New York: McGraw-Hill.
- Wolinski, K., & Baker, J. (2009). Theoretical predictions of enforced structural changes in molecules. *Molecular Physics*, *107*, 2403.

Michał Jaszuński, Antonio Rizzo, and Kenneth Ruud

Contents

Introduction	498
The Molecular Hamiltonian	499
The Molecular Breit–Pauli Hamiltonian	500
Small Terms Due to the Scalar Potential in the Hamiltonian	504
The Magnetic Vector Potential	505
Small Terms Due to the Vector Potential in the Hamiltonian	506
Response Theory	510
Molecular Response: Definitions, Symbols	510
Expansions of Energy and Multipole Moments	514
Electric Properties	515
Electric Multipole Moments	515
Dipole Polarizability	521
First Dipole Hyperpolarizability	523
Second Dipole Hyperpolarizability	525
Cauchy Moments	527
Long-Range Dispersion Interaction Coefficients	528

We dedicate this work to the memory of our longtime friend Andrzej Sadlej.

M. Jaszuński (✉)

Institute of Organic Chemistry, Polish Academy of Sciences, Warszawa, Poland
e-mail: michaljz@icho.edu.pl

A. Rizzo

Istituto per i Processi Chimico Fisici del Consiglio Nazionale delle Ricerche, Area della Ricerca di Pisa, Pisa, Italy
e-mail: rizzo@ipcf.cnr.it

K. Ruud

Centre for Theoretical and Computational Chemistry, Department of Chemistry, University of Tromsø, Tromsø, Norway
e-mail: kenneth.ruud@uit.no

Electric Field Gradient at the Nucleus, Nuclear Quadrupole Coupling Constant	529
One-Photon Absorption, Excitation Energies, and Transition Moments	530
Two-Photon Absorption	533
Magnetic Properties	535
Magnetizability	535
Rotational g Tensor	539
Birefringences and Dichroisms	542
Optical Rotation	544
Circular Dichroism	547
Faraday Effect	550
Hypermagnetizabilities, Cotton–Mouton Effect	551
Nuclear Spin–Related Properties	556
NMR Effective Spin Hamiltonian	557
Nuclear Magnetic Shielding	559
Nuclear Spin Rotation Constants	567
NMR Indirect Spin–Spin Coupling	568
Electron Spin–Related Properties	571
Spin–Orbit Coupling Constants	572
Phosphorescence	573
ESR Effective Spin Hamiltonian	574
ESR Electronic g -Tensor	576
Zero-Field Splitting	577
ESR Hyperfine Coupling Tensors	579
Conclusions	581
Bibliography	582

Abstract

The theory and applications of *ab initio* methods for the calculation of molecular properties are reviewed. A wide range of properties characterizing the interactions of molecules with external or internal sources of static or dynamic electromagnetic fields (including nonlinear properties and those related to nuclear and electron spins) is considered. Emphasis is put on the properties closely connected to the parameters used to describe experimentally observed spectra. We discuss the definitions of these properties, their relation to experiment, and give some remarks regarding various computational aspects. Theory provides a unified approach to the analysis of molecular properties in terms of average values, transition moments, and linear and nonlinear responses to the perturbing fields. Several literature examples are given, demonstrating that theoretical calculations are becoming easier, and showing that computed *ab initio* molecular properties are in many cases more accurate than those extracted from experimentally observed data.

Introduction

The term “molecular properties” is generally used to denote properties characterizing the interactions of molecules with internal or external sources of static or dynamic electromagnetic fields. In the description of a complex quantum-mechanical system consisting of a molecule and a field it is convenient,

both in theory and in experiment, to introduce an approximation that separates the molecule, the field, and the molecule–field interaction. One can then introduce an additional approximation and analyze the effects of the interaction only for the molecule, assuming a passive role of the interaction-induced changes in the field. In this way we can describe the interaction between the field and the molecule in terms of what we can consider intrinsic properties of the molecule.

Most spectroscopic methods assume that the experiment can be described in terms of these molecule–field interactions. There is, therefore, often a very close connection between the parameters used to describe the experimental spectra and the molecular properties that describe the molecule–field interactions probed by the experiment. The set of properties describing the influence of external electric and magnetic fields – static and time-dependent, spatially uniform or nonuniform – can be used to determine parameters that enter into the analysis of different experimental spectra. In particular, optical properties describe the interactions of a molecule with an external electromagnetic radiation field, and characterize both linear and nonlinear responses to the fields. Other sets of molecular properties describe the interactions of external magnetic fields and the magnetic fields due to the spins of all particles – that is, nuclear spins (specific for each nuclear isotope) and electron spin. These properties determine the parameters entering the analysis of Nuclear Magnetic Resonance (NMR) and Electron Spin Resonance (ESR) experiments. Additional fields can be created by the presence of neighboring molecules, and various effects due to weak long-range intermolecular interactions can also be described in terms of molecular properties. The treatment of molecular vibrational properties is in many cases based on the same principles as the formalism explored here; however, the calculation of such properties differs quite substantially from that of electronic properties due to the different quantum-mechanical treatments of the electronic and nuclear motions, and for this reason we will not consider vibrational properties in this chapter.

The Molecular Hamiltonian

Our starting point for electronic structure calculations of molecules is the time-independent, field-free nonrelativistic Schrödinger equation for the electronic degrees of freedom of a molecule

$$\hat{H}^{(0)}\Psi = E\Psi, \quad (1)$$

where Ψ is the wave function describing the electronic state of the molecule, E the energy of this state, and $\hat{H}^{(0)}$ the Hamiltonian operator for the system, given by

$$\hat{H}^{(0)} = \frac{1}{2m_e} \sum_i \hat{p}_i^2 - \frac{e^2}{4\pi\epsilon_0} \sum_{iK} \frac{Z_K}{r_{iK}} + \frac{e^2}{8\pi\epsilon_0} \sum_{i \neq j} \frac{1}{r_{ij}} + \frac{e^2}{8\pi\epsilon_0} \sum_{K \neq L} \frac{Z_K Z_L}{R_{KL}}, \quad (2)$$

where the first term corresponds to the kinetic energy of the electrons, $\mathbf{p}_i = -i\hbar\nabla_i$ is the linear momentum operator of particle i , the second term is the attractive

potential between the electrons and the nuclei, the third term the electron–electron repulsion potential, and the last term the repulsive potential for the interaction between the nuclear point charges. In Eq. 2, e is the elementary charge, $Z_K e$ is the charge of nucleus K , ϵ_0 is the electric constant (electric permittivity of vacuum), and m_e the mass of the electron. We use \mathbf{r}_i and \mathbf{R}_K to denote the vectors giving the position of electron i and nucleus K with respect to a chosen origin of the molecular frame. If needed, for a specific frame of interest, we use the notation

$$\mathbf{r}_{iO} = \mathbf{r}_i - \mathbf{O} \quad (3)$$

to emphasize that the positions are given relative to an origin \mathbf{O} . We define vectors between two different points in space with the directional convention

$$\mathbf{R}_{KL} = \mathbf{R}_K - \mathbf{R}_L, \quad (4)$$

corresponding to a vector pointing from nucleus L to nucleus K . In the summations over electrons or nuclei, we always assume – unless explicitly stated otherwise – that the summation runs over all the particles of a given type in the molecule. Whenever needed, we shall use for vector and tensor coordinates the following subscripts:

- Greek symbols for unspecified molecular frame coordinates; for these coordinates the Einstein summation convention is applied
- x, y, z for specific molecular frame coordinates (for instance, determined by molecular symmetry)
- X, Y, Z for laboratory frame coordinates
- 1, 2, 3 when the tensor is specified using its own principal axes and values
- a, b, c for coordinates defined by the moment of inertia \mathbf{I}

where $\mathbf{I} = \sum_K M_K \mathbf{R}_K \mathbf{R}_K^T$ and M_K is the mass of nucleus K .

The Hamiltonian in Eq. 2 does not take into account the interactions with external electric or magnetic fields. Moreover, so far we have neglected numerous other interactions, related to the existence of the electron and nuclear spins (and the corresponding magnetic moments), their interactions with external fields, and relativistic effects. Thus, for the study of molecular properties which depend on these interactions we need to consider a more general approach, based for instance on the Breit–Pauli Hamiltonian (Dyall and Faegri 2007; Moss 1973). In many cases the contributions to molecular properties due to the additional terms appearing in the Breit–Pauli Hamiltonian can be estimated using the available solutions of the (nonrelativistic) Schrödinger equation and perturbation theory.

The Molecular Breit–Pauli Hamiltonian

There are several ways to introduce and justify the use of the Breit–Pauli Hamiltonian. One can start from the Dirac equation which involves the relativistic

Hamiltonian for the free electron, and extend it by applying various approximations to many electrons in the field generated by the nuclei. We will not discuss here the steps taken in the reduction of the Dirac Hamiltonian to the Breit–Pauli Hamiltonian (Dyall and Faegri 2007; Moss 1973). Instead, we only list the various contributions appearing in the Breit–Pauli Hamiltonian. The approximate form of the Breit–Pauli Hamiltonian we will discuss is accurate to order α_{fs}^2 , where α_{fs} is the fine structure constant, and higher-order corrections are not considered at this stage (we will include later selected terms of the order α_{fs}^4 needed for the calculation of response properties in the presence of magnetic fields). Finally, even though some of the operators appearing in the electronic Breit–Pauli Hamiltonian are of little interest from the point of view of molecular properties, they are included here for completeness.

The purely electronic contributions to the Breit–Pauli Hamiltonian can be written as (Bethe and Salpeter 1957, p. 170; Moss 1973):

$$\hat{H}_{\text{el}}^{\text{BP}} = \sum_i \left(\frac{\pi_i^2}{2m_e} - e\phi \right. \quad (5)$$

$$\left. - \mathbf{m}_i \cdot \mathbf{B} \right. \quad (6)$$

$$\left. + \frac{\alpha_{\text{fs}}^2}{4m_e} \mathbf{m}_i \cdot (\pi_i \times \mathbf{F} - \mathbf{F} \times \pi_i) \right. \quad (7)$$

$$\left. + \frac{e\hbar^2 \alpha_{\text{fs}}^2}{8m_e^2} (\nabla \cdot \mathbf{F}) \right. \quad (8)$$

$$\left. - \frac{\alpha_{\text{fs}}^2}{8m_e^3} \pi_i^4 \right. \quad (9)$$

$$\left. + \frac{\alpha_{\text{fs}}^2}{2m_e^2} (\mathbf{m}_i \cdot \mathbf{B}) \pi_i^2 \right. \quad (10)$$

$$\left. + \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} \left\{ \frac{e^2}{2} \frac{1}{r_{ij}} \right. \right. \quad (11)$$

$$\left. - \frac{e^2 \alpha_{\text{fs}}^2}{4m_e^2} \left[\frac{\pi_i \cdot \pi_j}{r_{ij}} + \frac{(\pi_i \cdot \mathbf{r}_{ij})(\mathbf{r}_{ij} \cdot \pi_j)}{r_{ij}^3} \right] \right. \quad (12)$$

$$\left. + \frac{e\alpha_{\text{fs}}^2}{2m_e} \frac{\mathbf{m}_i \cdot (\mathbf{r}_{ij} \times \pi_i)}{r_{ij}^3} \right. \quad (13)$$

$$\left. - \frac{e\alpha_{\text{fs}}^2}{m_e} \frac{\mathbf{m}_i \cdot (\mathbf{r}_{ij} \times \pi_j)}{r_{ij}^3} \right. \quad (14)$$

$$+ \frac{\alpha_{\text{fs}}^2}{2} \left[\frac{\mathbf{m}_i \cdot \mathbf{m}_j}{r_{ij}^3} - \frac{3(\mathbf{m}_i \cdot \mathbf{r}_{ij})(\mathbf{r}_{ij} \cdot \mathbf{m}_j)}{r_{ij}^5} - \frac{8\pi}{3} \delta(\mathbf{r}_{ij}) \mathbf{m}_i \cdot \mathbf{m}_j \right] \quad (15)$$

$$- \frac{\pi e^2 \hbar^2 \alpha_{\text{fs}}^2}{2m_e^2} \delta(\mathbf{r}_{ij}) \left. \vphantom{\frac{\pi e^2 \hbar^2 \alpha_{\text{fs}}^2}{2m_e^2} \delta(\mathbf{r}_{ij})} \right\}, \quad (16)$$

where $\delta \mathbf{r}_{ij}$ is the Dirac delta function

$$\delta(\mathbf{r}_{ij}) = 0 \quad \text{for} \quad \mathbf{r}_{ij} \neq \mathbf{0}; \quad \int_V \delta(\mathbf{r}_{ij}) d^3 \mathbf{r} = 1. \quad (17)$$

In addition to these purely electronic contributions, we must also consider the interactions that involve simultaneously one of the electrons and one of the nuclei:

$$\hat{H}_{\text{eN}}^{\text{BP}} = \frac{1}{4\pi\epsilon_0} \sum_{iK} Z_K \left[-\frac{e^2}{r_{iK}} \right. \quad (18)$$

$$+ \frac{e\alpha_{\text{fs}}^2}{2m_e} \frac{\mathbf{m}_i \cdot (\mathbf{r}_{iK} \times \boldsymbol{\pi}_i)}{r_{iK}^3} \quad (19)$$

$$\left. + \frac{\pi e^2 \hbar^2 \alpha_{\text{fs}}^2}{2m_e^2} \delta(\mathbf{r}_{iK}) \right]. \quad (20)$$

Finally, we have the contributions arising from the nuclei only

$$\hat{H}_{\text{nucI}}^{\text{BP}} = \sum_K \left[Z_K e \phi_K - \mathbf{m}_K \cdot \mathbf{B} \right. \quad (21)$$

$$\left. + \sum_{L \neq K} \frac{e^2}{8\pi\epsilon_0} \frac{Z_K Z_L}{R_{KL}} \right]. \quad (22)$$

We have omitted in these equations all the terms dependent on the (inverse) masses of the nuclei. In the purely electronic and purely nuclear parts we have not included the rest mass of the particles. In the above equations we have introduced the mechanical momentum

$$\boldsymbol{\pi}_i = \mathbf{p}_i - q_i \mathbf{A}(\mathbf{r}_i), \quad (23)$$

where $\mathbf{A}(\mathbf{r}_i)$ is the magnetic vector potential experienced by electron i , and q_i is the charge of the particle, which for an electron is $-e$. The corresponding magnetic field induction is \mathbf{B} (see below). ϕ represents the external scalar electrostatic potential, and the electric field due to this external potential is $\mathbf{F} = -\nabla\phi$. We have also

introduced the magnetic moment of the electron \mathbf{m}_i , which is related to the electron spin through

$$\mathbf{m}_i = -g_e \frac{e\hbar}{2m_e} \mathbf{s}_i = -g_e \mu_B \mathbf{s}_i, \quad (24)$$

where g_e is the electron g factor and the Bohr magneton μ_B is the unit for magnetic moments. In a similar manner, \mathbf{m}_K is the nuclear magnetic moment, related to the nuclear spin \mathbf{I}_K as

$$\mathbf{m}_K = \hbar \gamma_K \mathbf{I}_K = \mu_N g_K \mathbf{I}_K, \quad (25)$$

where γ_K is the nuclear magnetogyric ratio, $\mu_N = e\hbar/2m_p$ is the nuclear magneton, m_p the proton mass, and g_K is the g factor of nucleus K , unique for each isotope. For a detailed description of the different terms appearing in the Breit–Pauli Hamiltonian we refer to Moss (1973). The Breit–Pauli Hamiltonian, Eqs. 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, and 22, does not contain explicitly any coupling between the nuclear magnetic moments \mathbf{m}_K and the magnetic moments of the electrons \mathbf{m}_i . In sections “The Magnetic Vector Potential” and “Small Terms Due to the Vector Potential in the Hamiltonian,” we introduce these interactions by establishing a magnetic vector potential for the nuclear magnetic moments in a similar manner to the magnetic vector potential for the external magnetic field induction. These interactions, important for determining the observable NMR and ESR spectra, will thus appear through the couplings arising when we expand π_i^2 (and the spin–Zeeman term leading to Eq. 6) using the full mechanical momentum operator given in Eq. 23.

The Breit–Pauli Hamiltonian describes a number of contributions that can be considered small compared to the nonrelativistic Hamiltonian given in Eq. 2. These contributions are thus ideally suited to be treated by perturbation theory, at least as long as we do not consider the heaviest atoms of the periodic table, where the relativistic effects become substantial and the lack of a variational bound for the Breit–Pauli Hamiltonian makes any perturbation approach fail. For the energy of molecules consisting of light atoms, the relativistic effects can to a first approximation often be treated considering only the mass-velocity (Eq. 9) and one-electron Darwin (Eq. 20) terms.

Some of the terms included in the Breit–Pauli Hamiltonian also describe small interactions that can be probed experimentally by inducing suitable excitations in the electron or nuclear spin space, giving rise to important contributions to observable NMR and ESR parameters. In particular, for molecular properties for which there are interaction mechanisms involving the electron spin, also the spin-orbit interaction (Eqs. 13 and 14) becomes important. The Breit–Pauli Hamiltonian in Eqs. 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, and 22, however, only includes molecule–external field interactions through the presence of a scalar

electrostatic potential ϕ (and the associated electric field \mathbf{F}) and the appearance of the magnetic vector potential in the mechanical momentum operator (Eq. 23). In order to extract in more detail the interaction between the electronic structure of a molecule and an external electromagnetic field, we need to consider in more detail the form of the scalar and vector potentials.

Small Terms Due to the Scalar Potential in the Hamiltonian

Let us begin with the simplest example, the interaction of a molecule with a static electric field. This interaction can be described by the Hamiltonian

$$\hat{H} = \hat{H}^{(0)} - e \sum_i \phi(\mathbf{r}_i) + e \sum_k Z_K \phi(\mathbf{R}_K), \quad (26)$$

where we have neglected the spin-orbit and relativistic terms given in Eqs. 7 and 8. We can expand the potential around a common origin

$$\phi(\mathbf{r}_i) = \phi(0) + \frac{\partial \phi}{\partial r_{i\alpha}} r_{i\alpha} + \frac{1}{2} \frac{\partial^2 \phi}{\partial r_{i\alpha} \partial r_{i\beta}} r_{i\alpha} r_{i\beta} + \dots \quad (27)$$

with a similar expansion for the potential at the position of the nuclei, $\phi(\mathbf{R}_K)$. Recalling that the electric field due to the external potential is $\mathbf{F} = -\nabla\phi$, we can rewrite Eq. 26 in terms of electric multipole operators

$$\hat{H} = \hat{H}^{(0)} + Q\phi(0) + \hat{H}^{\mathbf{F}} \cdot \mathbf{F} + \hat{H}^{\nabla\mathbf{F}} \cdot \nabla\mathbf{F} + \dots, \quad (28)$$

where Q is the total charge of the molecule. We note that the multipole operators are defined with respect to the chosen common origin. For a neutral system, the first two contributions describing the field effects are given by the:

- External electric field perturbation:

$$\hat{H}^{\mathbf{F}} = -\boldsymbol{\mu}, \quad (29)$$

where the electric dipole operator is

$$\hat{\mu}_\alpha = -e \sum_i \hat{r}_{i\alpha} + e \sum_K Z_K R_{K\alpha}. \quad (30)$$

- External electric field gradient perturbation:

$$\hat{H}^{\nabla\mathbf{F}} = -\frac{1}{3}\Theta, \quad (31)$$

where the electric quadrupole operator in its traceless form is

$$\hat{\Theta}_{\alpha\beta} = -\frac{e}{2} \sum_i (3\hat{r}_{i\alpha}\hat{r}_{i\beta} - \hat{r}_i^2\delta_{\alpha\beta}) + \frac{e}{2} \sum_K Z_K (3R_{K\alpha}R_{K\beta} - R_K^2\delta_{\alpha\beta}). \quad (32)$$

The interaction due to the external field gradient can also be expressed in terms of the second moments (traced form)

$$\hat{q}_{\alpha\beta} = -e \sum_i \hat{r}_{i\alpha}\hat{r}_{i\beta} + e \sum_K Z_K R_{K\alpha}R_{K\beta}. \quad (33)$$

Later we shall discuss also the effects due to the:

- Electric field gradient at the nucleus K:

$$\begin{aligned} \hat{V}_{\alpha\beta}^K &= \frac{1}{4\pi\epsilon_0} e \sum_i \frac{3\hat{r}_{iK,\alpha}\hat{r}_{iK,\beta} - \hat{r}_{iK}^2\delta_{\alpha\beta}}{r_{iK}^5} \\ &\quad - \frac{1}{4\pi\epsilon_0} e \sum_{L \neq K} Z_L \frac{3R_{LK,\alpha}R_{LK,\beta} - R_{LK}^2\delta_{\alpha\beta}}{R_{LK}^5}, \end{aligned} \quad (34)$$

since we will need it to describe the interaction with the nuclear quadrupole moments, \mathbf{Q}_K , described by the interaction Hamiltonian

$$\hat{H}^{VK} = -\frac{1}{2} \sum_K \mathbf{V}^K \mathbf{Q}_K. \quad (35)$$

The Magnetic Vector Potential

We consider a molecule in which the magnetic field arises from two primary sources, an external magnetic field induction and the permanent magnetic moments of nuclei possessing a spin. In the minimal coupling approximation, the mechanical momentum operator (Eq. 23) is given as

$$\pi_i = \mathbf{p}_i + e\mathbf{A}^{\mathbf{B}}(\mathbf{r}_i) + e \sum_K \mathbf{A}^{\text{m}K}(\mathbf{r}_i), \quad (36)$$

where the superscripts indicate the different sources of the magnetic vector potential. In Eq. 36 we have explicitly indicated that we are interested in the magnetic vector potential at the positions of the electrons.

The magnetic vector potential is not uniquely defined, and we have to choose the gauge (Jackson 1998). In the Coulomb gauge, that is, where the magnetic vector potential is divergence free at all points in space,

$$\nabla \cdot \mathbf{A}(\mathbf{r}) = 0, \quad (37)$$

we may choose the vector potential describing the interaction with the external magnetic field induction and with the magnetic moments of the nuclei in the form

$$\mathbf{A}(\mathbf{r}_i) = \frac{1}{2}\mathbf{B} \times \mathbf{r}_{iO} + \alpha_{fs}^2 \sum_K \frac{\mathbf{m}_K \times \mathbf{r}_{iK}}{r_{iK}^3}. \quad (38)$$

It can easily be verified that the magnetic vector potential (Eq. 38) is divergence free. Since the magnetic field induced by the magnetic vector potential is given by

$$\mathbf{B} = \nabla \times \mathbf{A}^{\mathbf{B}}(\mathbf{r}_i), \quad (39)$$

we can add any scalar differentiable potential $\nabla\Lambda$ to the magnetic vector potential in Eq. 38 without changing the description of the magnetic field induction experienced by the molecule (since $\nabla \times (\nabla\Lambda) = 0$). There are, therefore, multiple definitions of the magnetic vector potential that describe the same magnetic field induction, and thus the same physical situation. This invariance with respect to the choice of magnetic vector potential may be lost in approximate calculations. This in turn leads to what is known as the gauge dependence of computed magnetic properties, an undesirable feature of approximate calculations – the results depend on an arbitrary factor in the calculation. In practice, when interactions with external magnetic fields are considered, special computational methods are introduced to bypass this problem and to obtain reliable and unambiguous gauge-origin independent results. We will return to this point when we consider magnetic properties in section “[Magnetic Properties](#).”

Small Terms Due to the Vector Potential in the Hamiltonian

Having now established the form of the vector potential, we return to the molecular Hamiltonian. In the presence of magnetic fields, the kinetic energy is defined by the mechanical momentum operator, Eq. 23 (instead of the linear momentum operator), and in addition we need to consider the Zeeman term, Eq. 6. We can expand all the terms appearing in the kinetic energy, using the vector potential in the mechanical momentum, and introduce the magnetic field induction in the Zeeman term. This yields the following form of the Hamiltonian:

$$\hat{H} = \hat{H}^0 - i \frac{e\hbar}{m_e} \sum_i \mathbf{A}(\mathbf{r}_i) \cdot \nabla_i + \frac{e^2}{2m_e} \sum_i \mathbf{A}^2(\mathbf{r}_i) + \sum_i \mathbf{m}_i \cdot (\nabla_i \times \mathbf{A}(\mathbf{r}_i)), \quad (40)$$

where \hat{H}^0 indicates the Hamiltonian in Eq. 2. In this equation we have taken advantage of the commutator relation $[\mathbf{A}(\mathbf{r}_i), \nabla_i] = 0$ (this relation holds true for our choice of the Coulomb gauge, see Eq. 37). Inserting the magnetic vector potential, Eq. 38, into Eq. 40 we get

$$\hat{H} = \hat{H}^0 + \hat{H}^{\text{B}}\mathbf{B} + \sum_K \hat{H}^{K(\text{PSO})}\mathbf{m}_K \quad (41)$$

$$+ \mathbf{B}^T \hat{H}^{\text{B},\text{B}}\mathbf{B} + \sum_K \mathbf{B}^T \hat{H}^{B,K}\mathbf{m}_K + \frac{1}{2} \sum_{K \neq L} \mathbf{m}_K^T \hat{H}^{K,L(\text{DSO})}\mathbf{m}_L \quad (42)$$

$$+ \sum_i \mathbf{m}_i \hat{H}^{\text{s},\text{B}}\mathbf{B} + \sum_{iK} \mathbf{m}_i \hat{H}^{K(\text{FC})}\mathbf{m}_K + \sum_{iK} \mathbf{m}_i \hat{H}^{K(\text{SD})}\mathbf{m}_K. \quad (43)$$

The summation over the nuclear magnetic moments in the last term of Eq. 42 should include contributions quadratic in \mathbf{m}_K , but since these contributions do not correspond to any measurable quantity, we have restricted the summation to different nuclei. We note also that these neglected terms are divergent, which shows the limitations of the applied theory.

Let us analyze the explicit expressions for the different operators in Eqs. 41, 42, and 43, which are obtained by inserting the magnetic vector potential Eq. 38 into Eq. 40. The contributions in Eq. 41 arise from the unperturbed Hamiltonian and from the term linear in $\mathbf{A}(\mathbf{r}_i)$. The operators involved are:

- Magnetic dipole operator

$$\hat{H}^{\text{B}} = \frac{e}{2m_e} \sum_i \mathbf{l}_{iO} = \frac{\mu_B}{\hbar} \sum_i \mathbf{l}_{iO} = -i \mu_B \sum_i (\mathbf{r}_{iO} \times \nabla_i) = -\mathbf{m}. \quad (44)$$

The magnetic dipole moment is proportional to the angular momentum, the coefficient of proportionality being the magnetogyric ratio $\gamma_e = e/2m_e$. The orbital angular momentum of electron i is defined as

$$\mathbf{l}_{iO} = \mathbf{r}_{iO} \times \mathbf{p}_i = -i\hbar \mathbf{r}_{iO} \times \nabla_i. \quad (45)$$

The magnetic dipole operator describes the magnetic moment induced by the angular momentum of the electrons arising from the orbital motion, in contrast to the magnetic moment arising from the electron spin which we discuss when we consider the spin-Zeeman interaction.

- Paramagnetic spin-orbit operator

$$\hat{H}^{K(\text{PSO})} = \alpha_{\text{fs}}^2 \frac{2\mu_B}{\hbar} \sum_i \frac{\mathbf{l}_{iK}}{r_{iK}^3}. \quad (46)$$

The paramagnetic spin-orbit interaction, often called the orbital hyperfine interaction, describes the interaction of the magnetic moment induced by the orbital motion of the electrons with the nuclear magnetic moments.

The contributions in Eq. 42 arise from the term quadratic in $\mathbf{A}(\mathbf{r}_i)$ in Eq. 40. They involve the following operators:

- Diamagnetic magnetizability operator

$$\hat{H}^{\text{B,B}} = \frac{e^2}{8m_e} \sum_i (r_{iO}^2 \mathbf{1} - \mathbf{r}_{iO} \mathbf{r}_{iO}^T), \quad (47)$$

where $\mathbf{1}$ is a 3×3 unit matrix. This operator is closely related to the second moment of the electronic charge, and thus reflects the size of the electron density distribution. The diamagnetic magnetizability tensor is a symmetric 3×3 tensor operator, and has, therefore, in general six independent tensor operator elements.

- Diamagnetic shielding operator

$$\hat{H}^{\text{B,K}} = \alpha_{\text{fs}}^2 \frac{e^2}{2m_e} \sum_i \frac{(\mathbf{r}_{iO} \cdot \mathbf{r}_{iK}) \mathbf{1} - \mathbf{r}_{iO} \mathbf{r}_{iK}^T}{r_{iK}^3}. \quad (48)$$

The diamagnetic shielding tensor describes the interaction of the magnetic field induction with the nuclear magnetic moment, mediated by the orbital motion of the electrons. We note that the diamagnetic shielding tensor is a 3×3 nonsymmetric tensor operator, and thus in general has nine independent elements.

- Diamagnetic spin-orbit operator

$$\hat{H}^{\text{K,L(DSO)}} = \alpha_{\text{fs}}^4 \frac{e^2}{2m_e} \sum_i \frac{(\mathbf{r}_{iK} \cdot \mathbf{r}_{iL}) \mathbf{1} - \mathbf{r}_{iK} \mathbf{r}_{iL}^T}{r_{iK}^3 r_{iL}^3}. \quad (49)$$

The diamagnetic spin-orbit operator describes the direct interaction between the two orbital magnetic moments induced in the electron density by the presence of the magnetic moments of nuclei K and L . As for the diamagnetic shielding operator, this is a 3×3 nonsymmetric tensor operator, with in general nine independent elements.

Finally, the terms in Eq. 43 arise from the spin–Zeeman term, the last term in Eq. 40.

- Spin–Zeeman operator

$$\hat{H}^{\text{s,B}} = -\mathbf{1}. \quad (50)$$

The spin–Zeeman operator describes the direct interaction of an electron spin with the external magnetic field induction, and does not act on the spatial coordinates of the electrons, acting only in spin space (this contribution thus vanishes for closed-shell systems). We note that it is customary to describe this term in the Hamiltonian using the electron spin explicitly rather than the electron magnetic moment.

- Fermi contact (FC) operator

$$\hat{H}^{\text{K(FC)}} = -\alpha_{\text{fs}}^2 \frac{8\pi}{3} \sum_i \delta(\mathbf{r}_{iK}) \mathbf{1}. \quad (51)$$

The Fermi contact operator acts both in the spin and in the spatial space of the electronic coordinates, often leading to a very different behavior of the response properties than for instance the $\hat{H}^{K(\text{PSO})}$ operator. The Fermi contact operator represents the direct interaction of the electron spin-magnetic moment with the nuclear magnetic moment at the position of the nucleus, and can also be considered to be a measure of the spin-polarization at the nucleus.

- Spin–dipole (SD) operator

$$\hat{H}^{K(\text{SD})} = -\alpha_{\text{fs}}^2 \sum_i \frac{3\mathbf{r}_{iK}\mathbf{r}_{iK}^T - r_{iK}^2\mathbf{1}}{r_{iK}^5}. \quad (52)$$

The spin–dipole operator is in the form of a classical dipole–dipole interaction between the magnetic dipole of the nucleus and the magnetic dipole of the electron.

For the contributions involving the electronic spin operators through the electron spin-magnetic moment \mathbf{m}_i — that is, $\hat{H}^{K(\text{FC})}$ and $\hat{H}^{K(\text{SD})}$ — we need to keep track of the electron spin index and of the nuclear spin index. Due to the nature of the Dirac delta function appearing in the operator for $\hat{H}^{K(\text{FC})}$, this operator is diagonal in the two tensor indices, that is

$$\hat{H}_{\alpha\beta}^{K(\text{FC})} = \hat{H}_{\text{iso}}^{K(\text{FC})} \delta_{\alpha\beta}. \quad (53)$$

Both the Fermi contact and the spin–dipole operators can be derived considering the magnetic field induction created by the vector potential describing the nuclear magnetic moment

$$\begin{aligned} \nabla \times \mathbf{A}^{\mathbf{m}_K}(\mathbf{r}_i) &= \alpha_{\text{fs}}^2 \nabla \times \frac{\mathbf{m}_K \times \mathbf{r}_{iK}}{r_{iK}^3} \\ &= \alpha_{\text{fs}}^2 \left(\sum_i \frac{3\mathbf{r}_{iK}\mathbf{r}_{iK}^T - r_{iK}^2\mathbf{1}}{r_{iK}^5} + \frac{8\pi}{3} \sum_i \delta(\mathbf{r}_{iK}) \mathbf{1} \right) \mathbf{m}_K, \end{aligned} \quad (54)$$

which can be derived using standard vector product formulas.

We shall in later sections also discuss the effects of the spin-orbit interaction. The total spin-orbit (SO) interaction operator (McWeeny 1992) can be defined combining the terms included in Eqs. 13, 14, and 19 as

$$\begin{aligned} \hat{H}^{\text{SO}} &= -\frac{e}{8\pi\epsilon_0 m_e} \alpha_{\text{fs}}^2 \left[\sum_{iK} Z_K \frac{\mathbf{m}_i \cdot \mathbf{1}_{iK}}{r_{iK}^3} - \sum_{i \neq j} \frac{(\mathbf{m}_i + 2\mathbf{m}_j) \cdot \mathbf{1}_{ij}}{r_{ij}^3} \right] \\ &\equiv \hat{H}^{\text{SO}(1)} + \hat{H}^{\text{SO}(2)}. \end{aligned} \quad (55)$$

The corrections to the \hat{H}^{SO} operator, Eq. 55, which account for the presence of magnetic fields, arise through the angular momentum operators (Fukui et al. 1996):

$$\mathbf{l}_{iK} = \mathbf{r}_{iK} \times [-i\hbar\nabla_i + e\mathbf{A}(\mathbf{r}_i)] \quad (56)$$

$$\mathbf{l}_{ij} = \mathbf{r}_{ij} \times [-i\hbar\nabla_i + e\mathbf{A}(\mathbf{r}_i)]. \quad (57)$$

The first part of the vector potential, related to the external magnetic field, causes the \hat{H}^{SO} operator to contain two contributions

$$\hat{H}^{\text{SO,B}} = \hat{H}^{\text{SO,B(1)}} + \hat{H}^{\text{SO,B(2)}}, \quad (58)$$

$$\hat{H}^{\text{SO,B(1)}} = -\frac{e^2}{16\pi\epsilon_0 m_e} \alpha_{\text{fs}}^2 \left[\sum_{iK} Z_K \mathbf{m}_i \cdot \frac{(\mathbf{r}_{iO} \cdot \mathbf{r}_{iK}) \mathbf{1} - \mathbf{r}_{iO} \mathbf{r}_{iK}^T}{r_{iK}^3} \cdot \mathbf{B} \right], \quad (59)$$

$$\hat{H}^{\text{SO,B(2)}} = \frac{e^2}{16\pi\epsilon_0 m_e} \alpha_{\text{fs}}^2 \left[\sum_{i \neq j} (\mathbf{m}_i + 2\mathbf{m}_j) \cdot \frac{(\mathbf{r}_{iO} \cdot \mathbf{r}_{ij}) \mathbf{1} - \mathbf{r}_{iO} \mathbf{r}_{ij}^T}{r_{ij}^3} \cdot \mathbf{B} \right]. \quad (60)$$

Similarly, the second part of the vector potential related to the magnetic moment of nucleus K leads to

$$\hat{H}^{\text{SO,K}} = \hat{H}^{\text{SO,K(1)}} + \hat{H}^{\text{SO,K(2)}}, \quad (61)$$

$$\hat{H}^{\text{SO,K(1)}} = -\frac{e^2}{8\pi\epsilon_0 m_e} \alpha_{\text{fs}}^4 \left[\sum_M Z_M \sum_i \mathbf{m}_i \cdot \frac{(\mathbf{r}_{iK} \cdot \mathbf{r}_{iM}) \mathbf{1} - \mathbf{r}_{iK} \mathbf{r}_{iM}^T}{r_{iK}^3 r_{iM}^3} \cdot \mathbf{m}_K \right], \quad (62)$$

$$\hat{H}^{\text{SO,K(2)}} = \frac{e^2}{8\pi\epsilon_0 m_e} \alpha_{\text{fs}}^4 \left[\sum_{i \neq j} (\mathbf{m}_i + 2\mathbf{m}_j) \cdot \frac{(\mathbf{r}_{iK} \cdot \mathbf{r}_{ij}) \mathbf{1} - \mathbf{r}_{iK} \mathbf{r}_{ij}^T}{r_{iK}^3 r_{ij}^3} \cdot \mathbf{m}_K \right]. \quad (63)$$

Response Theory

Molecular Response: Definitions, Symbols

The formalism we use to compare theory with experiment is based on the analysis of the expectation value of some operator of interest. We denote the time-dependent reference state as $|\tilde{0}(t)\rangle$ and consider hermitian (in general time-dependent) perturbing operators. The response functions correspond to the Fourier coefficients in the expansion of the time-dependent expectation value of an operator A (Olsen and Jørgensen 1995):

$$\begin{aligned}
\langle \tilde{0}(t) | A | \tilde{0}(t) \rangle &= \langle 0 | A | 0 \rangle + \int_{-\infty}^{\infty} \langle \langle A; V^{\omega_1} \rangle \rangle e^{-i\omega_1 t} d\omega_1 \\
&+ \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \langle A; V^{\omega_1}, V^{\omega_2} \rangle \rangle e^{-i(\omega_1 + \omega_2)t} d\omega_1 d\omega_2 \\
&+ \frac{1}{6} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \langle A; V^{\omega_1}, V^{\omega_2}, V^{\omega_3} \rangle \rangle e^{-i(\omega_1 + \omega_2 + \omega_3)t} d\omega_1 d\omega_2 d\omega_3 + \dots,
\end{aligned} \tag{64}$$

where each V^{ω_i} represents a time-dependent perturbation operator. We have omitted the infinitesimal factors, which ensure that the perturbations are switched on adiabatically and vanish for $t = -\infty$.

In the analysis of the different terms in Eq. 64 for the case of exact electronic states, the zeroth-order response function is the expectation value of the operator A in the reference state $|0\rangle$, $\langle 0 | A | 0 \rangle$. The first-order, or linear, response function and its residue are given in the sum-over-states form by Olsen and Jørgensen (1995) and Sasagane et al. (1993):

$$\begin{aligned}
\langle \langle A; B \rangle \rangle_{\omega_1} &= -\frac{1}{\hbar} \sum_P \sum_n \frac{\langle 0 | A | n \rangle \langle n | B | 0 \rangle}{\omega_n - \omega_1} \\
&= -\frac{1}{\hbar} \sum_n \left[\frac{\langle 0 | A | n \rangle \langle n | B | 0 \rangle}{\omega_n - \omega_1} + \frac{\langle 0 | B | n \rangle \langle n | A | 0 \rangle}{\omega_n + \omega_1} \right].
\end{aligned} \tag{65}$$

The poles correspond to the excitation energies of the system. In addition, from the residues:

$$\lim_{\omega_1 \rightarrow \omega_e} \hbar (\omega_1 - \omega_e) \langle \langle A; B \rangle \rangle_{\omega_1} = \langle 0 | A | e \rangle \langle e | B | 0 \rangle, \tag{66}$$

$$\lim_{\omega_1 \rightarrow -\omega_e} \hbar (\omega_1 + \omega_e) \langle \langle A; B \rangle \rangle_{\omega_1} = -\langle 0 | B | e \rangle \langle e | A | 0 \rangle, \tag{67}$$

we can extract information about the transition matrix elements. The second-order, or quadratic, response function is defined as

$$\langle \langle A; B, C \rangle \rangle_{\omega_1, \omega_2} = -\frac{1}{\hbar^2} \sum_P \sum_{m, n} \frac{\langle 0 | A | m \rangle \langle m | B | n \rangle \langle n | C | 0 \rangle}{(\omega_m - \omega_\sigma) (\omega_n - \omega_2)}. \tag{68}$$

In Eq. 68 and in the expressions given below, ω_σ equals the sum of the other frequencies ($\omega_\sigma = \sum_i \omega_i$) and P permutes the operators and the associated frequencies. For example, in Eq. 68, P permutes $A(-\omega_\sigma)$, $B(\omega_1)$, $C(\omega_2)$. We note that different expressions for nonlinear response functions, involving operators of the form $\bar{B} = B - \langle 0 | B | 0 \rangle$, are obtained when the state $|0\rangle$ is omitted from the summations (see, e.g., Sasagane et al. 1993).

For the quadratic and cubic response functions there is a variety of residues, and we present only expressions for the residues that are of practical interest. A single residue of a quadratic response function is

$$\lim_{\omega_2 \rightarrow \omega_f} \hbar (\omega_2 - \omega_f) \langle\langle A; B, C \rangle\rangle_{\omega_1, \omega_2} = -\frac{1}{\hbar} \left[\sum_P \sum_m \frac{\langle 0 | A | m \rangle \langle m | B | f \rangle}{(\omega_m - \omega_\sigma)} \right] \langle f | C | 0 \rangle. \quad (69)$$

This type of residue is used for instance to evaluate two-photon transition amplitudes, the B term in magnetic circular dichroism or the transition moments in phosphorescence. A double residue of the quadratic response function, given by

$$\begin{aligned} \lim_{\omega_1 \rightarrow -\omega_f} \hbar (\omega_1 + \omega_f) \lim_{\omega_2 \rightarrow \omega_e} \hbar (\omega_2 - \omega_e) \langle\langle A; B, C \rangle\rangle_{\omega_1, \omega_2} \\ = -\langle 0 | B | f \rangle \langle f | (A - \langle 0 | A | 0 \rangle) | e \rangle \langle e | C | 0 \rangle, \end{aligned} \quad (70)$$

allows us to determine (from a knowledge of the reference state $|0\rangle$ only) the transition moment between two other states, $|e\rangle$ and $|f\rangle$.

For the particular choice of $|e\rangle = |f\rangle$, we obtain

$$\begin{aligned} \lim_{\omega_1 \rightarrow -\omega_e} \hbar (\omega_1 + \omega_e) \lim_{\omega_2 \rightarrow \omega_e} \hbar (\omega_2 - \omega_e) \langle\langle A; B, C \rangle\rangle_{\omega_1, \omega_2} \\ = -\langle 0 | B | e \rangle (\langle e | A | e \rangle - \langle 0 | A | 0 \rangle) \langle e | C | 0 \rangle, \end{aligned} \quad (71)$$

which allows us to extract the expectation value $\langle e | A | e \rangle$. In this way, using a double residue of the quadratic response function from the reference state function we can compute first-order properties of another state.

Finally, the third-order, or cubic, response function can be expressed as

$$\langle\langle A; B, C, D \rangle\rangle_{\omega_1, \omega_2, \omega_3} = -\frac{1}{\hbar^3} \sum_P \sum_{m, n, q} \frac{\langle 0 | A | m \rangle \langle m | B | n \rangle \langle n | C | q \rangle \langle q | D | 0 \rangle}{(\omega_m - \omega_\sigma) (\omega_n - \omega_2 - \omega_3) (\omega_q - \omega_3)}. \quad (72)$$

From the discussed residues, we can determine properties explicitly dependent on different electronic states, for instance all the transition moments coupling the reference state $|0\rangle$ and other states $|e\rangle$ (from linear response residues) and all the one-photon transition moments coupling other pairs of states (from quadratic response double residues). In addition, from single residues of the second- and third-order response functions we can identify two- and three-photon transition matrix elements.

The frequency(time)-independent properties are a special case of the response functions ($\omega_1 = \omega_2 \cdots = 0$) and correspond to energy derivatives. We shall use a simplified notation omitting the subscripts indicating the (zero) frequencies for these properties. The response function, is therefore, in this case written as $\langle\langle A; B, C, D \dots \rangle\rangle$.

For time-dependent oscillatory perturbations, the response functions can be related to the derivatives of the time-averaged quasienergy (Christiansen et al. 1998; Saue 2002). The derivation, based on a time-dependent variational principle and the time-dependent Hellman–Feynman theorem, shows explicitly the symmetry of the response functions with respect to the permutation of all the operators.

We do not describe in this chapter the standard basis sets and wave function models in use in quantum chemistry, referring instead the interested reader to other chapters in this handbook and to a recent textbook (Helgaker et al. 2000). We only note that response theory has been formulated and corresponding programs have been implemented for most of the standard approximations, such as:

- RHF, the restricted Hartree–Fock model
- MCSCF, the multiconfigurational self-consistent field model
- DFT, density-functional theory
- MP2, the Møller–Plesset second-order perturbation theory, which in the case of frequency-dependent perturbations has to be replaced by an iterative optimization of the perturbation amplitudes, as described by the CC2 model (Christiansen et al. 1995)
- CCSD, the coupled-cluster singles-and-doubles model
- CCSD(T), CCSD with a noniterative perturbative treatment of the triples correction, which in the case of frequency-dependent response functions has to be replaced by an iterative triples correction, as described by the CC3 model (Koch et al. 1997)
- FCI, full configuration-interaction theory

There are also a couple of models designed primarily for the purpose of calculating response functions – that is, they cannot be derived in terms of energy or quasienergy derivatives of the energy expression for an electronic structure method. We note here the second-order polarization propagator approximation [SOPPA (Enevoldsen et al. 1998; Nielsen et al. 1980)], SOPPA using CCSD correlation amplitudes [SOPPA(CCSD) (Sauer 1997)], as well as the equation-of-motion CCSD model [EOM-CCSD (Perera et al. 1994a; Stanton and Bartlett 1993)].

In general, the accuracy of the computed molecular properties reflects the accuracy of the approach used for the unperturbed reference state. There is, however, one significant difference between variational and perturbative, nonvariational approximations which should be kept in mind in the calculation of properties. The response equations which hold for the exact wave function are in general not fulfilled for nonvariational wave function models, even in the limit of a complete basis set. For instance, in the MP2 and coupled-cluster approximations, the static first-order properties, which should be computed as the first energy derivatives, do not correspond to simple expectation values – corrections arising from orbital relaxation must in general be taken into account. Similarly, in the calculation of higher-order molecular properties when the basis set depends explicitly on the applied perturbations (such as geometrical distortions, or magnetic field perturbations described using London atomic orbitals), orbital relaxation corrections should

be included (Gauss and Stanton 1995). We note, however, that as the excitation level is increased (triple excitations, etc.), the importance of orbital relaxation is reduced, since in the limit of the full CI wave function, which is exact, no orbital relaxation is necessary.

Expansions of Energy and Multipole Moments

The static properties can be defined by an expansion of the total energy

$$\begin{aligned}
 E = E^{(0)} &- \mu_\alpha F_\alpha - \frac{1}{2} \alpha_{\alpha\beta} F_\alpha F_\beta - \frac{1}{6} \beta_{\alpha\beta\gamma} F_\alpha F_\beta F_\gamma - \frac{1}{24} \gamma_{\alpha\beta\gamma\delta} F_\alpha F_\beta F_\gamma F_\delta + \dots \\
 &- \frac{1}{3} \Theta_{\alpha\beta} F_{\alpha\beta} - \frac{1}{3} A_{\alpha,\beta\gamma} F_\alpha F_{\beta\gamma} - \frac{1}{6} C_{\alpha\beta,\gamma\delta} F_{\alpha\beta} F_{\gamma\delta} + \dots \\
 &- \frac{1}{2} \xi_{\alpha\beta} B_\alpha B_\beta - \frac{1}{2} \zeta_{\gamma,\alpha\beta} B_\alpha B_\beta F_\gamma - \frac{1}{4} \eta_{\gamma\delta,\alpha\beta} B_\alpha B_\beta F_\gamma F_\delta + \dots \\
 &+ \sigma_{K,\alpha\beta} B_\alpha m_{K,\beta} + \dots
 \end{aligned} \tag{73}$$

For a molecule in a radiation field, the definitions of the frequency-dependent properties are related to the induced (in general also frequency-dependent) multipole moments. Let us consider a small harmonic perturbation of angular frequency ω from a plane-wave radiation field. We then have (Barron 2004):

- The induced electric dipole moment μ

$$\mu_\alpha = \alpha_{\alpha\beta} F_\beta + \frac{1}{\omega} G'_{\alpha\beta} \dot{B}_\beta + \frac{1}{3} A_{\alpha,\beta\gamma} F_{\beta\gamma} + \dots \tag{74}$$

$$+ \frac{1}{\omega} \alpha'_{\alpha\beta} \dot{F}_\beta + G_{\alpha\beta} B_\beta + \frac{1}{3\omega} A'_{\alpha,\beta\gamma} \dot{F}_{\beta\gamma} + \dots, \tag{75}$$

- The induced electric quadrupole moment Θ

$$\Theta_{\alpha\beta} = A_{\gamma,\alpha\beta} F_\gamma - \frac{1}{\omega} D'_{\gamma,\alpha\beta} \dot{B}_\gamma + C_{\alpha\beta,\gamma\delta} F_{\gamma\delta} + \dots \tag{76}$$

$$- \frac{1}{\omega} A'_{\gamma,\alpha\beta} \dot{F}_\gamma + D_{\gamma,\alpha\beta} B_\gamma + \frac{1}{\omega} C'_{\alpha\beta,\gamma\delta} \dot{F}_{\gamma\delta} + \dots, \tag{77}$$

- The induced magnetic dipole moment m'

$$m'_\alpha = \xi_{\alpha\beta} B_\beta + \frac{1}{3\omega} D'_{\alpha,\beta\gamma} \dot{F}_{\beta\gamma} - \frac{1}{\omega} G'_{\beta\alpha} \dot{F}_\beta + \dots \tag{78}$$

$$+ \frac{1}{\omega} \xi'_{\alpha\beta} \dot{B}_\beta + \frac{1}{3} D_{\alpha,\beta\gamma} F_{\beta\gamma} + G_{\beta\alpha} F_\beta + \dots, \tag{79}$$

where \dot{F} and \dot{B} denote the time derivatives of the fields and we have considered only linear response properties.

The properties we have included are the electric multipole polarizabilities – $\alpha_{\alpha\beta}$, dipole–dipole, $A_{\alpha, \beta\gamma}$, dipole–quadrupole; and $C_{\alpha\beta, \gamma\delta}$, quadrupole–quadrupole – as well as the mixed polarizabilities – $G_{\alpha\beta}$, electric dipole-magnetic dipole, $D_{\gamma, \alpha\beta}$, magnetic dipole-electric quadrupole and the magnetizability (magnetic dipole-magnetic dipole polarizability) $\xi_{\alpha\beta}$. The frequency-dependent response properties labeled with a prime vanish for $\omega = 0$, and the properties specified in the second rows of the equations (Eqs. 75, 77, 79) are zero for reference states described by real wave functions (in practice, closed-shell nondegenerate systems).

We note that in the expansion of the induced dipole moment, higher-order corrections may also be important, in particular in connection with electric-field perturbations. For instance, considering only electric dipole contributions, it is customary to also include contributions to μ_{α} arising from frequency-dependent hyperpolarizabilities (see Eq. 84 below).

Electric Properties

In this section we consider molecular properties which characterize the interactions with static and/or frequency-dependent electric fields. The electric properties of a molecule determine the electric properties of the bulk sample, such as the relative permittivity (dielectric constant) and the refractive index. In addition, the electric properties can be used to describe intermolecular forces.

Only static and dynamic molecular properties involving electric dipole and quadrupole operators will be discussed below. However, electric properties related to higher-order electric multipole operators can also be determined in a similar manner to the properties described here, in terms of expectation values, linear and nonlinear response functions. Nevertheless, it should be kept in mind that although the same formalism is applied in the calculation of response functions involving octupole, hexadecapole, and higher moments, in practice it may be difficult to obtain accurate results for these moments and related molecular properties. For instance, significantly larger basis sets with more polarization and diffuse functions may be required than for the properties involving lower-order electric moments.

Most theoretical results are obtained in atomic units (a.u.), whereas experimental data often are given in SI or esu units. In Table 1 some useful conversion factors are given.

Electric Multipole Moments

The permanent multipole moments characterize the charge distribution in a molecule. The zeroth-order (monopole) moment corresponds to the total charge, the first moment to the dipole, the second to the quadrupole, and so on. The multipole

Table 1 Unit conversion factors for selected electric properties

	1 a.u.	SI		esu	
F	$E_h e^{-1} a_0^{-1}$	$5.142\,206 \times 10^{11}$	V m^{-1}	$1.715\,3 \times 10^7$	statvolt cm^{-1}
μ	ea_0	$8.478\,353 \times 10^{-30}$	C m	$2.541\,7 \times 10^{-18}$	statvolt cm^{2a}
Θ	ea_0^2	$4.486\,551 \times 10^{-40}$	C m^2	$1.345\,0 \times 10^{-26}$	statvolt cm^3
α	$e^2 a_0^2 E_h^{-1}$	$1.648\,777 \times 10^{-41}$	$\text{C}^2 \text{ m}^2 \text{ J}^{-1}$	$1.481\,8 \times 10^{-25}$	cm^3
β	$e^3 a_0^3 E_h^{-2}$	$3.206\,362 \times 10^{-53} \text{ n}$	$\text{C}^3 \text{ m}^3 \text{ J}^{-2}$	$8.639\,2 \times 10^{-33}$	statvolt $^{-1}$ cm^4
γ	$e^4 a_0^4 E_h^{-3}$	$6.235\,381 \times 10^{-65}$	$\text{C}^4 \text{ m}^4 \text{ J}^{-3}$	$5.036\,7 \times 10^{-40}$	statvolt $^{-2}$ cm^5
V^K	$E_h e^{-1} a_0^{-2}$	$9.717\,365 \times 10^{21}$	V m^{-2}	$3.241\,4 \times 10^{15}$	statvolt cm^{-2b}

^a 1Debye = 10^{-18} statvolt cm^2

^bThe nuclear quadrupole moments, needed to compare the calculated NQCC with experiment (section “[Electric Field Gradient at the Nucleus, Nuclear Quadrupole Coupling Constant](#)”), are usually given in barn; 1b = 10^{-28} m^2

moments can be computed as expectation values of the multipole moment operators discussed in section “[Small Terms Due to the Scalar Potential in the Hamiltonian.](#)” As for the operators, we can use the primitive or traceless forms for the quadrupole and higher-order moments. The traceless multipole moments describe the deviation from an isotropic charge distribution. For instance, the traceless quadrupole moment vanishes for a spherical charge distribution when the center of the sphere coincides with the origin, since the transformation from the primitive (traced) quadrupole moment $q_{\alpha\beta}$ to the traceless quadrupole moment $\Theta_{\alpha\beta}$ eliminates the isotropic contribution. The traceless multipole moment of order n is defined by $2n + 1$ components, whereas the traced tensor involves $(n + 2)(n + 1)/2$ independent quantities. A tabulation of nonzero symmetry-independent elements of multipole moments (μ = dipole, Θ = quadrupole, Ω = octupole, and Φ = hexadecapole) for different point groups is given by Kielich (1972).

In general, when the multipole expansion of the potential is used to describe the influence of a static electric field, only the first nonzero moment is unambiguously defined. The next, higher-order multipole moments depend on the choice of origin, and when computed with respect to the origin of the coordinate system, they depend on the position of the molecule in space. This can be seen from Eq. 30, where a shift of the chosen common origin for the expansion of the potential by a vector \mathbf{D} for a molecule with the total charge Q leads to a change of $Q\mathbf{D}$ in the dipole moment; similarly, for a neutral molecule with a nonzero dipole moment, the quadrupole moment will change.

Dipole and Quadrupole Moments

From the multipole expansion of the energy in the presence of a static electric field, Eq. 73, we find that the α -component of the static permanent dipole moment is defined as the electric field derivative of $E(\mathbf{F})$

$$\mu_\alpha = - \left. \frac{\partial E(\mathbf{F})}{\partial F_\alpha} \right|_{\mathbf{F}=\mathbf{0}}. \quad (80)$$

For wave functions that fulfill the Hellman–Feynman theorem we can compute the dipole moment as the expectation value

$$\mu_\alpha = \langle 0 | \hat{\mu}_\alpha | 0 \rangle, \quad (81)$$

where $\hat{\mu}_\alpha$ is the α -component of the dipole moment operator defined in Eq. 30. Obviously, the expectation value expression is correct for the exact wave function. Furthermore, for most approximate wave functions it is much easier to compute the expectation value than to calculate the energy derivative. However, the results differ when the method used does not fulfill the Hellman–Feynman theorem. In this case, one can either derive the analytic expressions which reproduce the correct definition of the electric field derivative of $E(\mathbf{F})$ (see for instance the discussion of the coupled cluster approach in Helgaker et al. 2000) or use the finite field method and numerical differentiation.

The finite-field method appears to be very simple and has been often applied, but it has some drawbacks. First, the values of $E(\mathbf{F})$ depend not only on the dipole moment, but also on the polarizability and hyperpolarizabilities, and high accuracy of the computed values of $E(\mathbf{F})$ is required to perform the numerical differentiation accurately. Moreover, the presence of the finite electric field may lower the symmetry of the system, making the calculations much more time-consuming.

Figure 1, taken from Bak et al. (2000), illustrates the accuracy of equilibrium geometry dipole moments calculated for a series of small molecules. In Bak et al. (2000), the dipole moments of 11 molecules were studied at various levels of approximation using a sequence of basis sets, with the CCSD(T)/aug-cc-pVQZ results providing the benchmark values. The normal distribution of errors with respect to the reference values, shown in Fig. 1, demonstrates the systematic improvements of the computed results with increasing basis set and improved treatment of electron correlation effects.

The expectation value of $\hat{\Theta}$, that is, the traceless electric quadrupole operator as defined in Eq. 32 following Buckingham (1967),

$$\Theta_{\alpha\beta} = \langle 0 | \hat{\Theta}_{\alpha\beta} | 0 \rangle = -3 \frac{\partial E(\nabla\mathbf{F})}{\partial F_{\alpha\beta}}, \quad (82)$$

can be related to the expectation values of second moments, $q_{\alpha\beta} = \langle 0 | \hat{q}_{\alpha\beta} | 0 \rangle$, by a simple linear transformation, the same as that connecting the corresponding operators, Eqs. 32 and 33. Both Θ and \mathbf{q} are rank two, symmetric tensors.

For molecules with an n -fold rotation axis with $n \geq 3$:

$$\Theta \equiv \Theta_{zz} = -2\Theta_{yy} = -2\Theta_{xx}, \quad (83)$$

where z coincides with the rotation axis. In experiment, two diagonal components of Θ are measured.

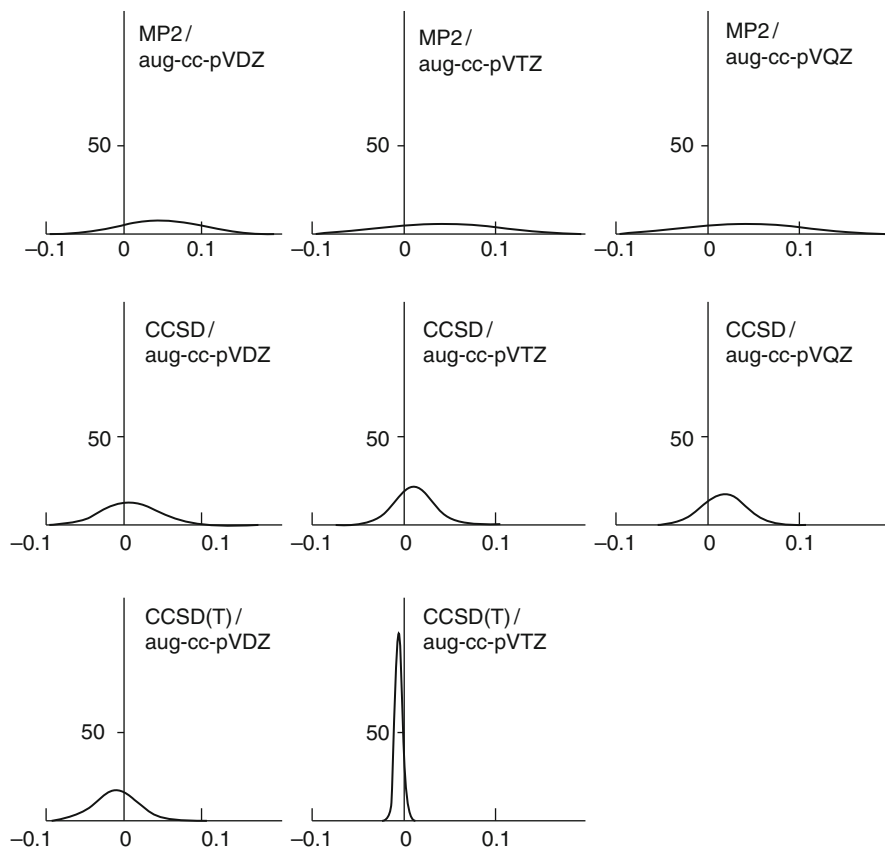


Fig. 1 Correlation and basis set dependence of the dipole moment – normal distribution of errors (in Debye) (Reproduced by permission from Bak et al. (2000))

The Origin of the Frequency-Dependent Electric Dipole Polarizability and Hyperpolarizabilities

The electric dipole moment $\tilde{\mu}$ in the presence of an applied spatially homogeneous non-monochromatic electric field $\sum_i \mathbf{F}^{\omega_i}$ (where ω_i indicates the circular frequency) can be written as

$$\begin{aligned}
 \tilde{\mu}_\alpha^{\omega_\sigma} &= \mu_\alpha \delta(\omega_\sigma = 0) + \alpha_{\alpha\beta}(-\omega_\sigma; \omega_1) F_\beta(\omega_1) \\
 &+ \frac{1}{2} \sum_{\omega_1, \omega_2} \beta_{\alpha\beta\gamma}(-\omega_\sigma; \omega_1, \omega_2) F_\beta(\omega_1) F_\gamma(\omega_2) \\
 &+ \frac{1}{6} \sum_{\omega_1, \omega_2, \omega_3} \gamma_{\alpha\beta\gamma\delta}(-\omega_\sigma; \omega_1, \omega_2, \omega_3) F_\beta(\omega_1) F_\gamma(\omega_2) F_\delta(\omega_3) + \dots,
 \end{aligned} \tag{84}$$

where μ_α is the unperturbed electric dipole moment. In this phenomenological expansion (which we use to describe elastic light scattering processes) the summations over optical frequencies will be restricted in order to preserve a constant ω_σ equal to the sum of optical frequencies

$$\omega_\sigma = \sum_i \omega_i. \quad (85)$$

Equation 84 defines the components of the dynamic electric dipole polarizability tensor, $\alpha_{\mu\nu}(-\omega_\sigma; \omega_1)$, those of the first electric dipole hyperpolarizability, $\beta_{\mu\nu\eta}(-\omega_\sigma; \omega_1, \omega_2)$, and those of the second electric dipole hyperpolarizability $\gamma_{\mu\nu\eta\xi}(-\omega_\sigma; \omega_1, \omega_2, \omega_3)$. The combination of frequencies and the presence or absence of static fields characterize various nonlinear processes. The calculations can be done for any combination of frequencies satisfying Eq. 85, but only selected examples of the most commonly studied optical phenomena will be discussed below (see, e.g., Willets et al. 1992 and Bogaard and Orr 1975 for more details).

Figure 2 shows how some of these processes can be interpreted in terms of electronic transitions between ground and excited states. The excited states can be treated as virtual states, but if one of the excitation energies is close to the frequency of the light, the contribution of that state is often dominant. Simplified models that include in the analysis only a few excited states can be used when these computed states describe well the real, physical states of the molecule which are relevant for the property studied. However, it is important to emphasize that even though these few-states models in many cases may give qualitative insight into the electronic processes, the convergence of the sum-over-states expansion for the (hyper)polarizabilities (see Eqs. 65, 68, and 72) with respect to the number of excited states included in the summation is in general very slow (Wiberg et al. 2006). We should also keep in mind that if an excitation energy equals a frequency, we have a very different situation in which we have to consider the finite lifetime of the excited state and also deal with the absorption process.

In experiment, the molecule is often perturbed by a combination of a static and a dynamic electric field. We then have, assuming a monochromatic dynamic field,

$$F_\alpha(t) = F_\alpha^0 + F_\alpha^\omega \cos(\omega t), \quad (86)$$

and therefore the time dependence of the polarization can be described as

$$\mu_\alpha(t) = \mu_\alpha^0 + \mu_\alpha^\omega \cos(\omega t) + \mu_\alpha^{2\omega} \cos(2\omega t) + \mu_\alpha^{3\omega} \cos(3\omega t) + \dots, \quad (87)$$

where $\mu_\alpha^{n\omega}$ are the Fourier amplitudes to be determined. In order to compare the measured polarization at a certain frequency to the theoretical expressions, we insert Eq. 86 into Eq. 84 and use the relations between $\cos^n(\omega t)$ and $\cos(n\omega t)$. We obtain

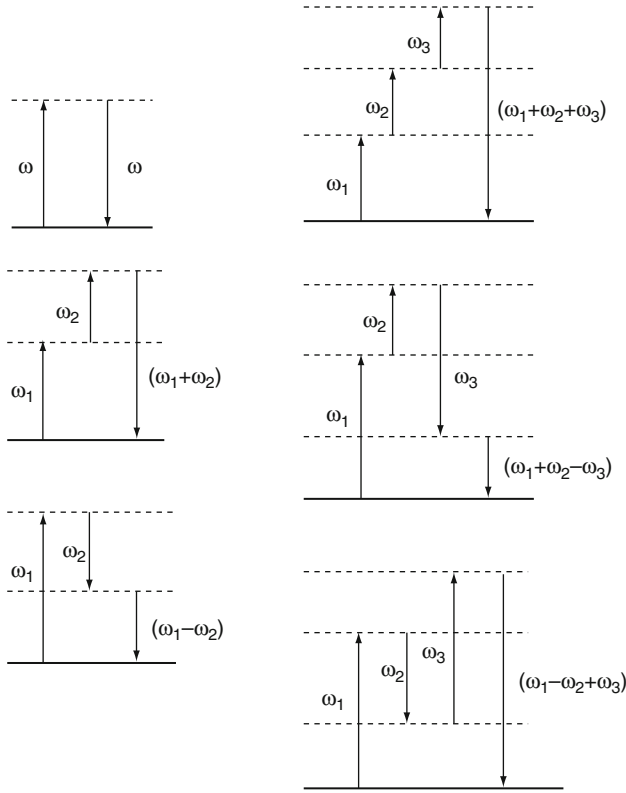


Fig. 2 Linear and nonlinear optical properties – a schematic illustration of the involved transitions for (left, top to bottom): $\alpha(-\omega; \omega)$; $\beta(-(\omega_1 + \omega_2); \omega_1, \omega_2)$; and $\beta(-(\omega_1 - \omega_2); \omega_1, -\omega_2)$; (right, top to bottom) $\gamma(-(\omega_1 + \omega_2 + \omega_3); \omega_1, \omega_2, \omega_3)$; $\gamma(-(\omega_1 + \omega_2 - \omega_3); \omega_1, \omega_2, -\omega_3)$, and $\gamma(-(\omega_1 - \omega_2 + \omega_3); \omega_1, -\omega_2, \omega_3)$

$$\begin{aligned} \mu_\alpha^0 &= \mu_\alpha + \alpha_{\alpha\beta}(0; 0) F_\beta^0 + \frac{1}{2}\beta_{\alpha\beta\gamma}(0; 0, 0) F_\beta^0 F_\gamma^0 + \frac{1}{8}\gamma_{\alpha\beta\gamma\delta}(0; 0, 0, 0) F_\beta^0 F_\gamma^0 F_\delta^0 \\ &+ \frac{1}{4}\beta_{\alpha\beta\gamma}(0; -\omega, \omega) F_\beta^\omega F_\gamma^\omega + \frac{1}{4}\gamma_{\alpha\beta\gamma\delta}(0; 0, -\omega, \omega) F_\beta^0 F_\gamma^\omega F_\delta^\omega, \end{aligned} \quad (88)$$

$$\begin{aligned} \mu_\alpha^\omega &= \alpha_{\alpha\beta}(-\omega; \omega) F_\beta^\omega + \beta_{\alpha\beta\gamma}(-\omega; \omega, 0) F_\beta^\omega F_\gamma^0 \\ &+ \frac{1}{2}\gamma_{\alpha\beta\gamma\delta}(-\omega; \omega, 0, 0) F_\beta^\omega F_\gamma^0 F_\delta^0 + \frac{1}{8}\gamma_{\alpha\beta\gamma\delta}(-\omega; \omega, -\omega, \omega) F_\beta^\omega F_\gamma^\omega F_\delta^\omega, \end{aligned} \quad (89)$$

$$\mu_\alpha^{2\omega} = \frac{1}{4}\beta_{\alpha\beta\gamma}(-2\omega; \omega, \omega) F_\beta^\omega F_\gamma^\omega + \frac{1}{4}\gamma_{\alpha\beta\gamma\delta}(-2\omega; \omega, \omega, 0) F_\beta^\omega F_\gamma^\omega F_\delta^0, \quad (90)$$

$$\mu_\alpha^{3\omega} = \frac{1}{24}\gamma_{\alpha\beta\gamma\delta}(-3\omega; \omega, \omega, \omega) F_\beta^\omega F_\gamma^\omega F_\delta^\omega, \quad (91)$$

where we have truncated the expansions at $\cos(3\omega t)$. Using the expansions given above and the expansion in Eq. 84 we can compare the experimentally measured polarization and the theoretical estimates.

Dipole Polarizability

The frequency-dependent electric dipole polarizability is, next to the dipole moment, the most important effect characterizing the response of a molecule to electromagnetic radiation. It is a second-order symmetric tensor determined by the linear response function

$$\alpha_{\alpha\beta}(-\omega; \omega) = -\langle\langle \hat{\mu}_\alpha; \hat{\mu}_\beta \rangle\rangle_\omega. \quad (92)$$

The frequency dependence of the polarizability may be described using the Cauchy moments (section “[Cauchy Moments](#)”).

For the static dipole polarizability $\alpha(0;0)$ we have

$$\alpha_{\alpha\beta}(0;0) = -\left. \frac{\partial^2 E(\mathbf{F})}{\partial F_\alpha \partial F_\beta} \right|_{\mathbf{F}=\mathbf{0}}, \quad (93)$$

and the energy derivative can be evaluated using the finite-field approach (i.e., computing numerically the differences between $E(\mathbf{F})$ and $E(0)$).

Numerous molecular properties which describe nonlinear effects, such as the Kerr effect (section “[Second Dipole Hyperpolarizability](#)”) or magnetic circular dichroism (section “[Magnetic Circular Dichroism](#)”), arising in the presence of radiation and additional electric or magnetic fields, are interpreted as derivatives of the dipole polarizability (Michl and Thulstrup 1995). They can be calculated as higher-order response functions. Similarly, relativistic corrections to the polarizabilities for heavy atoms can be estimated from higher-order response functions including the mass-velocity and Darwin operators, Eqs. 9 and 20, as additional perturbations (Kirpekar et al. 1995).

In isotropic fluids, the experimentally measured quantity is usually the scalar component of the α tensor given by the isotropic average, defined as

$$\alpha_{\text{iso}} = \frac{1}{3} \alpha_{\eta\eta}. \quad (94)$$

The polarizability anisotropy may be defined as

$$(\alpha_{\text{anis}})^2 = \frac{(3\alpha_{\mu\nu}\alpha_{\mu\nu} - \alpha_{\mu\mu}\alpha_{\nu\nu})}{2}, \quad (95)$$

which, using the components of the diagonalized tensor, leads to

$$\alpha_{\text{anis}} = 2^{-1/2} \left[(\alpha_{xx} - \alpha_{yy})^2 + (\alpha_{yy} - \alpha_{zz})^2 + (\alpha_{zz} - \alpha_{xx})^2 \right]^{1/2}. \quad (96)$$

For linear molecules, α_{anis} reduces to the difference between the parallel and perpendicular components,

$$\alpha_{\text{anis}} = \alpha_{\parallel} - \alpha_{\perp}. \quad (97)$$

The macroscopic property related to the molecular polarizability α is the dielectric constant ε , defined by the ratio of the permittivity of the medium to the electric constant, ε_0 . It is also related to the refractive index n ($n = \sqrt{\varepsilon}$, assuming the vacuum magnetic permeability to be equal to 1). The quantity measured in a dielectric constant experiment is

$$\tilde{\alpha} = \left(\frac{\tilde{\mu}_Z^F}{F} \right)_{F \rightarrow 0} = \langle \alpha_{II} \rangle_{\Omega} + \frac{\mu_z^2}{3kT}, \quad (98)$$

where we use the notation $\langle \dots \rangle_{\Omega}$ for the average over all the orientations, k is the Boltzmann constant, and T is the temperature. The temperature-dependent term is proportional to the square of the dipole moment, the temperature-independent term is the isotropic average of the polarizability (using Eq. 94 we can obtain a corresponding expression in terms of properties calculated in the molecular frame of reference).

The anisotropy of the electric dipole polarizability can be determined in Rayleigh scattering experiments, where the observed depolarization of the incident light is a function of $(\alpha_{\text{anis}})^2$ (Bonin and Kresin 1997).

The static electric dipole polarizability of the He atom has been studied in more detail than any other response property of any other atomic or molecular system. Not only is the nonrelativistic value known accurately, and the relativistic corrections computed, but – in a recent benchmark calculation – the quantum-electrodynamic corrections have also been analyzed and numerous effects related to the nuclear mass have been taken into account (Łach et al. 2004).

Standard energy-optimized basis sets are not suitable for accurate calculations of electric polarizabilities. The simplest solution – adding the necessary polarization and diffuse functions – makes the basis sets too large to enable efficient calculations for large molecules. Significantly smaller basis sets, designed considering the electric-field dependence of the orbitals (Benkova et al. 2005; Sadlej 1988), provide results of similar or better quality at lower computational cost.

A practical problem that may be encountered in calculations using very extended basis sets is the appearance of linear dependencies in the basis set. This may make the response equations (as well as the equations determining the molecular energy/density) ill-conditioned and the calculations slowly convergent; in exceptional cases it may even be impossible to converge the equations. The problem is normally resolved by removing the linear dependencies in the basis set, or by removing manually the most diffuse basis functions from the basis set.

Electron correlation effects on the isotropic polarizability are in general moderate, being at the most 5–10 % of the uncorrelated value and for nondipolar systems

always leading to an increased polarizability. Correlation effects are usually larger on the individual tensor components than on the isotropic polarizability, and can, therefore, be expected to be more important for the polarizability anisotropies. In general, Hartree–Fock polarizability anisotropies cannot be considered accurate enough to allow for a quantitative comparison with experimental observations, or for the interpretation of the ellipticity arising in some birefringences, where the polarizability anisotropies often constitute an important (and dominant) orientational contribution.

DFT in general reproduces quite well correlation effects in the polarizability of small molecular systems. In particular, polarizability anisotropies calculated using DFT will usually be in better agreement with experiment. However, due to the local nature of the exchange–correlation functionals and the lack of general implementations of current density functionals, DFT does not perform so well for extended conjugated systems (Champagne et al. 2000). We will return briefly to this point when discussing calculations of hyperpolarizabilities.

First Dipole Hyperpolarizability

The first electric dipole hyperpolarizability, given by the quadratic response function

$$\beta_{\alpha\beta\gamma}(-\omega_{\sigma}; \omega_1, \omega_2) = -\langle\langle \hat{\mu}_{\alpha}; \hat{\mu}_{\beta}, \hat{\mu}_{\gamma} \rangle\rangle_{\omega_1, \omega_2}, \quad (99)$$

has physical meaning only when $\omega_{\sigma} = \omega_1 + \omega_2$.

The static dipole hyperpolarizability $\beta(0;0,0)$ can be obtained evaluating the third derivative of the energy. When we consider dynamic fields, the processes of particular interest are:

Second Harmonic Generation (SHG)	$\beta_{\mu\nu\eta}^{\text{SHG}}(-2\omega; \omega, \omega) = -\langle\langle \hat{\mu}_{\mu}; \hat{\mu}_{\nu}, \hat{\mu}_{\eta} \rangle\rangle_{\omega, \omega},$
dc-Pockels Effect (dc-P) or	
Electro-Optic Pockels Effect (EOPE)	$\beta_{\mu\nu\eta}^{\text{EOPE}}(-\omega; \omega, 0) = -\langle\langle \hat{\mu}_{\mu}; \hat{\mu}_{\nu}, \hat{\mu}_{\eta} \rangle\rangle_{\omega, 0},$
Optical Rectification (OR)	$\beta_{\mu\nu\eta}^{\text{OR}}(0; \omega, -\omega) = -\langle\langle \hat{\mu}_{\mu}; \hat{\mu}_{\nu}, \hat{\mu}_{\eta} \rangle\rangle_{\omega, -\omega}.$

According to Eqs. 88, 89, 90, and 91, in these processes we deal with a perturbation of frequency ω (and a static field in the Pockels effect) and we consider the resulting dipole moment oscillating with frequency 2ω , frequency ω or static, respectively.

The general definition of the quadratic response function, Eq. 68, indicates its symmetry with respect to permutation of operators. Thus, for all the dipole hyperpolarizabilities we have

$$\beta_{\mu\nu\eta}(-\omega_{\sigma}; \omega_1, \omega_2) = \beta_{\mu\eta\nu}(-\omega_{\sigma}; \omega_2, \omega_1). \quad (100)$$

Also the relation

$$\beta_{\mu\nu\eta}(-\omega_\sigma; \omega_1, \omega_2) = \beta_{\nu\mu\eta}(\omega_1; -\omega_\sigma, \omega_2) \quad (101)$$

holds, and we have, for instance, $\beta_{\mu\nu\eta}^{\text{OR}}(0; \omega, -\omega) = \beta_{\eta\nu\mu}^{\text{EOPE}}(-\omega; \omega, 0)$. We note that this relation does not hold for macroscopic samples due to the differences in the local fields experienced by the molecule in these two different experimental setups.

In the case of the first hyperpolarizability, the measured quantity is the vector component of β in the direction of the permanent dipole moment μ , defining the molecular z axis. The relevant averages are given by:

$$\beta_{\parallel} = \frac{1}{5} (\beta_{z\eta\eta} + \beta_{\eta z\eta} + \beta_{\eta\eta z}), \quad (102)$$

$$\beta_{\perp} = \frac{1}{5} (2\beta_{z\eta\eta} - 3\beta_{\eta z\eta} + 2\beta_{\eta\eta z}), \quad (103)$$

where the same sequence of optical frequencies (not given explicitly) is used for the laboratory axes and the molecular axes.

The number of independent nonzero tensor elements depends on the nonlinear optical process and on the symmetry of the molecule, see for instance Bogaard and Orr (1975). For example, β^{SHG} is symmetric with respect to the permutation of the second and third indices (see Eq. 100) and this can be used to simplify the equations for the parallel and perpendicular components. For nonzero frequencies, the number of independent tensor components to be computed decreases when Kleinman's symmetry (Kleinman 1972) is assumed – that is, we assume that we can permute the indices of the incoming light without changing the corresponding frequencies,

$$\beta_{\mu\nu\eta}(-\omega_\sigma; \omega_1, \omega_2) \approx \beta_{\mu\eta\nu}(-\omega_\sigma; \omega_1, \omega_2). \quad (104)$$

Although for $\omega_1 \neq \omega_2$ this is only an approximation and these two tensor components are not equal, Kleinman's symmetry is often applied for low frequencies, where it is usually found to be a reasonable approximation.

The basis set requirements for the calculation of first hyperpolarizabilities are much the same as for the linear polarizability. However, as the first hyperpolarizability probes even higher-order electric-field-perturbed densities of the molecule, care should be exercised to ensure that the basis set is sufficiently saturated with respect to diffuse polarizing functions. Special basis sets have been developed for the calculation of hyperpolarizabilities by Pluta and Sadlej (1998), though the same basis sets that can be used for polarizabilities in most cases give reliable estimates also for first hyperpolarizabilities.

Due to the fact that the first hyperpolarizability involves multiple virtual excited states (Eq. 68), the property has proven to be more sensitive to electron correlation effects than the polarizability (Christiansen et al. 2006). Molecules with push-pull

conjugated structures are often characterized by large first hyperpolarizabilities. Due to the problems facing many commonly used local exchange-correlation functionals, DFT has been shown to have difficulties describing correctly nonlinear optical properties such as the first hyperpolarizability for extended systems (Champagne et al. 2000). The recently introduced Coulomb-attenuated B3LYP functional (CAM-B3LYP), see Yanai et al. (2004) has, in contrast, shown quite good performance also for rather extended systems (Paterson et al. 2006).

Second Dipole Hyperpolarizability

The second dipole hyperpolarizability is described by the cubic response function

$$\gamma_{\alpha\beta\gamma\delta}(-\omega_\sigma; \omega_1, \omega_2, \omega_3) = -\langle\langle \hat{\mu}_\alpha; \hat{\mu}_\beta, \hat{\mu}_\gamma, \hat{\mu}_\delta \rangle\rangle_{\omega_1, \omega_2, \omega_3}, \quad (105)$$

where $\omega_\sigma = \omega_1 + \omega_2 + \omega_3$.

Similarly to the lower-order static polarizabilities, the hyperpolarizability $\gamma(0;0,0,0)$ corresponds to an energy derivative and can be computed using the finite-field approach. Among the possible frequency-dependent third-order processes, the most relevant are:

Electric Field – Induced Second Harmonic Generation (dc-SHG, ESHG, EFISH)	
$\gamma_{\mu\nu\eta\xi}^{\text{ESHG}}(-2\omega; \omega, \omega, 0)$	$= -\langle\langle \hat{\mu}_\mu; \hat{\mu}_\nu, \hat{\mu}_\eta, \hat{\mu}_\xi \rangle\rangle_{\omega, \omega, 0}$
Third Harmonic Generation (THG)	
$\gamma_{\mu\nu\eta\xi}^{\text{THG}}(-3\omega; \omega, \omega, \omega)$	$= -\langle\langle \hat{\mu}_\mu; \hat{\mu}_\nu, \hat{\mu}_\eta, \hat{\mu}_\xi \rangle\rangle_{\omega, \omega, \omega}$
Electro-optical or dc-Kerr effect (EOKE, dc-Kerr)	
$\gamma_{\mu\nu\eta\xi}^{\text{dc-Kerr}}(-\omega; \omega, 0, 0)$	$= -\langle\langle \hat{\mu}_\mu; \hat{\mu}_\nu, \hat{\mu}_\eta, \hat{\mu}_\xi \rangle\rangle_{\omega, 0, 0}$
Optical or ac-Kerr effect (OKE, ac-Kerr)	
$\gamma_{\mu\nu\eta\xi}^{\text{ac-Kerr}}(-\omega_1; \omega_1, \omega_2, -\omega_2)$	$= -\langle\langle \hat{\mu}_\mu; \hat{\mu}_\nu, \hat{\mu}_\eta, \hat{\mu}_\xi \rangle\rangle_{\omega_1, \omega_2, -\omega_2}$
Intensity-Dependent Refractive Index or Degenerate Four Wave Mixing (IDRI or DFWM)	
$\gamma_{\mu\nu\eta\xi}^{\text{IDRI}}(-\omega; \omega, \omega, -\omega)$	$= -\langle\langle \hat{\mu}_\mu; \hat{\mu}_\nu, \hat{\mu}_\eta, \hat{\mu}_\xi \rangle\rangle_{\omega, \omega, -\omega}$
dc Optical Rectification (dc-OR, EFior)	
$\gamma_{\mu\nu\eta\xi}^{\text{dc-OR}}(0; \omega, -\omega, 0)$	$= -\langle\langle \hat{\mu}_\mu; \hat{\mu}_\nu, \hat{\mu}_\eta, \hat{\mu}_\xi \rangle\rangle_{\omega, -\omega, 0}$

The experimentally measured quantities in isotropic fluids are usually the scalar components of the tensor γ given by the isotropic average:

$$\gamma_{\parallel} = \frac{1}{15} (\gamma_{\xi\eta\eta\xi} + \gamma_{\xi\eta\xi\eta} + \gamma_{\xi\xi\eta\eta}), \quad (106)$$

$$\gamma_{\perp} = \frac{1}{15} (2\gamma_{\xi\eta\eta\xi} - \gamma_{\xi\xi\eta\eta}), \quad (107)$$

where the same sequence of optical frequencies, not given explicitly above, is used for the laboratory axes and the molecular axes.

As for the first hyperpolarizability, the number of independent nonzero tensor elements depends on the optical process and on the symmetry of the molecule. For example, γ^{THG} is symmetric with respect to the second, third, and fourth indices, and this can be used to simplify Eqs. 106–107. For instance, the average value that should be compared with the experimental THG parallel component becomes

$$\gamma_{\parallel}^{\text{THG}} = \frac{1}{5} \gamma_{\xi\xi\eta\eta}^{\text{THG}}. \quad (108)$$

Experimentally, if all fields have parallel polarization, one can measure the parallel components of the first and second hyperpolarizabilities which take into account the classical orientational averaging. In the case of ESHG, with the optical field polarized perpendicular to the static field, one measures the perpendicular components, and in the case of a dc-Kerr experiment, the differences between the parallel and perpendicular components.

In the ESHG experiment a laser beam passes through the sample in a static electric field and a weak, collinear, frequency-doubled beam is detected. Absolute values for the hyperpolarizabilities cannot be extracted; the signal from the sample is compared to that of a known buffer gas (ultimately helium, for which there are accurate theoretical values (Bishop and Pipin 1989)) or a solid. In analogy with the derivation for $\tilde{\alpha}$, the classical thermal averaging yields

$$\tilde{\gamma}^{\text{ESHG}} = \left(\tilde{\beta}_{ZZZ}^F / F \right)_{F \rightarrow 0} = \langle \gamma_{ZZZZ} \rangle_{\Omega} + \frac{\mu_z \langle \beta_{ZZZ} \rangle_{\Omega}}{3kT}. \quad (109)$$

The molecular hyperpolarizabilities γ^{ESHG} and β^{SHG} can be obtained in this experiment. For non-centrosymmetric molecules a series of measurements over a range of temperatures has to be performed, whereas for centrosymmetric molecules $\beta = 0$ and thus a single measurement at one temperature is sufficient.

The majority of existing experimental data on hyperpolarizabilities are derived from ESHG and dc-Kerr measurements (Shelton and Rice 1994). The dc-Kerr effect differs from the other nonlinear optical processes as it allows for absolute measurements without the need for a reference measurement. The measured molar Kerr constant is (Shelton and Rice 1994):

$$A^{\text{dc-Kerr}} = \frac{N_A}{81\epsilon_0} \left[\frac{3}{2} (\langle \gamma_{ZZZZ} \rangle_{\Omega} - \langle \gamma_{ZXXZ} \rangle_{\Omega}) + \mu_z (\langle \beta_{ZZZ} \rangle_{\Omega} - \langle \beta_{XZX} \rangle_{\Omega}) / kT \right. \\ \left. + \frac{3}{10kT} \left(\alpha_{\alpha\beta} \alpha_{\alpha\beta}^0 - \langle \alpha_{ZZ} \rangle_{\Omega} \langle \alpha_{ZZ} \rangle_{\Omega} \right) + \frac{1}{kT} \mu_z^2 (\alpha_{zz} - \langle \alpha_{ZZ} \rangle_{\Omega}) \right], \quad (110)$$

where N_A is Avogadro's number, and it involves both dynamic hyperpolarizabilities ($\gamma^{\text{dc-Kerr}}$ and β^{EOPE}) as well as static (superscript 0) and dynamic polarizabilities.

The computational requirements for the second hyperpolarizability are more or less the same as for the first hyperpolarizability. Although the number of studies that analyze the importance of electron correlation effects (not including here DFT) is rather limited, the available results confirm in general the findings made for the first hyperpolarizabilities (Christiansen et al. 2006).

Cauchy Moments

Response functions for a system in its electronic ground state are analytic functions of the frequency arguments, except at the poles that occur when a frequency or a sum of frequencies is equal to an excitation energy. Thus, for frequencies below the first pole, the linear, quadratic, or cubic response functions can be expanded in power series (Hättig and Jørgensen 1998). Let us consider in more detail the simplest case when only electric dipole operators are involved and the frequency dependence of the (hyper)polarizabilities is of interest. The Cauchy series – that is, the power series expansion of the frequency-dependent polarizability – is usually written as:

$$\alpha(-\omega; \omega) = \sum_{k=0}^{\infty} \omega^{2k} S(-2k-2) = S(-2) + \omega^2 S(-4) + \omega^4 S(-6) + \dots, \quad (111)$$

where $S(-2) = \alpha(0; 0)$. This expansion is valid also for purely imaginary frequency arguments.

Similar expansions may be applied to describe the dispersion effects for hyperpolarizabilities. Often, the power series expansion in the frequency arguments may be truncated at second order leading to:

$$\alpha_{\text{iso}}(-\omega_{\sigma}; \omega_1) = \alpha_{\text{iso}}(0; 0) [1 + \mathcal{A}^{\alpha} (\omega_{\sigma}^2 + \omega_1^2) + \mathcal{O}(\omega_i^4)], \quad (112)$$

$$\beta_{\parallel}(-\omega_{\sigma}; \omega_1, \omega_2) = \beta_{\parallel}(0; 0, 0) [1 + \mathcal{A}^{\beta} (\omega_{\sigma}^2 + \omega_1^2 + \omega_2^2) + \mathcal{O}(\omega_i^4)], \quad (113)$$

$$\gamma_{\parallel}(-\omega_{\sigma}; \omega_1, \omega_2, \omega_3) = \gamma_{\parallel}(0; 0, 0, 0) [1 + \mathcal{A}^{\gamma} (\omega_{\sigma}^2 + \omega_1^2 + \omega_2^2 + \omega_3^2) + \mathcal{O}(\omega_i^4)]. \quad (114)$$

Each expansion contains only a single second-order dispersion coefficient \mathcal{A} . Since these coefficients for different optical processes of a given order are equivalent (Bishop 1995), they express the dispersion of frequency-dependent properties in a way that is transferable between different optical processes, for instance \mathcal{A}^{β} describes both the second-harmonic generation and the (dc-)electro-optic Pockels effect.

Long-Range Dispersion Interaction Coefficients

For neutral, closed-shell weakly interacting systems, one of the dominant contributions to the interaction energy at large intermolecular distances is the dispersion energy. It can be analyzed using perturbation theory, and for two atoms or molecules A and B , we find (McWeeny 1992):

$$E_{\text{disp}}^{AB} = -\frac{1}{\hbar} \sum_{n \neq 0}^A \sum_{m \neq 0}^B \frac{\left| \langle 0^A 0^B | \hat{V}^{AB} | n^A m^B \rangle \right|^2}{\omega_{n0}^A + \omega_{m0}^B}, \quad (115)$$

where \hat{V}^{AB} denotes the intersystem interaction, $|0^A\rangle$ and $|n^A\rangle$ represent the ground and excited states of system A , ω_{n0}^A is the excitation energy, and similarly for system B .

We would like to express the dispersion energy in terms of response properties of molecules A and B since this allows for a physical interpretation of the dispersion interaction. Such an expression is obtained applying an integral transform to avoid the summation of the A and B excitation energies in the denominator of Eq. 115 and using the multipole expansion of the perturbing operator \hat{V}^{AB} .

For the second-order dispersion interaction between closed-shell atoms A and B we have:

$$E_{\text{disp}}^{AB}(R) = -\frac{C_6^{AB}}{R^6} - \frac{C_8^{AB}}{R^8} - \frac{C_{10}^{AB}}{R^{10}} - \dots, \quad (116)$$

where the coefficient for the dipole–dipole term can be written as (Langhoff et al. 1971):

$$C_6^{AB} = \frac{\hbar}{(4\pi\epsilon_0)^2} \frac{3}{\pi} \int_0^\infty \alpha^A(-i\omega; i\omega) \alpha^B(-i\omega; i\omega) d\omega. \quad (117)$$

This equation is known as the Casimir–Polder formula. The polarizability at imaginary frequencies, $\alpha(-i\omega; i\omega)$, can be obtained for instance using Eq. 111 with an imaginary value of ω . The C_n^{AB} , $n > 6$ coefficients depend on similar higher-order multipole polarizabilities of one or both atoms. The same approach can be applied for atom–molecule and molecule–molecule interactions considering individual tensor components of the required polarizabilities. It can also be used for three-body interactions, and to describe the dispersion contributions to the pair polarizability function (Fowler et al. 1994).

To determine the long-range dispersion coefficients we need the values of $\alpha(-i\omega; i\omega)$ for the whole range of imaginary frequencies, whereas the radius of convergence of the power series expansion of the polarizability (Eq. 111) is determined by the lowest excitation energy. However, using analytic continuation of the Cauchy expansion (Langhoff and Karplus 1970) we can obtain all the required values of $\alpha(-i\omega; i\omega)$ and thus we can determine accurate values of the dispersion interaction

coefficient C_6^{AB} . The polarizabilities at imaginary frequencies can also be obtained directly, solving the complex linear response equations (Norman et al. 2003).

Qualitatively correct estimates of the dispersion energy are obtained in the perturbational approach even using the dispersion coefficients C_n calculated at the DFT level. It has been shown that a successful description of intermolecular forces for large molecules may be obtained in this manner (Misquitta et al. 2005; Podszwa and Szalewicz 2008).

Electric Field Gradient at the Nucleus, Nuclear Quadrupole Coupling Constant

The *electric field gradient* (EFG) tensor at the nucleus $K - V_{\alpha\beta}^K$ – corresponds to the expectation value (Baker et al. 1989; Fowler et al. 1989) of the EFG operator \hat{V}^K given by Eq. 34:

$$V_{\alpha\beta}^K = \langle 0 | \hat{V}_{\alpha\beta}^K | 0 \rangle, \quad (118)$$

where the traceless form of the operator \hat{V}^K is used. The components of the diagonalized EFG tensor, given in its own (EFG) principal axis system, fulfill therefore the equation

$$V_{zz}^K + V_{yy}^K + V_{xx}^K = 0. \quad (119)$$

They are by definition arranged so that

$$|V_{zz}^K| \geq |V_{yy}^K| \geq |V_{xx}^K|. \quad (120)$$

For the asymmetry parameter η_K defined as

$$\eta_K = \frac{V_{xx}^K - V_{yy}^K}{V_{zz}^K}, \quad (121)$$

we thus have $0 \leq \eta_K \leq 1$.

On the other hand, for comparison with experimental data derived from the hyperfine structure of rotational spectra, tensor components defined in the principal axes of inertia system must be used.

For quadrupolar nuclei – that is, nuclei for which the nuclear spin quantum number $I_K \geq 1$, the *nuclear quadrupole coupling* tensor is defined as $-eQ_K \mathbf{V}^K / \hbar$, where eQ_K is the electric quadrupole moment of the nucleus. There is an interaction of the nuclear quadrupole moment tensor with the EFG at the nucleus, mediated by the nuclear spin \mathbf{I}_K and described by the operator:

$$\mathcal{H}^{\text{Nq}} = \frac{1}{2} \sum_K \frac{-eQ_K}{I_K(2I_K - 1)} \mathbf{I}_K^T \mathbf{V}^K \mathbf{I}_K. \quad (122)$$

The *nuclear quadrupole coupling constant* (NQCC) is defined as $-e Q_K V_{zz}^K / \hbar$. The EFG at the nucleus affects the NMR relaxation processes, thus it is relevant in the analysis of the linewidths of NMR signals. The linewidth is proportional to the inverse relaxation time (Abragam 1961), that is:

$$\Delta\nu_{1/2} \propto \frac{1}{T_q} = \frac{3}{40} \frac{2I_K + 3}{I_K^2 (2I_K - 1)} \left(1 + \frac{\eta_K^2}{3}\right) \left(\frac{e Q_K}{\hbar} V_{zz}^K\right)^2 \tau_c, \quad (123)$$

where $\Delta\nu_{1/2}$ is the width of the peak at half of the maximum intensity and τ_c is the correlation time for the quadrupole coupling relaxation.

From the formal $1/r_K^3$ dependence of the operator in Eq. 34, we realize that the basis used to calculate this property must be flexible in the core region and the outer-core/inner-valence region. This is an important requirement, as many modern basis sets are quite heavily contracted in these regions (since the orbitals do not change much in this region when going from the atomic to the molecular system). In general, much of the same basis set requirements as those arising in the calculation of nuclear shielding constants (and to some extent also the spin–spin coupling constants) apply, and we refer to section “[Nuclear Spin–Related Properties](#)” for further discussion.

An important application of theoretically calculated electric field gradients at the nucleus is in the determination of nuclear quadrupole moments. The interaction between the nuclear quadrupole moments and the electric field gradient at the nucleus can be measured with unprecedented accuracy (compared to other experimental approaches for determining nuclear quadrupole moments) in microwave spectroscopy. These interactions give rise to the fine structure in the rotational spectrum of the molecule. Highly accurate estimates for the electric field gradient at the nucleus can often be calculated theoretically (since this is a first-order molecular property). By combining these theoretical results with experimental observations, accurate values of the nuclear quadrupole moments can be obtained, and this approach has been used to revise previously estimated nuclear quadrupole moments (see, e.g., Kellö and Sadlej 1998).

One-Photon Absorption, Excitation Energies, and Transition Moments

Calculations of one-photon absorption enable the determination of the extinction coefficient ε , describing the attenuation of the intensity of an incoming light beam when passing through a sample of absorbing species. In the lowest-order perturbative expansion of the interaction between light and matter, ε is related to the excitation energy and $\langle 0 | \hat{\mu} | n \rangle$, the transition moment between the ground $| 0 \rangle$ and excited $| n \rangle$ electronic states. The derived quantities are the dipole strength

$${}^n D_{\alpha\beta} = \langle 0 | \hat{\mu}_\alpha | n \rangle \langle n | \hat{\mu}_\beta | 0 \rangle, \quad (124)$$

and the oscillator strength. In the length gauge, the latter can be written as

$$f_{0n}^r = \frac{2m_e\omega_{0n}}{3\hbar e^2} \langle 0 | \hat{\mu}_\alpha | n \rangle \langle n | \hat{\mu}_\alpha | 0 \rangle = \frac{2m_e}{3e^2} \lim_{\omega \rightarrow \omega_{0n}} (\omega - \omega_{0n}) \omega \langle \langle \hat{\mu}_\alpha; \hat{\mu}_\alpha \rangle \rangle_\omega. \quad (125)$$

Using the hypervirial relationship

$$\langle 0 | \hat{p}_\alpha | n \rangle = i m_e \omega_{0n} \langle 0 | \hat{r}_\alpha | n \rangle, \quad (126)$$

we obtain in the velocity gauge

$$f_{0n}^v = \frac{2}{3\hbar m_e \omega_{0n}} \langle 0 | \hat{p}_\alpha | n \rangle \langle n | \hat{p}_\alpha | 0 \rangle = \frac{2}{3m_e} \lim_{\omega \rightarrow \omega_{0n}} (\omega - \omega_{0n}) \omega^{-1} \langle \langle \hat{p}_\alpha; \hat{p}_\alpha \rangle \rangle_\omega, \quad (127)$$

and in the mixed length-velocity gauge

$$f_{0n}^{rv} = -\frac{2i}{3\hbar e} \langle 0 | \hat{\mu}_\alpha | n \rangle \langle n | \hat{p}_\alpha | 0 \rangle = -\frac{2i}{3e} \lim_{\omega \rightarrow \omega_{0n}} (\omega - \omega_{0n}) \langle \langle \hat{\mu}_\alpha; \hat{p}_\alpha \rangle \rangle_\omega, \quad (128)$$

respectively (Hansen and Bouman 1980).

To illustrate the differences between oscillator strengths computed in different approaches, we present in Table 2 the results for the $1^1\Sigma_g^+ \rightarrow 1^1\Pi_u$ singlet transition in the N_2 molecule. The HF results and the results for the hierarchy of coupled-cluster approximations (CCS, CC2, CCSD and CC3) show the role of the wave function model. In addition, the role of the basis set is shown in the sequence of the results for the d-aug-cc-pV XZ correlation-consistent basis sets (Kendall et al. 1992), with $X = D, T, Q,$ and 5 , abbreviated in the table as daDZ, daTZ, daQZ, and da5Z, respectively. At the HF level the hypervirial relationship (Eq. 126) is fulfilled in complete basis sets and the differences between computed oscillator strengths are only due to the incompleteness of the basis set. We observe that all the HF results converge with X , and the da5Z values are almost identical. In the CC methods, in addition to the differences arising due to basis set incompleteness, $f_{0n}^r, f_{0n}^v,$ and f_{0n}^{rv} may differ even for a complete basis set since Eq. 126 is not satisfied. This difference should be diminishing as we approach the full CI limit, and in fact for a given one-electron basis the differences between $f_{0n}^r, f_{0n}^v,$ and f_{0n}^{rv} follow the sequence $CCS > CC2 \approx CCSD > CC3$. Finally, we note that for evaluating the accuracy of the results the gauge invariance of a calculated oscillator strength is a necessary, but not sufficient requirement. Gauge invariance has been obtained for large basis sets both at the HF and at the CC3 level, whereas the correlated values differ significantly from the HF results, and only the CC3 results can be expected to be of high accuracy.

Table 2 Oscillator strength and excitation energy (ω_{0n} , in eV) for the singlet transition $1^1\Sigma_g^+ \rightarrow 1^1\Pi_u$ of N_2 ; results taken from Pawłowski et al. (2004)

Model	Basis set	f_{0n}^r	f_{0n}^v	f_{0n}^{rv}	$f_{0n}^r - f_{0n}^v$	ω_{0n}
HF	daDZ	0.1669	0.1725	0.1697	-0.0055	13.26
	daTZ	0.1702	0.1724	0.1713	-0.0022	13.22
	daQZ	0.1709	0.1717	0.1713	-0.0008	13.22
	da5Z	0.1714	0.1715	0.1715	-0.0001	13.22
CCS	daDZ	0.1629	0.1485	0.1556	0.0144	13.27
	daTZ	0.1662	0.1504	0.1582	0.0158	13.24
	daQZ	0.1670	0.1511	0.1589	0.0158	13.23
	da5Z	0.1675	0.1512	0.1592	0.0163	13.23
CC2	daDZ	0.1202	0.1144	0.1172	0.0058	12.42
	daTZ	0.1239	0.1211	0.1225	0.0028	12.58
	daQZ	0.1248	0.1230	0.1239	0.0018	12.65
	da5Z	0.1263	0.1238	0.1251	0.0025	12.69
CCSD	daDZ	0.1671	0.1617	0.1644	0.0055	12.80
	daTZ	0.1757	0.1729	0.1743	0.0029	13.00
	daQZ	0.1794	0.1769	0.1782	0.0025	13.08
	da5Z	0.1822	0.1788	0.1805	0.0034	13.11
CC3	daDZ	0.1844	0.1806	0.1825	0.0037	12.71
	daTZ	0.2046	0.2045	0.2046	0.0002	12.87
	daQZ	0.2141	0.2147	0.2144	-0.0006	12.92
	da5Z	0.2196	0.2194	0.2195	0.0002	12.94

The Thomas–Reiche–Kuhn (TRK) sum rule, which can be written in terms of the dipole oscillator strengths as

$$\sum_n f_{0n} = N_{e1}, \quad (129)$$

where N_{e1} is the total number of electrons, can be used as another test of the accuracy of the calculations.

Of particular concern are Rydberg states, characterized by a very diffuse orbital occupied by the excited electron. DFT in general fails to properly describe Rydberg-excited states, as many of the common exchange-correlation functionals do not display the correct asymptotic behavior for the $-1/r$ operator. This deficiency of modern exchange-correlation functionals can be partially rectified by introducing the correct asymptotic behavior for the exchange-correlation functional (Tozer and Handy 1998). It is worth noting that the time-dependent Hartree–Fock (TDHF) theory does not display the same problems in describing Rydberg states, due to use of the exact exchange operator. However, the lack of electron correlation effects in TDHF may limit the applicability of this approach for excited states in general.

To account for the broadening of the experimental spectrum (related not only to the finite lifetime of the electronic state, but also to the rovibrational structure,

collisions, and other aspects of the interaction between light and matter), and in particular to investigate whether certain bands may be hidden in the experimental spectrum due to overlapping bands, simple Lorentzian line broadening is often added in the form

$$L(\nu) \approx \frac{\tilde{\Gamma}}{\pi} \left[\frac{1}{(\nu_0 - \nu)^2 + \tilde{\Gamma}^2} \right], \quad (130)$$

where $\tilde{\Gamma}$ is related to the lifetime of the excited state, and ν_0 is the frequency of the electronic excitation. Assuming an isolated absorption band, $\tilde{\Gamma}$ is equal to half the width of the absorption band at half height (the value of $\tilde{\Gamma}$ is often adjusted in the calculations to fit the experimental data). It is common also to employ a Gaussian function to reproduce the line broadening

$$L(\nu) = \frac{2}{\Gamma\sqrt{2\pi}} e^{-2(\nu-\nu_0)^2/\Gamma^2}, \quad (131)$$

and we have $\tilde{\Gamma} = \Gamma\sqrt{(\ln 2)/2}$.

Finally, we recall that the transition moment between two excited states $\langle m | A | n \rangle$ can be obtained using only the reference state wave function from a double residue of the quadratic response function Eq. 70 (and if $|m\rangle = |n\rangle$ we may in this way determine the expectation value of A in the excited state).

Two-Photon Absorption

In experimental studies of two- and multi-photon absorption processes, the multi-photon transition strength, a function of all the frequencies of all photons absorbed, is analyzed. For two laser sources of circular frequency ω_1 and ω_2 with associated wavelengths λ_1 and λ_2 , the two-photon transition strength $\delta(\omega_1, \omega_2)$ for the transition between states $|0\rangle$ and $|n\rangle$ in isotropic samples is given by, see e.g., Craig and Thirunamachandran (1984):

$$\begin{aligned} \delta_{0n}^{\text{TPA}}(\omega_1, \omega_2) = & \mathbf{F} \mathcal{S}_{\alpha\alpha}^{0n}(\omega_1, \omega_2) \mathcal{S}_{\beta\beta}^{0n,*}(\omega_1, \omega_2) + \mathbf{G} \mathcal{S}_{\alpha\beta}^{0n}(\omega_1, \omega_2) \mathcal{S}_{\alpha\beta}^{0n,*}(\omega_1, \omega_2) + \\ & + \mathbf{H} \mathcal{S}_{\alpha\beta}^{0n}(\omega_1, \omega_2) \mathcal{S}_{\beta\alpha}^{0n,*}(\omega_1, \omega_2), \end{aligned} \quad (132)$$

where \mathbf{F} , \mathbf{G} , and \mathbf{H} are numbers depending on the polarization state of the two photons and on the geometrical setup (mutual direction of the laser beams) (Craig and Thirunamachandran 1984; McClain 1974) and, in the dipole approximation,

$$\mathcal{S}_{\alpha\beta}^{0n}(\omega_1, \omega_2) = \frac{1}{\hbar} \sum_m \left\{ \frac{(\hat{\mu}_\alpha)_{0m}(\hat{\mu}_\beta)_{mn}}{\omega_{m0} - \omega_1} + \frac{(\hat{\mu}_\beta)_{0m}(\hat{\mu}_\alpha)_{mn}}{\omega_{m0} - \omega_2} \right\} \quad (133)$$

is the second-rank, two-photon tensor. In Eq. 133 the summation runs over the whole set of excited states, the energy conservation relation $\omega_1 + \omega_2 = \omega_{n0}$ applies, and off-resonance conditions are implied – that is, the frequencies ω_1 and ω_2 are sufficiently far off the values at which the denominators vanish. The tensor is nonsymmetric in the exchange of the two frequencies except for $\omega_1 = \omega_2$. For the special case of a one-color beam – that is, a monochromatic light source – the transition matrix is symmetric and (using $\omega = \omega_1 = \omega_2$)

$$\delta_{0n}^{\text{TPA}}(\omega) = S_{\lambda\lambda}^{0n}(\omega) S_{\mu\mu}^{0n,*}(\omega) + 2S_{\lambda\mu}^{0n}(\omega) S_{\lambda\mu}^{0n,*}(\omega) \quad (134)$$

for linear polarization of the incident light and

$$\delta_{0n}^{\text{TPA}}(\omega) = -S_{\lambda\lambda}^{0n}(\omega) S_{\mu\mu}^{0n,*}(\omega) + 3S_{\lambda\mu}^{0n}(\omega) S_{\lambda\mu}^{0n,*}(\omega) \quad (135)$$

for circular polarization, respectively.

It can be shown that the two-photon absorption transition rate (cross section) can be obtained from the single residue of the cubic response function (Hättig et al. 1998a). Two-photon absorption transition amplitudes $S_{\alpha\beta}^{0n}(\omega_1, \omega_2)$, Eq. 133, can also be extracted from the single residue of a quadratic response function, see Eq. 69 above.

The following relations hold for the matrix elements of the two-photon transition tensor

$$S_{\alpha\beta}^{0n}(\omega) = S_{\beta\alpha}^{0n}(\omega_{n0} - \omega), \quad (136)$$

$$S_{\alpha\beta}^{0n}(\omega) = S_{\alpha\beta}^{n0}(-\omega)^*, \quad (137)$$

For methods which do not fulfill Eq. 137 (such as the coupled cluster approach), one can instead use the two-photon transition strength (Hättig et al. 1998b):

$$F_{\alpha\beta,\gamma\delta}^{0n}(\omega) = \frac{1}{2} \left\{ S_{\alpha\beta}^{0n}(-\omega) S_{\gamma\delta}^{n0}(\omega) + S_{\gamma\delta}^{0n,*}(-\omega) S_{\alpha\beta}^{0n,*}(\omega) \right\}, \quad (138)$$

where $S_{\alpha\beta}^{0n}(\omega)$ and $S_{\alpha\beta}^{n0}(\omega)$ are the left and right transition moments.

Three-photon absorption has been described through the single residue of the cubic response function (Cronstrand et al. 2003). Similarly to two-photon absorption, one can discuss the third-order (left and right) transition moments and strengths (Hättig et al. 1998a) and define the three-photon transition strengths in this manner for nonvariational wave functions. There has been a rather limited number of theoretical studies of three-photon absorption processes and these will therefore not be discussed any further here.

In general, the computational requirements for the calculation of molecular properties from the residues of the response functions inherit the requirements from the response functions themselves; that is, the selection of the basis sets has to

be done considering the operators appearing in the expression for the transition moments and excited-state properties. However, as the residues are connected to specific excited states, the nature of the probed excited state also needs to be considered.

Magnetic Properties

Let us now consider perturbations involving external static magnetic fields only. In this case, the number of observable molecular properties that arise is rather limited. This is due in part to the fact that the strength of the magnetic field perturbation of the molecular energy is much smaller than the electric field perturbational strength, but also because only even-order interactions will survive for closed-shell molecules. This is normally referred to as quenching of the induced magnetic moment, and can also be seen as an explanation why there are no permanent magnetic dipoles in such systems. The quenching of the magnetic moment in closed-shell molecules can easily be understood by considering that for any orbitally nondegenerate state, one can always choose the wave function to be real. For such a state, the expectation value of the magnetic moment is given as

$$\langle 0 | \hat{H}_\alpha^B | 0 \rangle = \langle 0 | \hat{H}_\alpha^{B*} | 0 \rangle^* = -\langle 0 | \hat{H}_\alpha^B | 0 \rangle^* = 0, \quad (139)$$

since the magnetic dipole operator is Hermitian and purely imaginary (see Eq. 44).

Magnetizability

The magnetizability yields the first nonvanishing (second order in the external magnetic field) contribution to the energy of interaction between an external magnetic field and a closed-shell molecule

$$\Delta E(\mathbf{B}) = -\frac{1}{2} B_\alpha \xi_{\alpha\beta} B_\beta, \quad (140)$$

where ξ is a symmetric tensor of rank two with a total of six independent elements. For isotropic samples, the magnetizability is expressed in the principal axis system in which the tensor is diagonal, and in addition to the isotropic magnetizability defined in a standard manner as

$$\xi_{\text{iso}} = \frac{1}{3} \xi_{\alpha\alpha}, \quad (141)$$

two anisotropies:

$$\xi_{\text{anis1}} = \xi_{33} - (\xi_{11} + \xi_{22}) / 2, \quad (142)$$

$$\xi_{\text{anis2}} = \xi_{11} - (\xi_{22} + \xi_{33}) / 2, \quad (143)$$

Table 3 Unit conversion factors for magnetizability

	1 a.u.	SI		CGS	
ξ	$(e\hbar/m_e)^2 E_h^{-1}$	7.89104×10^{-29}	JT^{-2}	4.7521	$\text{ppm cgs/ppm cm}^3 \text{ mol}^{-1}$

are usually defined, with $|\xi_{33}| \leq |\xi_{22}| \leq |\xi_{11}|$. However, results reported in the literature often do not follow these definitions, and care should be exercised when comparing with experiment to ensure that the same conventions are used for defining the anisotropies. For unit conversion factors, see Table 3.

Two contributions to the magnetizability appear in the nonrelativistic electronic Hamiltonian in the presence of a magnetic vector potential Eq. 40. One arises as an expectation value of the diamagnetic magnetizability operator, see Eq. 47. The second involves a linear response contribution arising from the interaction of the magnetic dipole operator Eq. 44 with itself. We can, therefore, calculate the magnetizability from the expression:

$$\xi_{\alpha\beta} = -2 \left\langle 0 \left| \hat{H}_{\alpha\beta}^{\text{B,B}} \right| 0 \right\rangle - \left\langle \left\langle \hat{H}_{\alpha}^{\text{B}}, \hat{H}_{\beta}^{\text{B}} \right\rangle \right\rangle = - \left. \frac{\partial^2 E(\mathbf{B})}{\partial B_{\alpha} \partial B_{\beta}} \right|_{\mathbf{B}=0}, \quad (144)$$

where the last equality shows that the magnetizability can be obtained as an energy derivative and it holds because the magnetic fields discussed here are static.

A negative magnetizability corresponds to an induced magnetic moment of the molecule opposing the applied magnetic field, and is experimentally manifested by the molecule trying to get away from the poles of a magnet and thus align itself across the applied magnetic field, an effect that is referred to as *diamagnetism*. The opposite effect, where the induced magnetic moment enforces the magnetic field, is referred to as *paramagnetism*, and in this case the molecule will seek to orient itself in the direction of the poles of the magnet and therefore along the magnetic field.

The diagonal elements of the first contribution to the magnetizability in Eq. 144 are negative since

$$\xi_{\alpha\beta}^{\text{dia}} = -2 \left\langle 0 \left| \hat{H}_{\alpha\beta}^{\text{B,B}} \right| 0 \right\rangle = - \frac{e^2}{4m_e} \left\langle 0 \left| \hat{r}_O^2 \delta_{\alpha\beta} - \hat{r}_{O,\alpha} \hat{r}_{O,\beta} \right| 0 \right\rangle, \quad (145)$$

and this part is for this reason referred to as the diamagnetic contribution. The linear response function contribution to the magnetizability (including the negative sign, see Eq. 144) is always positive and is therefore referred to as the paramagnetic contribution. For closed-shell molecules, the expectation value almost always dominates, and these molecules are therefore diamagnetic. Exceptions occur for molecules for which there are near degeneracies in the electronic ground state. Then the denominator of the linear response function (65) may become small, leading to such an increase of the magnitude of the paramagnetic term that the overall magnetizability becomes positive. A well known and often studied example is the BH molecule (Ruud et al. 1995; Sauer et al. 1993).

Although the expression for the magnetizability in Eq. 144 is conceptually simple and can in principle be used straightforwardly, the gauge origin problem turns out to be difficult. In order to ensure gauge-origin independent magnetizabilities, as well as significantly improve the basis set convergence, London atomic orbitals should be used when calculating magnetizabilities.

A London orbital is obtained from a conventional basis function by multiplying it by an imaginary phase factor

$$\omega_\mu(\mathbf{B}, \mathbf{R}_M) = \exp\left(-\frac{i}{\hbar}e\mathbf{A}_M \cdot \mathbf{r}\right) \chi_\mu(\mathbf{R}_M). \quad (146)$$

χ_μ is here an ordinary basis function, such as a Gaussian or a Slater-type basis function. The magnetic vector potential appearing in the imaginary phase factor is defined as

$$\mathbf{A}_M = \frac{1}{2}\mathbf{B} \times \mathbf{R}_{OM} = \frac{1}{2}\mathbf{B} \times (\mathbf{R}_M - \mathbf{O}). \quad (147)$$

The effect of this complex phase factor is to move the global gauge origin \mathbf{O} to the “best” gauge origin for the basis function located at center \mathbf{R}_M , namely, the nucleus M to which the basis function is attached.

When London atomic orbitals (LAOs, known also as gauge-including atomic orbitals – GIAOs) are used, the definition of dia- and paramagnetic contributions to the magnetizability in Eq. 144 can no longer be used. We can instead define (Ruud and Helgaker 1997):

$$\xi_{\alpha\beta}^{\text{dia}}(\mathbf{O}) = -2\left\langle 0 \left| \hat{H}_{\alpha\beta}^{\text{B,B}}(\mathbf{O}) \right| 0 \right\rangle, \quad (148)$$

$$\xi_{\alpha\beta}^{\text{para,Lon}}(\mathbf{O}) = \xi_{\alpha\beta}^{\text{Lon}} - \xi_{\alpha\beta}^{\text{dia}}(\mathbf{O}). \quad (149)$$

This partitioning leads to dia- and paramagnetic contributions that display the correct gauge origin dependence also for finite basis sets, since the diamagnetic contribution, $\xi_{\alpha\beta}^{\text{dia}}$ in Eq. 148, is an expectation value, and therefore it has the correct origin dependence. Furthermore, expectation values in general show fast basis set convergence. Therefore the significant improvement in convergence of both the total magnetizability and the diamagnetic magnetizability ensures fast basis set convergence also for the paramagnetic contribution, Eq. 149.

To illustrate the importance of using London orbitals when calculating magnetizabilities, we consider in Table 4 the magnetizability (and rotational g tensor, vide infra) of the PF_3 molecule. We note that the isotropic magnetizability calculated using the aug-cc-pVDZ basis set and without London atomic orbitals is wrong by a factor of almost three compared to the estimated Hartree–Fock limit. As shown in Table 4, even when the aug-cc-pV5Z basis containing more than 500 basis functions is used, the results obtained using Eq. 144 are still more than 10 % off

Table 4 Basis set convergence of the SCF magnetizability (in a.u.) and rotational g tensor of PF_3 calculated without and with London atomic orbitals, Ruud and Helgaker (1997)

	aug-cc-pVDZ		aug-cc-pVTZ		aug-cc-pVQZ		aug-cc-pV5Z	
	No-Lon	Lon	No-Lon	Lon	No-Lon	Lon	No-Lon	Lon
ξ^{iso}	-16.55	-6.80	-11.50	-6.67	-9.06	-6.61	-7.66	-6.57
$\Delta\xi$	-3.66	-0.25	-0.77	-0.32	0.48	-0.33	-0.28	-0.34
g_{\parallel}	0.0845	-0.0618	0.0159	-0.0636	-0.0181	-0.0643	-0.0461	-0.0649
g_{\perp}	0.0890	-0.0378	0.0162	-0.0380	-0.0181	-0.0383	-0.0274	-0.0385

the results obtained with London atomic orbitals, which can be expected to be very close to the Hartree–Fock limit considering the excellent basis set convergence of the London atomic orbital magnetizabilities. In contrast, all the calculations using London orbitals are within 3 % of the LAO aug-cc-pV5Z results. More importantly, the anisotropy is off by an order of magnitude at the aug-cc-pVDZ level, and even has the wrong sign at the aug-cc-pVQZ basis (containing 345 basis functions) if London orbitals are not used.

Computational studies have shown that unless the ground electronic state is nearly degenerate with the lowest excited states, electron correlation effects on the isotropic magnetizability are in general negligible. All published results of correlated calculations of the magnetizability of closed-shell molecules without close-lying excited states have found electron correlation effects to be less than 3 % in most cases, much smaller than any uncertainty in experimentally determined magnetizabilities for such molecular systems. Results within 1–2 % of the Hartree–Fock limit can be obtained using the very modest-sized aug-cc-pVDZ basis set of Woon and Dunning (1994). The important point in the construction of a suitable basis set is the inclusion of diffuse polarizing functions in order to ensure highly accurate magnetizabilities.

Commonly adopted exchange-correlation functionals (e.g., B3LYP, BLYP) are in general less efficient in recovering correlation effects on molecular magnetic properties than they are in recovering correlation effects on vibrational spectra or electric properties. Indeed, in many cases DFT may perform worse than Hartree–Fock in the calculation of isotropic magnetizabilities. A set of new functionals designed specifically for the calculation of magnetic properties has been developed by Keal and Tozer (2003). However, recent CCSD benchmark calculations indicate that the results obtained with one of these functionals deteriorate when vibrational corrections and solvent effects are taken into account (Lutnæs et al. 2009).

The isotropic magnetizabilities are in general very little affected by nonelectronic contributions such as zero-point vibrational corrections, dielectric medium effects, or weak intermolecular forces. The situation is very different for the magnetizability anisotropy, which may display both fairly large electron correlation effects as well as sizeable contributions from nonelectronic contributions.

Although highly accurate theoretical calculations using coupled-cluster wave functions have demonstrated that correlation effects are in general small (Gauss

et al. 2007), and that nonelectronic contributions are small as noted above, agreement with experimentally determined magnetizabilities is in general poor. Part of the reason for this is the general lack of experimental gas-phase magnetizabilities. However, more important for the apparent disagreement between theory and experiment may be the fact that isotropic magnetizabilities are almost exclusively determined *relative* to the magnetizability of reference compounds. It has been suggested that the reference values used to determine experimental magnetizabilities are inaccurate (Ruud et al. 2000).

Rotational g Tensor

The magnetizability describes the magnetic moment induced in a molecule by an external magnetic field induction, and how this induced magnetic moment in turn can interact with the external magnetic field and lead to an energy correction. A molecule also has its own source of a magnetic moment, arising from the rotational motion. Since the molecule consists of charged particles, the nuclei, and the electrons, the overall rotation of these particles will generate a small magnetic moment. In the case of the external-field-induced magnetic moment, the vector potential and the electron magnetogyric ratio determine the form of the magnetic dipole operator in Eq. 44. For the magnetic moment induced by the molecular rotation, the magnitude of the rotational angular momentum \mathbf{J} is determined by the rotational quantum numbers of the molecule, giving rise to a rotationally induced magnetic moment

$$m_a^J = \frac{\mu_N}{\hbar} \sum_{a,b} g_{ab}^J J_b, \quad (150)$$

where μ_N is the nuclear magneton and the rotational g tensor (also called the molecular g factor) is defined as

$$g^J = g^{J,\text{nucl}} + g^{J,\text{el}} = m_p \mathbf{I}^{-1} \sum_K Z_K (R_{K,\text{CM}}^2 \mathbf{1} - \mathbf{R}_{K,\text{CM}} \mathbf{R}_{K,\text{CM}}^T) - \frac{4m_p m_e}{e^2} \mathbf{I}^{-1} \xi^{\text{para}}(\mathbf{R}_{\text{CM}}). \quad (151)$$

In Eq. 151 we have introduced the moment of inertia tensor \mathbf{I} and we have used the definition of ξ^{para} as given in Eq. 149, which is applicable both when using London orbitals and when using conventional basis sets. Both the nuclear positions and magnetic dipole operators are defined with respect to the center of mass of the molecule, as this is the point about which the molecule rotates.

There are two contributions to the rotationally induced magnetic moment, one arising from the rotation of the nuclear framework, and one arising because of the rotation of the electron density, and these contributions are opposite in sign because

the linear response function ξ^{para} in Eq. 151 is positive. This rotationally induced magnetic moment is an example of a nonadiabatic effect. If the electrons were able to adjust their positions and momenta instantaneously to any change in the nuclear configuration, the magnetic moment created by the electrons and the nuclei would have the same magnitude but opposite sign, due to the different charges of the particles. However, this description, corresponding to the Born–Oppenheimer approximation, is not exact, and in fact a small permanent magnetic moment will arise from the decoupling of the electronic and nuclear rotation.

The rotational magnetic moment can interact with either internal or external sources of magnetic moments or magnetic fields. If the rotational magnetic moment interacts with the nuclear magnetic moment it gives rise to a contribution to the energy, discussed in more detail in section “[Nuclear Spin Rotation Constants](#),” which involves the nuclear spin–rotation constant. If we apply an external magnetic field and consider the interaction of this field with the magnetic moment created by the molecular rotation, the interaction energy is (Flygare 1974; Gauss et al. 1996):

$$\Delta E = -\frac{\mu_N}{\hbar} \sum_{a,b} B_a g_{ab}^J J_b, \quad (152)$$

where J_a is the rotational quantum number along the principal rotational axis a .

The rotational g tensor is a tensor of rank two with in principle nine independent elements. The rotational magnetic moment is only well defined in the principal rotational axis system, where the moment of inertia tensor is diagonal, and only the diagonal elements of g^J in this principal axes system contribute to the rotational Zeeman effect (Flygare 1974; Gauss et al. 1996).

We also note that if we can determine both the magnetizability and the rotational g factor, we can indirectly determine the molecular quadrupole moment from the equation:

$$q_{cc} = \frac{2m_e}{e} (\xi_{aa} + \xi_{bb} - 2\xi_{cc}) + \frac{e}{2m_p} (g_{aa}^J I_{aa} + g_{bb}^J I_{bb} - 2g_{cc}^J I_{cc}). \quad (153)$$

This relation shows that microwave Zeeman spectroscopy can be an important source of molecular quadrupole moments. From a theoretical point of view, this equation does not provide any additional information compared to calculating the molecular quadrupole moment directly as an expectation value, since the relationship (Eq. 153) is fulfilled exactly for any wave function.

The electronic contribution to the rotational g factor is important in high-resolution microwave spectroscopy even in the absence of an external magnetic field. In Eq. 150 we introduced the magnetic moment induced by the rotation of the molecular framework. Instead of letting this induced magnetic moment interact with an external magnetic field induction, we can let it interact with the rotational moment of the molecule. This gives a correction to the molecular energy which is quadratic in the rotational quantum number J . In ordinary microwave spectroscopy, the leading energy contribution is related to the moment of inertia tensor and is

also quadratic in J . Therefore, the energy correction related to g^J can be treated as a correction to the moment of inertia tensor, and it is customary to introduce an effective inverse moment of inertia tensor defined as (Flygare 1974):

$$\mathbf{I}_{\text{eff}}^{-1} = \mathbf{I}^{-1} \left(\mathbf{1} - \frac{m_e}{m_p} g^{J,\text{el}} \right). \quad (154)$$

When comparing highly accurate theoretical estimates of moments of inertia directly with the results observed in high-resolution microwave spectroscopy, it is essential to take into account these nonadiabatic corrections. It is particularly noteworthy that the experimental determination of the equilibrium geometry involves the estimation of the effects of the rotational g tensor on the moment of inertia tensor, and that it is currently possible to calculate these effects with an accuracy much higher than obtainable from experimental data alone. Therefore, the use of highly accurate experimental data, for instance from rotational spectra, combined with highly accurate calculations of small corrections to these quantities, is currently the most accurate way to obtain experimental equilibrium geometries (Stanton et al. 2001) (for a related analysis of diatomic systems, see Sauer et al. 2005).

As the rotational g tensor includes a linear response function involving two magnetic dipole operators, similar computational requirements apply for this property as for the magnetizability. We can therefore expect that the direct calculation of the rotational g tensor using Eq. 151 will display very slow basis set convergence. However, London orbitals are very efficient in getting fast basis set convergence, as illustrated for PF_3 in Table 4.

In microwave or molecular beam experiments, all of the three diagonal elements of the rotational g tensor are determined. As for the magnetizability, the individual tensor components show some dependence on electron correlation effects, and in order to allow for the calculated results to be within the very tight experimental errors bars for the rotational g tensor elements, electron correlation and the effects of zero-point vibrational corrections need to be included. This is illustrated for the water molecule in Table 5.

Table 5 Electron correlation effects and zero-point vibrational corrections to the rotational g tensor components of H_2O . The molecule is located in the xz -plane with the dipole moment along the z -axis (g^J is dimensionless)

Computational method	g_{xx}^J	g_{yy}^J	g_{zz}^J
Hartree–Fock (Ruud et al. 1998)	0.6635	0.6814	0.7342
RASSCF (Ruud et al. 1998)	0.651(1)	0.667(3)	0.728(3)
CCSD (Gauss et al. 2007)	0.6545	0.6707	0.7304
CCSD(T) (Gauss et al. 2007)	0.6543	0.6690	0.7308
RASSCF + ZPV (Ruud et al. 1998)	0.637(1)	0.640(3)	0.709(3)
Exp. (Molecular beam) (Verhoeven and Dymanus 1970)	0.645(6)	0.657(1)	0.718(7)
Exp. (Microwave Zeeman) (Kukolich 1969)	0.6465(20)	0.6650(20)	0.7145(20)

Birefringences and Dichroisms

The consequences of the fact that an electromagnetic wave has both a frequency-dependent electric field component and a frequency-dependent magnetic field component have so far not been discussed. The two components are perpendicular to each other, as well as perpendicular to the direction of the propagation of the light beam. Moreover, we have not discussed the implications of the possible polarization of the light beam. Light experienced in everyday life is mostly nonpolarized – that is, the tip of the electric or magnetic field vectors moves randomly in the plane perpendicular to the propagation direction. With the use of appropriate tools, unpolarized light can be turned into a wide variety of polarization states, becoming linearly, elliptically, or circularly polarized. In these cases, the tip of the field vector(s) oscillates in the plane containing the direction of propagation (linear polarization) or rotates along a circle (circular polarization) or an ellipse (elliptical polarization) in the plane perpendicular to it. The polarization of light plays an important role in the interactions of light with a molecule.

Birefringences and dichroisms are two aspects of this interaction which are closely related to the polarization status of light. They arise through a variety of different mechanisms and interactions between the electromagnetic field and the molecule, but they are all characterized by the simultaneous measurements of the refractive (birefringences) or absorptive (dichroisms) index of light along two directions, different with respect to some predefined laboratory or molecular frame. The difference in the dispersion or absorption in these two directions is described by a quantity of a specific sign and magnitude. All these properties give detailed information about the molecular structure, and in particular some of them can distinguish between a sample and its mirror image. Distinguishable mirror images can occur naturally, in chiral molecules where the mirror image of the molecule cannot be superimposed on the original molecule itself – that is, for so-called enantiomers – or they can occur by introducing an asymmetry in the experimental setup that makes the mirror image of the experimental design impossible to superimpose on the original setup. In the latter case, since the experimental design is itself chiral, the observed effect does not vanish even if the molecule is not chiral. We note that only molecules possessing no improper or rotation-reflection axes may be chiral.

It is instructive to start by giving some examples of birefringences, before discussing the details of the most important birefringences and dichroisms. Let us consider first an isotropic liquid sample subject to a strong external static electric field. The refractive index n_{\parallel} experienced by the component of the polarization vector aligned parallel to the direction of the external field \mathbf{F} in this case differs from the index of refraction n_{\perp} for the component aligned perpendicular to the direction of the vector \mathbf{F} (Kerr 1875a, b). As a consequence, the initially linearly polarized light beam will exit the region of the sample with an ellipticity. This is an example of a *linear birefringence*

$$\Delta n^{\text{lin}} = n_{\parallel} - n_{\perp}, \quad (155)$$

and the effect is known as electric-field-induced optical birefringence or more often as the Kerr (electro-optical) effect. It exhibits a quadratic dependence on the strength of the applied electric field. Its existence is due mainly to the fact that the external electric field, in a system possessing an anisotropic electric dipole polarizability tensor, tends to align the molecules preferentially in its direction.

Other mechanisms in general also contribute to the emergence of an anisotropy, the most important being the effect of electronic rearrangements. They play a role through the higher-order polarizabilities of the system, in particular through the second electric-dipole hyperpolarizability γ and, in the presence of permanent electric dipoles, through the first electric-dipole hyperpolarizability β (Kerr 1875a, b). Other examples of linear birefringences are the Cotton–Mouton and Buckingham effects, briefly discussed below, and the Jones and magneto-electric birefringences, which occur when linearly polarized light goes through an isotropic fluid perpendicularly to both static electric and magnetic fields.

Circular birefringence is observed when the two circular components of a linearly polarized beam propagate with different circular velocities, and therefore an anisotropy arises between the two refractive indices n_+ and n_- :

$$\Delta n^{\text{circ}} = n_+ - n_- \quad (156)$$

The net result is a rotation of the plane of polarization. An example of a circular birefringence is *optical activity*, which is observed when a medium composed of chiral molecules (with an excess of one enantiomer) is subject to linearly polarized electromagnetic radiation. The dependence of the corresponding anisotropy on the wavelength is described by the optical rotatory dispersion (ORD). The Faraday effect, discussed below, is another well-known example of a circular birefringence.

An *axial birefringence* occurs with unpolarized light, and an example is the so-called magnetochiral birefringence (Barron and Vrbancich 1984; Kalugin et al. 1999). It can be seen in isotropic samples composed of chiral molecules, and it is observed when a static magnetic induction field is switched on parallel to the direction of propagation of the unpolarized probe beam. The refractive index experienced by the beam propagating parallel to the external magnetic field, $n_{\uparrow\uparrow}$, becomes different from that experienced by a beam propagating antiparallel to the field, $n_{\uparrow\downarrow}$

$$\Delta n^{\text{ax}} = n_{\uparrow\uparrow} - n_{\uparrow\downarrow} \quad (157)$$

The corresponding anisotropy has a different sign for the different enantiomers.

An external field will have an orientational effect on the molecules of the sample, and this is a mechanism responsible for the emergence of birefringences. In the presence of external fields the sample becomes anisotropic, and the extent to which this happens depends on the temperature. Indeed, the general form of the optical anisotropy (for measurements taken at a fixed pressure) is usually of the form

$$\Delta n \propto A_0 + \frac{A_1}{T} + \frac{A_2}{T^2} + \dots, \quad (158)$$

where the term A_0 includes all contributions arising from the reorganization of the electrons due to the action of the external field(s), whereas the terms A_1, A_2, \dots are connected to different mechanisms of reorientation of the molecules, involving the interaction of the field(s) with permanent electric or magnetic multipole moments. The temperature-dependent terms vanish for systems of spherical symmetry; the contributions exhibiting a nonlinear dependence are usually connected to the presence of permanent electric or magnetic dipole moments, or higher-order processes involving more complicated interactions between fields and multipoles.

When the inducing field is time dependent, the birefringence is said to be “optically induced.” In general, the information provided by birefringences induced by static fields can differ from that obtained observing the corresponding optically induced phenomena. For example, in the static Kerr effect of dipolar molecules, an important role is played by the permanent dipoles, which tend to be aligned by the static field. This contribution vanishes for dynamic inducing fields, where the birefringence arises only from the rearrangement of the anisotropy of the electric dipole polarizability. The contribution resulting from the interaction with permanent multipoles averages over time to zero.

Optical Rotation

The first contribution to the induced electric dipole moment in Eq. 74 from the time dependence of the magnetic field, $\omega^{-1}G'$, gives rise to two different observable properties. In the dispersive region, this property determines the optical rotatory power, or just optical rotation (OR) for short, and in the absorptive region it determines the rotational strength observed in electronic circular dichroism (ECD).

There are no terms bilinear in the electric and magnetic fields in the Hamiltonian, therefore G' , the mixed electric dipole–magnetic dipole polarizability (see also Eq. 74), can be expressed as a linear response function:

$$G'_{\alpha\beta}(-\omega; \omega) = -\frac{2\omega}{\hbar} \sum_{n \neq 0} \mathcal{I}m \frac{\langle 0 | \hat{\mu}_\alpha | n \rangle \langle n | \hat{m}_\beta | 0 \rangle}{\omega_n^2 - \omega^2} = -\mathcal{I}m \langle \langle \hat{\mu}_\alpha; \hat{m}_\beta \rangle \rangle_\omega. \quad (159)$$

We note that G' vanishes if the electromagnetic field is static ($\omega = 0$), see the discussion in section “[Expansions of Energy and Multipole Moments](#).” For unit conversion factors, see Table 6.

Measurements of optical rotation are almost exclusively carried out on liquid samples (though gas-phase measurements recently became possible (Müller

Table 6 Unit conversion factors for electric dipole–magnetic dipole polarizability

	1 a.u.	SI	CGS
G'	$e^2 a_0^3 \hbar^{-1}$	$3.60702 \times 10^{-35} \text{ C}^2 \text{ m}^3 \text{ J}^{-1} \text{ s}^{-1}$	$1.08136 \times 10^{-27} \text{ Fr cm G}^{-1}$

et al. 2000)) and we will, therefore, be concerned primarily with the rotational average of the G' tensor, which is described by the quantity ${}^{\text{ORD}}\beta$ defined as

$${}^{\text{ORD}}\beta = -\frac{1}{3\omega} G'_{\alpha\alpha}. \quad (160)$$

The most common experimental setup involves the determination of the rotation of plane-polarized light as it passes through a sample in which there is an excess of one enantiomer. The standard optical rotation $[\alpha]_D^{25}$, proportional to ${}^{\text{ORD}}\beta$, is reported for light with a frequency corresponding to the sodium D-line (589.3 nm) at a temperature of 25 °C.

Because the mixed electric dipole–magnetic dipole polarizability involves the magnetic dipole operator, in approximate calculations G' carries an origin dependence. Indeed, the individual tensor elements of G' are origin dependent. The trace of G' must be origin independent, since the optical rotation is an experimental observable. In non-isotropic media, contributions to the optical rotation tensor arise from the mixed electric dipole–electric quadrupole polarizability A ,

$$A_{\alpha,\beta\gamma}(-\omega; \omega) = -\left\langle\left\langle \hat{\mu}_{\alpha}; \hat{\Theta}_{\beta\gamma} \right\rright\rangle_{\omega}, \quad (161)$$

(see the last term in Eq. 74), and it is the combination of the G' and A contributions that is gauge-origin independent for exact wave functions. The contribution from A vanishes in isotropic samples, since this mixed electric dipole–electric quadrupole polarizability is traceless.

For approximate variational wave functions, the origin independence of the trace of G' is only achieved in the limit of a complete basis set. One way to overcome this problem is to introduce local gauge origins by using London atomic orbitals (Helgaker et al. 1994).

The dispersion of the optical rotation was for a long time also the focus of much experimental attention through optical rotatory dispersion measurements. Even after it became customary to restrict the optical rotation measurements to a single frequency, ORD served as an important tool for determining excitation energies in chiral molecules, although it has now been surpassed by electronic circular dichroism for these purposes (see section “Circular Dichroism”).

At the sodium frequency, the optical rotation is rather small for most molecules. However, the individual diagonal elements of the mixed electric dipole–magnetic dipole polarizability may be fairly large in absolute value, often canceling each other in the trace. This is illustrated for a few selected molecules in Table 7. Consequently, the optical rotation is highly sensitive to numerical errors in the tensor components, because small residual errors in the individual tensor components, arising from the solution of the linear response equations, may lead to substantial errors in ${}^{\text{ORD}}\beta$. Another consequence of this cancellation is that ${}^{\text{ORD}}\beta$ is very sensitive to the choice of molecular geometry, as well as zero-point vibrational effects (Ruud et al. 2001), since the small changes introduced by these effects in the molecular charge

Table 7 The diagonal elements of the mixed electric dipole–magnetic dipole polarizability and the trace of the tensor for a few selected molecules (in a.u., $[\alpha]_D^{25}$ in $\text{deg cm}^3 \text{ dm}^{-1} \text{ g}^{-1}$)

Molecule	$\omega^{-1}G'_{xx}$	$\omega^{-1}G'_{yy}$	$\omega^{-1}G'_{zz}$	$^{\text{ORD}}\beta$	$[\alpha]_D^{25}$
H ₂ O ₂ ($\phi = 120^\circ$)	−1.201	0.387	0.960	−0.049	−55.3
(+)-methyloxirane	−0.820	−0.812	1.587	0.015	10.2
(+)-camphor	0.240	4.891	−5.847	0.239	60.5

DFT/B3LYP results obtained with the Sadlej polarized basis sets, taken from Giorgio et al. (2004)

Table 8 The specific optical rotation ($[\alpha]_D^{25}$ in $\text{deg cm}^3 \text{ dm}^{-1} \text{ g}^{-1}$) for different conformers of (S)(−)-paraconic acid calculated using different basis sets and at DFT/B3LYP and MP2 optimized geometries^a

Conformer	DFT geometries			MP2 geometry
	6-311++G**	aug-cc-pVDZ	aug-cc-pVDZ	aug-cc-pVDZ
A	−31.27	−42.35 ^b	−47.91	−10.58
B	152.79	144.53 ^b	145.76	^c
C	112.01	95.73 ^b	91.47	109.74
D	−250.61	−245.03 ^b	−253.88	−251.71
E	−19.21	−10.78 ^b	0.36	−13.49
F	−188.13	−182.86 ^b	−182.67	−217.33
$[\alpha]_D^{25}$	+14.78	+6.72	+4.40	−85.23

^aThe bottom line of the table gives the total specific optical rotation $[\alpha]_D^{25}$ (all conformations weighted by the Boltzmann distribution). All the results taken from Marchesan et al. (2005)

^bThe optical rotations were calculated for the B3LYP/6-311++G** geometries

^cAn energy minimum corresponding to a **B**-like conformer could not be found

distribution, and thus on the different diagonal elements of G' , can give rather large effects on $^{\text{ORD}}\beta$.

The sensitivity of the optical rotation is perhaps most clearly illustrated in the case of conformationally flexible molecules – that is, molecules that have a significant population of multiple stable minima. As exemplified for paraconic acid (a substituted γ -butyrolactone) in Table 8, different molecular conformations can have optical rotations that differ by orders of magnitude and even in sign. A thorough conformational search is therefore mandatory before the absolute sign of the optical rotation (or any birefringence or dichroism for that matter) is determined from theoretical calculations for conformationally flexible molecules. The optical rotation of the flexible molecule can then be determined by Boltzmann averaging over the dominant molecular conformations.

In order to ensure reasonably well-converged results for the optical rotation, basis sets of polarized valence double-zeta quality are required, and sets such as the aug-cc-pVDZ of Woon and Dunning (1994), or the polarized triple-zeta basis of Sadlej (1988) have been shown to perform well for calculations of optical rotation. Most importantly, diffuse p functions in the outer regions of the electron density, which in most cases is described by the electron density of hydrogen atoms, are required to ensure qualitatively correct results (Zuber and Hug 2004).

Since the optical rotation is very sensitive even to small changes in the electron density, electron correlation effects should also be taken into account in order to get accurate results. Due to the fact that chiral molecules in general have very low symmetry, if any symmetry at all, the only viable approach for calculating electron correlation effects in chiral molecules is currently density functional theory (Ruud et al. 2003; Stephens et al. 2001). Methods such as MP2 and CCSD could also be used to describe electron correlation, but due to their nonvariational nature it is not ensured that gauge-origin independent results can be obtained in the conventional length-gauge formulation even in the limit of a complete basis set. In principle, the use of the dipole velocity gauge will ensure that the calculated results are independent of the gauge origin. For all but the smallest basis sets, the differences between the velocity and the length gauges are not very large, with the length gauge in general performing better. However, in the case of the optical rotation an additional complication arises: whereas G' given by Eq. 159 will vanish in the limit of a static field, this is not the case for the dipole-velocity analogue of the optical rotation. It has been suggested that the much slower basis set convergence of the optical rotation compared to other properties calculated using the velocity-gauge formulation can be improved by subtracting the static-limit value of the corresponding response function (Pedersen et al. 2004), such that

$$G' \propto \left\langle \left\langle \hat{p}_\alpha; \hat{l}_\beta \right\rangle \right\rangle_\omega - \left\langle \left\langle \hat{p}_\alpha; \hat{l}_\beta \right\rangle \right\rangle_0. \quad (162)$$

Since there are no empirical rules relating the stereochemistry of a molecule to the observed sign of the optical rotation, an important area of application for optical rotation calculations would be the combined theoretical and experimental determination of absolute configurations. However, this is a difficult task due in part to the large variations with geometry in the optical rotation of conformationally flexible molecules, and in part because of the small magnitude of $^{\text{ORD}}\beta$. These factors make the sign of the optical rotation hard to determine with confidence. One way to increase the predictive power of the calculations is to change the frequency of the incident light to shorter wavelengths. The magnitude of the optical rotation increases dramatically as the frequency approaches that of a resonance. Thus, if both theory and experiment could determine the optical rotation at frequencies closer to electronic excitation energies, theoretical calculations could provide a much more reliable proof of the absolute configuration of the molecule (Giorgio et al. 2004).

Circular Dichroism

Natural (Electronic) Circular Dichroism

Let us now turn to the lowest-order absorption process involving mixed electric and magnetic perturbations. This property, which is the absorptive analogue of the optical rotation, is known as *electronic circular dichroism* (ECD) or just CD for short.

The differential absorption of circularly polarized light, corresponding to the difference between the absorptive index of the two circular components of linearly polarized light, is proportional to the rotational strength, which is normally calculated as the residue of the linear response mixed electric dipole–magnetic dipole polarizability:

$$\lim_{\omega \rightarrow \omega_{n0}} \hbar (\omega - \omega_{n0}) \langle\langle \hat{\mu}_\alpha; \hat{m}_\beta \rangle\rangle_\omega = \langle 0 | \hat{\mu}_\alpha | n \rangle \langle n | \hat{m}_\beta | 0 \rangle. \quad (163)$$

Since this expression corresponds to the infinite lifetime approximation for the excited state, only a single number will be obtained at the frequency of the electronic excitation.

In general ${}^n R$, the rotatory strength for the transition $|0\rangle \rightarrow |n\rangle$, includes an electric dipole–magnetic dipole contribution

$${}^n R_{\alpha\beta}^m = -\frac{3ie^2}{4m_e} \left(\delta_{\alpha\beta} \langle 0 | \mathbf{r} | n \rangle \langle n | \mathbf{I}^T | 0 \rangle - \langle 0 | \hat{r}_\beta | n \rangle \langle n | \hat{l}_\alpha | 0 \rangle \right), \quad (164)$$

and an electric dipole–electric quadrupole contribution

$${}^n R_{\alpha\beta}^Q = -\frac{3\omega_{n0}e^2}{4} \epsilon_{\alpha\gamma\delta} \langle 0 | \hat{r}_\gamma | n \rangle \langle n | \hat{q}_{\delta\beta} | 0 \rangle. \quad (165)$$

For randomly oriented molecules, the averaging leaves only the electric dipole–magnetic dipole contribution and the scalar rotatory strength is given by

$${}^n R = -\frac{ie^2}{2m_e} \langle 0 | \mathbf{r}^T | n \rangle \langle n | \mathbf{I} | 0 \rangle, \quad (166)$$

corresponding to Eq. 163. These expressions are given in the length gauge. In the velocity gauge

$${}^n R_{\alpha\beta}^m = \frac{3e^2}{4m_e^2\omega_{n0}} \left(\delta_{\alpha\beta} \langle 0 | \mathbf{p} | n \rangle \langle n | \mathbf{I}^T | 0 \rangle - \langle 0 | \hat{p}_\beta | n \rangle \langle n | \hat{l}_\alpha | 0 \rangle \right), \quad (167)$$

and

$${}^n R_{\alpha\beta}^Q = \frac{3e^2}{4m_e^2\omega_{n0}} \epsilon_{\alpha\gamma\delta} \langle 0 | \hat{p}_\gamma | n \rangle \langle n | \hat{T}_{\delta\beta}^+ | 0 \rangle, \quad (168)$$

where \mathbf{T}^+ indicates the “velocity” form of the electric quadrupole operator

$$\mathbf{T}^+ = -(\mathbf{rp} + \mathbf{pr}). \quad (169)$$

This form has an advantage in comparison to the length form. Although with a translation of the reference frame the magnetic dipole and electric quadrupole components change, the total tensor in the velocity gauge is invariant to such a

change of origin. For the length gauge, this invariance depends in addition on the fulfillment of the hypervirial relation, Eq. 126.

Two-Photon Circular Dichroism

Two-photon circular dichroism (Tinoco 1975) arises in chiral systems due to the differential absorption of two photons, of which at least one is circularly polarized (De Boni et al. 2008). In this sense it can be seen as the nonlinear extension of ECD. The observable, the anisotropy of the two-photon transition strength, is proportional to the two-photon rotatory strength (Jansík et al. 2005; Rizzo et al. 2006; Tinoco 1975):

$${}^n R^{\text{TPCD}}(\omega) = -b_1[\mathcal{B}_1(\omega)]_{0n} - b_2[\mathcal{B}_2(\omega)]_{0n} - b_3[\mathcal{B}_3(\omega)]_{0n}, \quad (170)$$

where b_1 , b_2 , and b_3 are numbers, simple combinations of the analogous polarization and setup-related coefficients \mathbf{F} , \mathbf{G} , and \mathbf{H} given for two-photon absorption in Eq. 132. The molecule-related parameters \mathcal{B}_1 , \mathcal{B}_2 , and \mathcal{B}_3 take the form:

$$[\mathcal{B}_1(\omega)]_{0n} = \frac{1}{\omega^3} \mathcal{M}_{\rho\sigma}^{\text{p},0n}(\omega) \mathcal{P}_{\rho\sigma}^{\text{p}*,0n}(\omega), \quad (171)$$

$$[\mathcal{B}_2(\omega)]_{0n} = \frac{1}{2\omega^3} \mathcal{T}_{\rho\sigma}^{+,0n}(\omega) \mathcal{P}_{\rho\sigma}^{\text{p}*,0n}(\omega), \quad (172)$$

$$[\mathcal{B}_3(\omega)]_{0n} = \frac{1}{\omega^3} \mathcal{M}_{\rho\rho}^{\text{p},0n}(\omega) \mathcal{P}_{\sigma\sigma}^{\text{p}*,0n}(\omega), \quad (173)$$

and they are, therefore, appropriate contractions of generalized two-photon second-rank tensors, like the one given in Eq. 133. Indeed, these tensors are defined (for the general case of two photons of different frequency) as follows:

$$\mathcal{P}_{\alpha\beta}^{\text{p},0n}(\omega_1, \omega_2) = \frac{1}{\hbar} \sum_m \left\{ \frac{(\hat{\mu}_\alpha^{\text{p}})_{0m} (\hat{\mu}_\beta^{\text{p}})}{\omega_{m0} - \omega_1} + \frac{(\hat{\mu}_\beta^{\text{p}})_{0m} (\hat{\mu}_\alpha^{\text{p}})_{mn}}{\omega_{m0} - \omega_2} \right\}, \quad (174)$$

$$\mathcal{M}_{\alpha\beta}^{\text{p},0n}(\omega_1, \omega_2) = \frac{1}{\hbar} \sum_m \left\{ \frac{(\hat{\mu}_\alpha^{\text{p}})_{0m} (\hat{m}_\beta)_{mn}}{\omega_{m0} - \omega_1} + \frac{(\hat{m}_\beta)_{0m} (\hat{\mu}_\alpha^{\text{p}})_{mn}}{\omega_{m0} - \omega_2} \right\}, \quad (175)$$

$$\mathcal{T}_{\alpha\beta}^{+,0n}(\omega_1, \omega_2) = \frac{1}{\hbar} \varepsilon_{\beta\rho\sigma} \sum_m \left\{ \frac{(\hat{T}_{\alpha\rho}^+)_{0m} (\hat{\mu}_\sigma^{\text{p}})_{mn}}{\omega_{m0} - \omega_1} + \frac{(\hat{\mu}_\sigma^{\text{p}})_{0m} (\hat{T}_{\alpha\rho}^+)_{mn}}{\omega_{m0} - \omega_2} \right\}, \quad (176)$$

where we have introduced the velocity form of the dipole operator,

$$\hat{\mu}_\alpha^{\text{p}} = -\frac{e}{m_e} \sum_i \hat{p}_{i\alpha}. \quad (177)$$

Within the formalism of response theory, the second-rank tensors of interest are obtained from the single residues of appropriate quadratic response functions, see Eq. 69. The quadratic response functions of relevance for two-photon circular dichroism are:

$$\left\langle\left\langle\hat{\mu}_{\alpha}^{\text{P}};\hat{\mu}_{\beta}^{\text{P}},V^{\omega_n}\right\rright\rangle_{\omega_1,\omega_2}\Rightarrow\mathcal{P}_{\alpha\beta}^{\text{P},0n}(\omega_1,\omega_2),\quad(178)$$

$$\left\langle\left\langle\hat{\mu}_{\alpha}^{\text{P}};\hat{m}_{\beta},V^{\omega_n}\right\rright\rangle_{\omega_1,\omega_2}\Rightarrow\mathcal{M}_{\alpha\beta}^{\text{P},0n}(\omega_1,\omega_2),\quad(179)$$

$$\varepsilon_{\beta\rho\sigma}\left\langle\left\langle\hat{T}_{\alpha\rho}^{+};\hat{\mu}_{\sigma}^{\text{P}},V^{\omega_n}\right\rright\rangle_{\omega_1,\omega_2}\Rightarrow\mathcal{T}_{\alpha\beta}^{+,0n}(\omega_1,\omega_2),\quad(180)$$

where V^{ω_n} is an arbitrary operator (corresponding to the excitation vector to the state n). Single residues of quadratic response functions are efficiently and accurately computed nowadays with a number of wave function models. Nevertheless, DFT has been used almost exclusively in the few theoretical studies of two-photon circular dichroism that have been published (Jansík et al. 2007, 2008).

Faraday Effect

In 1846, Faraday observed that when a static magnetic field was applied with its direction parallel to the direction of propagation of linearly polarized light traversing a transparent sample (glass, liquid, gas), the plane of polarization of the beam was rotated (Faraday 1846a, b). It was the first ever experimental evidence of a birefringence phenomenon, and in particular of a circular birefringence, and the effect is often referred to as the *Faraday effect* or magnetic field-induced optical rotation. The rotation is proportional to the optical anisotropy, see Eq. 156, and this depends linearly on the strength of the applied magnetic induction field. The mechanisms governing this effect involve the interaction of the static external magnetic induction field with the oscillating electric field of the light wave and with the electric multipoles of the sample. The term “magnetic optical rotatory dispersion” (MORD) is often used to denote the frequency dispersion of the Faraday effect, and thus of the corresponding circular birefringence.

For a fluid sample, the rotation with respect to the molecular frame becomes (Barron 2004; Buckingham and Stephens 1966):

$$\phi = V(\omega) B_z l \propto \omega \varepsilon_{\alpha\beta\gamma} \left(\alpha'_{\alpha\beta,\gamma}^{\text{B}} + \frac{1}{kT} \langle 0 | \hat{m}_{\gamma} | 0 \rangle \alpha'_{\alpha\beta} \right) B_z l, \quad(181)$$

where l is the path length, we have introduced the Verdet constant $V(\omega)$, and α'^{B} denotes the magnetic field derivative of α' , the antisymmetric electric dipole polarizability. The second term in the parentheses vanishes for closed-shell systems, as the magnetic dipole is quenched, $\langle 0 | \hat{m}_{\gamma} | 0 \rangle = 0$.

Table 9 Unit conversion factors for the Verdet constant

	1 a.u.	SI	CGS
$V(\omega)$	$\text{rad } e a_0 \hbar^{-1}$	$8.03961 \times 10^4 \text{ rad T}^{-1} \text{ m}^{-1}$	$2.76382 \times 10^2 \text{ min G}^{-1} \text{ cm}^{-1}$

The Verdet constant of a closed-shell system, free of orbital degeneracies both in the ground and in the excited state, is therefore obtained from the response function (Jaszuński et al. 1994; Parkinson and Oddershede 1997):

$$V(\omega) \propto i\omega \varepsilon_{\alpha\beta\gamma} \langle \langle \hat{\mu}_\alpha; \hat{\mu}_\beta, \hat{m}_\gamma \rangle \rangle_{\omega,0}. \quad (182)$$

This contribution is sometimes referred to as the *B* term contribution to the Verdet constant. For molecules containing orbital degeneracies additional contributions arise (the *A* and *C* terms), in analogy with the closely related phenomenon – magnetic circular dichroism, discussed in section “[Magnetic Circular Dichroism](#).” For unit conversion factors, see Table 9.

Hypermagnetizabilities, Cotton–Mouton Effect

The *Cotton–Mouton effect* (CME) is a magnetic induction field-induced linear birefringence, observed in the presence of a strong magnetic induction field \mathbf{B} with a component perpendicular to the direction of propagation of the beam. It is mainly due to the tendency of both the electric field associated with the light beam and the magnetic induction field to align the molecules exhibiting both an anisotropic electric dipole polarizability and an anisotropic magnetizability tensor (according to the Langevin–Born theory (Born 1918; Langevin 1910)). The rearrangement of the electron density again plays a part through the mixed electric and magnetic hyperpolarizabilities, or hypermagnetizabilities.

The experimental quantity connected to the hypermagnetizability is the molar Cotton–Mouton constant ${}_mC$. Semiclassically, for fluids composed of closed-shell systems (Buckingham and Pople 1956; Fowler and Buckingham 1989; Jamieson 1991):

$${}_mC = \frac{2\pi N_A}{27} [\Delta\eta + Q(T)], \quad (183)$$

where $\Delta\eta$ is the hypermagnetizability anisotropy

$$\Delta\eta = \frac{1}{15} (3\eta_{\alpha\beta,\alpha\beta} - \eta_{\alpha\alpha,\beta\beta}), \quad (184)$$

and

$$Q(T) = \frac{1}{15kT} (3\alpha_{\alpha\beta}\xi_{\alpha\beta} - \alpha_{\alpha\alpha}\xi_{\beta\beta}). \quad (185)$$

In Eq. 185, α is the (frequency-dependent) electric polarizability tensor, and ξ the magnetizability tensor. The two contributions to the molar Cotton–Mouton constant in Eq. 183 are, therefore, a temperature-independent term related to the hypermagnetizability $\eta_{\alpha\beta,\gamma\delta}$ and a temperature-dependent molecular orientational part – the “Langevin term.”

When the magnetizability is expressed as a power series in the perturbing electric field:

$$\xi_{\alpha\beta}(\mathbf{F}) = \xi_{\alpha\beta} + \zeta_{\gamma,\alpha\beta} F_\gamma + \frac{1}{2} \eta_{\gamma\delta,\alpha\beta} F_\gamma F_\delta + \dots, \quad (186)$$

the expansion coefficients $\zeta_{\gamma,\alpha\beta}$ and $\eta_{\gamma\delta,\alpha\beta}$ define the molecular first and second hypermagnetizability. The first indices of ζ and η refer to the electric field and the last two to the magnetic field. In fluids, the experimentally measured property is related to the anisotropy of η . From the expansion above it can be seen that the static hypermagnetizability η corresponds to the fourth derivative of the molecular energy:

$$\eta_{\alpha\beta,\gamma\delta} = - \left. \frac{\partial^4 E(\mathbf{F}, \mathbf{B})}{\partial F_\alpha \partial F_\beta \partial B_\gamma \partial B_\delta} \right|_{\mathbf{F}, \mathbf{B} = \mathbf{0}}. \quad (187)$$

It can also be expressed in terms of $\xi(\mathbf{F})$ – the magnetizability dependent on \mathbf{F} :

$$\eta_{\alpha\beta,\gamma\delta} = \left. \frac{\partial^2 \xi_{\gamma\delta}}{\partial F_\alpha \partial F_\beta} \right|_{\mathbf{F} = \mathbf{0}}, \quad (188)$$

or $\alpha(\mathbf{B})$ – the polarizability dependent on \mathbf{B} :

$$\eta_{\alpha\beta,\gamma\delta}(\omega) = \left. \frac{\partial^2 \alpha(-\omega; \omega)_{\alpha\beta}}{\partial B_\gamma \partial B_\delta} \right|_{\mathbf{B} = \mathbf{0}}, \quad (189)$$

and we have indicated in the latter equation that it holds also for the frequency-dependent polarizability.

The hypermagnetizability is given as a sum of paramagnetic and diamagnetic contributions:

$$\eta_{\alpha\beta,\gamma\delta}(\omega) = \eta_{\alpha\beta,\gamma\delta}^{\text{para}}(-\omega; \omega, 0, 0) + \eta_{\alpha\beta,\gamma\delta}^{\text{dia}}(-\omega; \omega, 0), \quad (190)$$

corresponding to a sum of a cubic response and a quadratic response function (Rizzo et al. 1997):

$$\eta_{\alpha\beta,\gamma\delta}^{\text{para}}(-\omega; \omega, 0, 0) = - \left\langle \left(\hat{\mu}_\alpha; \hat{\mu}_\beta, \hat{m}_\gamma, \hat{m}_\delta \right) \right\rangle_{\omega, 0, 0}, \quad (191)$$

$$\eta_{\alpha\beta,\gamma\delta}^{\text{dia}}(-\omega; \omega, 0) = \left\langle \left(\hat{\mu}_\alpha; \hat{\mu}_\beta, \hat{\xi}_{\gamma\delta}^{\text{B,B}} \right) \right\rangle_{\omega, 0}, \quad (192)$$

where (see for comparison Eq. 145) the “diamagnetic magnetizability operator” in the second term is

$$\hat{\xi}_{\alpha\beta}^{\text{B,B}} = -2\hat{H}_{\alpha\beta}^{\text{B,B}}. \quad (193)$$

We recall here that the magnetic dipole moment and diamagnetic magnetizability operators, and thus the partitioning of the hypermagnetizability into para- and diamagnetic components, may depend on the chosen gauge origin.

Finite field approaches corresponding to Eqs. 188 and 189 have been used (Rizzo et al. 1997). In these calculations $\eta_{\alpha\beta,\gamma\delta}$ is obtained combining the available analytic linear response methods with the numerical finite difference approach. The frequency-dependent electric dipole polarizability can be computed in the presence of a magnetic induction field, and the hypermagnetizability is obtained by numerical differentiation. This approach (Cybulski and Bishop 1994; Tellgren et al. 2008) in general requires complex algebra as the perturbation is purely imaginary, whereas most electronic structure codes use real wave functions. In the more often used finite field approach based on Eq. 188, one computes the magnetizability ξ with and without an external electric field perturbation. However, in this approach only the static value of the required hypermagnetizability is obtained. Analytic calculations of the quadratic and cubic responses Eqs. 192 and 191 also including London orbitals to ensure gauge origin independence can nowadays be carried out (Thorvaldsen et al. 2008). The effect of the London orbitals on improving the basis set convergence appears to be small.

For atoms, choosing the origin of the gauge at the nucleus automatically makes the dependence on the gauge vanish, and there is no advantage in using LAOs. Moreover, the evaluation of $\Delta\eta$ may be reduced to the calculation of the Cauchy moments $S(-2n)$ and of the dipole–dipole–quadrupole hyperpolarizability frequency dispersion coefficients (Coriani et al. 1999a; Rizzo et al. 1997). For unit conversion factors, see Table 10.

Electric-Field-Gradient-Induced Birefringence

As a magnetic induction field can give rise to an optical anisotropy through its interaction with the anisotropic magnetizability of a molecule, so can an electric field gradient, through its coupling with a permanent molecular quadrupole moment. The corresponding birefringence (Buckingham 1959; Buckingham and Disch 1963) is commonly called *electric-field-gradient-induced birefringence*, more recently *Buckingham birefringence*, and it exhibits a linear dependence on the strength of the electric field gradient. Besides the orientational Langevin-Born-type effect, acting on quadrupolar molecules and due to both the electric field associated with the light

Table 10 Unit conversion factors for hypermagnetizability

	1 a.u.	SI	CGS
η	$(ea_0^2)^2 (e\hbar/m_e)^2 E_h^{-3}$	$2.98425 \times 10^{-52} \text{ C}^2 \text{ m}^2 \text{ J}^{-1} \text{ T}^{-2}$	$2.68211 \times 10^{-44} \text{ cm}^3 \text{ G}^{-2}$

wave and to the static external field gradient, it also involves the rearrangement of the electron density as a consequence of the electron interaction with the electric and magnetic wave vectors associated with the beam and with the externally applied electric field gradient. In this case, mixed electric-dipole, electric-quadrupole, and magnetic-dipole hyperpolarizabilities play their role in determining the strength of the effect. Buckingham birefringence has been often employed for the determination of molecular quadrupole moments (Buckingham 1959; Buckingham and Disch 1963).

For an ideal gas at constant pressure, for light propagating along the Z axis

$$\Delta n = n_X - n_Y = \frac{N_A \nabla F}{15V_m \varepsilon_0} s = \frac{3 \nabla F}{2V_m} mQ(\omega, T), \quad (194)$$

$$mQ(\omega, T) = \frac{2N_A}{45\varepsilon_0} s, \quad (195)$$

Where V_m is the molar volume, ∇F is the external electric field gradient arranged so that $\nabla F = \nabla F_{XX} = -\nabla F_{YY}$ and $\nabla F_{ZZ} = 0$. Equation 195 introduces the Buckingham constant mQ , which depends through the quantity s on the circular frequency of the light and on the temperature. For non-dipolar systems, where the quadrupole moment does not depend on the choice of origin,

$$s = b(\omega) + \frac{1}{kT} \Theta_{\alpha\beta} \alpha_{\alpha\beta}(-\omega; \omega), \quad (196)$$

Involving the molecular quadrupole moment $\Theta_{\alpha\beta}$ and the frequency-dependent electric dipole polarizability $\alpha_{\alpha\beta}(-\omega; \omega)$. The temperature-independent contribution is

$$b(\omega) = B_{\alpha\beta, \alpha\beta}(-\omega; \omega, 0) - \mathcal{B}_{\alpha, \alpha\beta, \beta}(-\omega; \omega, 0) - \frac{5}{\omega} \varepsilon_{\alpha\beta\gamma} J'_{\alpha, \beta, \gamma}(-\omega; \omega, 0), \quad (197)$$

And is a combination of three mixed hyperpolarizabilities:

$$B_{\alpha, \beta, \gamma\delta}(-\omega; \omega, 0) = \left\langle \left\langle \hat{\mu}_\alpha; \hat{\mu}_\beta, \hat{\Theta}_{\gamma\delta} \right\rangle \right\rangle_{\omega, 0}, \quad (198)$$

$$\mathcal{B}_{\alpha, \beta\gamma, \delta}(-\omega; \omega, 0) = \left\langle \left\langle \hat{\mu}_\alpha; \hat{\Theta}_{\beta\gamma}, \hat{\mu}_\delta \right\rangle \right\rangle_{\omega, 0}, \quad (199)$$

$$J'_{\alpha, \beta, \gamma}(-\omega; \omega, 0) = i \left\langle \left\langle \hat{\mu}_\alpha; \hat{M}_\beta, \hat{\mu}_\gamma \right\rangle \right\rangle_{\omega, 0}, \quad (200)$$

The expressions are somewhat more complex for dipolar fluids (Buckingham and Longuet-Higgins 1968), where the quadrupole moment becomes origin dependent. The formal expressions for s , Eq. 196, does not change provided that we refer to a

frequency-dependent origin for the quadrupole operator, commonly labeled as the “effective quadrupole center” (Rizzo and Coriani 2005; Rizzo et al. 2007). For any other choice of the origin we have

$$s = b(\omega) + \frac{1}{kT} \left\{ \Theta_{\alpha\beta\alpha\beta}(-\omega; \omega) - \mu_\alpha \left[A_{\beta,\alpha\beta}(-\omega; \omega) + \frac{5}{\omega} \varepsilon_{\alpha\beta\gamma} G'_{\beta\gamma}(-\omega; \omega) \right] \right\}. \quad (201)$$

The effective quadrupole center is defined as the point in space where the last term in Eq. 201, involving the molecular dipole moment μ_α and the mixed polarizabilities $A_{\alpha,\beta\gamma}(-\omega; \omega)$ and $G'_{\alpha\beta}(-\omega; \omega)$, vanishes.

Magnetic Circular Dichroism

For a medium that is isotropic in the absence of magnetic fields, the Faraday A , B , and C terms determining the MCD of a transition from the electronic state to the electronic state $|n\rangle$ are defined as (Buckingham and Stephens 1966; Schatz and McCaffery 1969; Stephens 1976):

$$A(0 \rightarrow n) = \frac{1}{2} [\langle n | \mathbf{m} | n \rangle - \langle 0 | \mathbf{m} | 0 \rangle] \cdot \mathcal{I}m(\langle 0 | \boldsymbol{\mu} | n \rangle \times \langle n | \boldsymbol{\mu} | 0 \rangle), \quad (202)$$

$$B(0 \rightarrow n) = \mathcal{I}m \left\{ \sum_{k \neq 0} \frac{\langle k | \mathbf{m} | 0 \rangle}{\omega_{k0}} \cdot \langle 0 | \boldsymbol{\mu} | n \rangle \times \langle n | \boldsymbol{\mu} | k \rangle + \sum_{k \neq n} \frac{\langle n | \mathbf{m} | k \rangle}{\omega_{kn}} \cdot \langle 0 | \boldsymbol{\mu} | n \rangle \times \langle k | \boldsymbol{\mu} | 0 \rangle \right\}, \quad (203)$$

$$C(0 \rightarrow n) = \frac{1}{2} \langle 0 | \mathbf{m} | 0 \rangle \cdot \mathcal{I}m(\langle 0 | \boldsymbol{\mu} | n \rangle \times \langle n | \boldsymbol{\mu} | 0 \rangle), \quad (204)$$

where an average over the molecular orientations has been made. These tensors correspond to the single residues of the tensors that determine the Verdet constant of the Faraday effect (section “Faraday Effect”). With the convention used here, the Faraday terms given in Eqs. 202, 203, and 204 are three times those of Buckingham and Stephens (1966). The A term exists only if either $|0\rangle$ or $|n\rangle$ is degenerate, whereas the C term is nonvanishing only if $|n\rangle$ is degenerate.

The B terms exist in all cases, independently of the appearance of any ground- and excited-state degeneracies. We can rearrange Eq. 203 recalling that $(\hat{\mu}_\alpha)_{lm}^* = (\hat{\mu}_\alpha)_{lm}$, $(\hat{m}_\alpha)_{lm}^* = -(\hat{m}_\alpha)_{lm}$, and $\mathcal{I}m(z) = -i(z - z^*)/2$ we find that the $B(0 \rightarrow n)$ term can be written in terms of a single residue of a quadratic response function (Coriani et al. 1999b; Olsen and Jørgensen 1985) as:

$$B(0 \rightarrow n) = i \varepsilon_{\alpha\beta\gamma} \left(\lim_{\omega \rightarrow \omega_{n0}} \hbar(\omega - \omega_{n0}) \langle \langle \hat{\mu}_\gamma; \hat{m}_\beta, \hat{\mu}_\alpha \rangle \rangle_{0,\omega} \right). \quad (205)$$

By analogy to the spectral representation expression for the single residue we may also write (Olsen and Jørgensen 1985):

$$\lim_{\omega \rightarrow \omega_{n0}} \hbar (\omega - \omega_{n0}) \langle \langle \hat{\mu}_\gamma; \hat{m}_\beta, \hat{\mu}_\alpha \rangle \rangle_{0,\omega} = M_{\mu_\alpha}^{0 \leftarrow n} M_{m_\beta \mu_\gamma}^{n \leftarrow 0}(0), \quad (206)$$

where $M_{\mu_\alpha}^{0 \leftarrow n}$ and $M_{m_\beta \mu_\gamma}^{n \leftarrow 0}(0)$ indicate specific one- and (formally) two-photon transition matrix elements between state $|0\rangle$ and state $|n\rangle$, respectively.

In a recent experimental study of the ethylene MCD spectrum (Snyder et al. 2004), the assignment of the transitions was revised. The proposed new assignment is in agreement with earlier ab initio calculations of the *B* term in ethylene (Coriani et al. 1999b). Recently, also calculations of the *A* term of magnetic circular dichroism appeared (Seth et al. 2004). We also note that both the *A* and *B* terms come out directly from calculations using the complex polarization propagator approach (Solheim et al. 2008).

Nuclear Spin-Related Properties

In the electronic Schrödinger equation defined by the Hamiltonian of Eq. 2, there are no terms dependent on the isotopic species of the nuclei. The electron–nucleus and nucleus–nucleus interactions are described only as interactions of point charges. In particular, no hyperfine interactions dependent on the nuclear magnetic moments are included. To determine molecular properties dependent on the nuclear magnetic dipole moments we thus considered an extension of this Hamiltonian, Eqs. 41, 42, and 43, assuming that the magnetically active nuclei are represented by point magnetic dipoles. The related properties of the nucleus – its spin and magnetogyric constant – enter the definitions of the molecular properties as known, constant factors.

All the effects which determine the existence of nuclear magnetic resonance (NMR) spectra are due to the magnetic moments of the nuclei. Each magnetically active nucleus contributes to the magnetic field through the magnetic vector potential Eq. 38. In the NMR spectrum we observe the effects due to the interaction of this locally induced magnetic field with the applied external magnetic field and with other magnetically active nuclei in the molecule. The richness of the NMR spectrum arises from the dependence of both these interactions on the electronic structure. Within a molecule, each nucleus is shielded by the electrons, and due to this shielding the effective, local magnetic field differs for different nuclei. Similarly, the interaction of two magnetic dipole moments – determining the spin–spin coupling in NMR spectra – depends not only on the distance between two nuclei, but also on the electronic structure of the molecule. Moreover, in contrast to the direct dipole–dipole interaction, this electron-mediated contribution – the indirect spin–spin coupling tensor – does not vanish in isotropic media. We focus in this section on the calculation of the linear response properties that determine these effects.

Due to the importance of NMR spectroscopy in modern chemical research, the shielding and spin–spin coupling constants are undoubtedly the most important properties which depend on the nuclear spin. Nevertheless, we analyze here also other phenomena proportional to the magnetic moment of a nucleus. The existence of an external magnetic field affects the rotational spectrum; similarly, splittings in the rotational spectrum will also appear due to the field associated with the nuclear magnetic dipole moment and these splittings are described by the nuclear spin–rotation constant (NSRC). Indeed, in high-resolution molecular beam and microwave spectroscopies, the effective Hamiltonians describing the rotational spectrum include the contributions to the rotational energy levels due to the presence of the spin–rotation interactions, as well as the interactions between two nuclear spin–magnetic moments, and the electric field gradients at the nuclei (see Vaara et al. (2002) and references therein for more details).

NMR Effective Spin Hamiltonian

Let us first consider the standard interpretation of the NMR spectrum. The parameters used to define it – the shielding constants and spin–spin coupling constants – are obtained from an effective NMR Hamiltonian \mathcal{H}^{NMR} . This operator acts in the space defined by all possible arrangements of the nuclear spins. The eigenvalues of \mathcal{H}^{NMR} determine the transition energies as the differences between different nuclear spin states in the molecule, and the eigenvectors determine the arrangement of the nuclear spins with respect to the external field. The effective NMR Hamiltonian is written as (Abragam 1961):

$$\mathcal{H}^{\text{NMR}} = -\sum_K \mathbf{B}^T (1 - \sigma_K) \mathbf{m}_K + \frac{1}{2} \sum_{K \neq L} m_K^T (\mathbf{D}_{KL} + \mathbf{K}_{KL}) \mathbf{m}_L, \quad (207)$$

where σ_K is the shielding tensor (see section “Nuclear Magnetic Shielding”) and \mathbf{K}_{KL} is the reduced indirect spin–spin coupling tensor (see section “NMR Indirect Spin–Spin Coupling”). We recall that the nuclear magnetic dipole moment is related to the nuclear spin, $\mathbf{m}_K = \mu_N g_K \mathbf{I}_K$, see Eq. 25. Finally, \mathbf{D}_{KL} is the direct interaction of the magnetic dipoles. The effective Hamiltonian does not include the \mathcal{H}^{Nq} operator, Eq. 122, which describes the interactions due to the quadrupole moment of the nucleus and the electric field gradient at the nucleus (as discussed in section “Electric Field Gradient at the Nucleus, Nuclear Quadrupole Coupling Constant,” \mathcal{H}^{Nq} can be related to the contributions quadratic in \mathbf{I}_K). The reason for this is that in isotropic media the nuclear quadrupole interaction does not contribute to the nuclear spin energy levels, but only to the relaxation processes in the nuclear spin system – thus, it only affects the shapes of the spectral lines, and not their positions.

The effective NMR Hamiltonian \mathcal{H}^{NMR} has no explicit dependence on the electronic structure. It includes two electron-independent contributions, the nuclear Zeeman term $-\mathbf{B}^T \cdot \mathbf{m}_k$ and the classical direct dipolar coupling of the nuclear magnetic moments $\mathbf{m}_K^T \mathbf{D}_{KL} \mathbf{m}_L$. All the effects of the electronic structure are

incorporated in the nuclear magnetic shielding tensors σ_K , and the reduced indirect nuclear spin–spin coupling tensors \mathbf{K}_{KL} .

In order to calculate the shielding and indirect spin–spin coupling tensors, we employ the perturbation expansion of the molecular Hamiltonian, discussed in section “[The Molecular Hamiltonian](#).” As indicated by Eq. 207, in the perturbation expansion of the Schrödinger equation we have to consider all contributions that are either linear or bilinear in \mathbf{B} and \mathbf{m}_K for the shielding constants, and those linear or bilinear in \mathbf{m}_K and \mathbf{m}_L for the indirect spin–spin coupling constants. Since the perturbing operators are time independent, both the shielding and spin–spin coupling tensors can be expressed as energy derivatives.

For a rotating molecule in an isotropic medium, the NMR spin Hamiltonian may be written as (assuming the external magnetic field is directed along the Z axis):

$$\mathcal{H}_{\text{iso}}^{\text{NMR}} = -\sum_K B (1 - \sigma_K) m_{K,Z} + \frac{1}{2} \sum_{K \neq L} K_{KL} \mathbf{m}_K^T \cdot \mathbf{m}_L. \quad (208)$$

The direct spin–spin couplings D_{KL} are purely anisotropic and vanish in the rotational averaging, while for the shielding and indirect spin–spin coupling tensors the averaging leads to what is commonly referred to as the nuclear shielding constants σ_K and the reduced indirect nuclear spin–spin coupling constants K_{KL} :

$$\sigma_K = \frac{1}{3} \text{Tr} \sigma_K, \quad (209)$$

$$K_{KL} = \frac{1}{3} \text{Tr} \mathbf{K}_{KL}. \quad (210)$$

The name “scalar coupling” is sometimes used for the indirect spin–spin interaction; this is rather unfortunate, and one should keep in mind that we discuss a tensor even when the individual components are not observed.

The numerical values of the shielding and indirect spin–spin coupling constants are determined in experiment by fitting the solutions of the nuclear spin equation with the effective Hamiltonian Eq. 208 to the observed spectrum. The NMR spin Hamiltonian is usually not expressed in terms of the reduced indirect spin–spin coupling tensors \mathbf{K}_{KL} , but rather in terms of the indirect spin–spin coupling tensors \mathbf{J}_{KL} , which are related to the reduced tensors as:

$$\mathbf{J}_{KL} = \frac{1}{h} \mu_N g_K \mu_N g_L \mathbf{K}_{KL}. \quad (211)$$

The reduced coupling tensors \mathbf{K}_{KL} are independent of the nuclear g factors, and therefore, in contrast to \mathbf{J}_{KL} , can be used to compare the strengths of the couplings between nuclei with different g_K values.

An example of an NMR spectrum is shown in Fig. 3. The positions of the three groups of peaks are determined by the shielding constants and the fine structure within each group of peaks is determined by the spin–spin coupling constants. Both

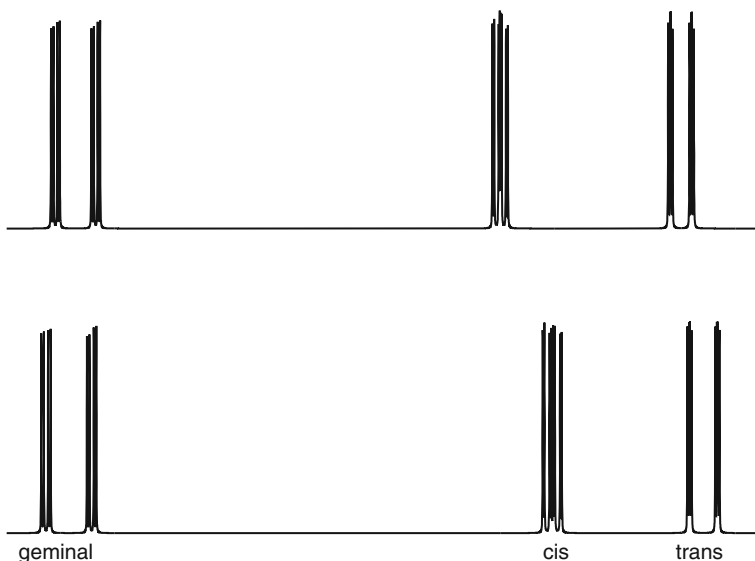


Fig. 3 Simulated 500 MHz proton NMR spectra of C_2H_3F . *Top* – ab initio, *bottom* – experimental values of shielding and spin–spin coupling constants. The labels define the position of the relevant proton with respect to the fluorine. The shielding constants and the Fermi contact contributions to the coupling constants were computed at the CCSD level, the other theoretical values are taken from Helgaker et al. (1997), where the experimental data were also discussed

proton spectra have been simulated assuming no coupling to C atoms (i.e., for ^{12}C). We note here that the changes in the energy levels due to the screening of the nuclear magnetic moments, reflected in the magnetic shielding constants, depend on the external magnetic field induction, whereas the changes induced by the spin–spin coupling constants (the splitting within each group of lines) are independent of the external magnetic field induction.

Nuclear Magnetic Shielding

The nuclear magnetic shielding tensor σ_K can be determined from the equation:

$$\left. \frac{\partial^2 E(\mathbf{B}, \mathbf{m}_K)}{\partial \mathbf{B} \partial \mathbf{m}_K} \right|_{\mathbf{B}=0, \mathbf{m}_K=0} = \sigma_K - \mathbf{1}, \quad (212)$$

where the second term represents the direct Zeeman interaction between the nucleus and the field. From the form of the energy derivative, as well as the bilinear nature of the interaction described by the shielding tensor appearing in the effective NMR Hamiltonian, it is clear that in the molecular Hamiltonian the terms that are linear and bilinear in the external magnetic field induction and the nuclear magnetic

moments should be considered. Collecting them and writing the contributions in terms of expectation values and response functions, we find for closed-shell systems:

$$\sigma_{K,\alpha\beta} = \left\langle 0 \left| \hat{H}_{\alpha\beta}^{B,K} \right| 0 \right\rangle + \left\langle \left\langle \hat{H}_{\alpha}^B; \hat{H}_{\beta}^{K(\text{PSO})} \right\rangle \right\rangle = \sigma_{K,\alpha\beta}^{\text{dia}} + \sigma_{K,\alpha\beta}^{\text{para}}. \quad (213)$$

As for the magnetizability, there are two contributions to the shielding: the so-called diamagnetic contribution σ_K^{dia} , which is obtained as the expectation value $\langle 0 | \hat{H}^{B,K} | 0 \rangle$ and the paramagnetic contribution σ_K^{para} , calculated as a linear response function $\langle \langle \hat{H}^B; \hat{H}^{K(\text{PSO})} \rangle \rangle$.

The shielding tensor is a nonsymmetric tensor of rank two. The principal axis system (PAS) is determined by the diagonalization of the symmetric part of the tensor. The shielding constant, already introduced in Eq. 209, is defined as the isotropic part; we can also define the anisotropy (Smith et al. 1992)

$$\sigma_{\text{anis}} = \sigma_{33}^{\text{PAS}} - \frac{1}{2} (\sigma_{11}^{\text{PAS}} + \sigma_{22}^{\text{PAS}}), \quad (214)$$

as well as the asymmetry:

$$\eta_{\text{asym}} = \frac{(\sigma_{22}^{\text{PAS}} - \sigma_{11}^{\text{PAS}})}{(\sigma_{33}^{\text{PAS}} - \sigma_{\text{iso}})}, \quad (215)$$

where it is assumed for the principal axis components of the shielding tensor that $\sigma_{33}^{\text{PAS}} \geq \sigma_{22}^{\text{PAS}} \geq \sigma_{11}^{\text{PAS}}$. Unfortunately, there are different conventions for the ordering of the principal shielding axes which can lead to much confusion when comparing literature values of anisotropies and asymmetries. To circumvent these problems it has been proposed that shielding tensor results should be reported in terms of the span and the skew, defined as (Mason 1993):

$$\Omega_{\text{span}} = \sigma_{33}^{\text{PAS}} - \sigma_{11}^{\text{PAS}}, \quad (216)$$

$$\kappa_{\text{skew}} = \frac{3(\sigma_{\text{iso}} - \sigma_{22}^{\text{PAS}})}{\Omega_{\text{span}}}. \quad (217)$$

The antisymmetric part does not affect the standard NMR spectrum (that is why the rank-1 tensor has been disregarded above); however, it affects the relaxation rates (Anet et al. 1990).

Since the magnetic dipole operator appears in the expression for the paramagnetic contribution to the shielding tensor, the calculated quantity will depend on the choice of a global gauge origin. Similarly, for the diamagnetic contribution an origin dependence can easily be seen from the definition of the operator $\hat{H}^{B,K}$ in Eq. 48. Hence, the partitioning of the nuclear shielding into dia- and paramagnetic contributions depends on the choice of global gauge origin. Any

interpretation attributed to the individual contributions is, therefore, limited to the applied computational approximation – in particular, care should be exercised when results for different nuclei in different molecules are compared. When the gauge origin is chosen to be the nucleus of study, the paramagnetic contribution to the shielding tensor is related to the electronic contribution to the nuclear spin-rotation constant; see Eq. 231 below. We note, however, that this relation is true only in a nonrelativistic approach.

The dependence of the dia- and paramagnetic contributions to the shielding on a global gauge origin means that in approximate calculations the nuclear magnetic shielding tensors will display a dependence on a gauge origin. This can be avoided through the introduction of local gauge origins, as described in section “Magnetic Properties.” In addition to the London orbital approach other methods have also been applied to enable the calculation of origin-independent nuclear magnetic shielding constants, such as the individual gauge for localized orbitals (IGLO) method developed by Kutzelnigg (1980; Schindler and Kutzelnigg 1982) and various continuous transformation of the origin of the current density (CTOCD) methods developed by Lazzeretti, Malagoli, and Zanasi (Lazzeretti and Zanasi 1996; Lazzeretti et al. 1994; Zanasi et al. 1995).

The shielding constant is dimensionless, and is expressed in ppm (parts per million). This unit reflects the role of the shielding in comparison to the electron-independent direct Zeeman interaction of the nuclear magnetic moment with the external field (Eq. 208). We can get a physical understanding of the nuclear shielding by considering that the electrons in the molecule create an additional magnetic field, acting as a correction proportional to the external field and leading to the local magnetic field

$$\mathbf{B}_K^{\text{loc}} = (\mathbf{1} - \sigma_K) \mathbf{B}. \quad (218)$$

When σ_K is positive, it means that the presence of electrons leads to a decreased local magnetic field experienced by the nuclear magnetic moments – that is, the electron density shields the nuclear magnetic moments from the external magnetic field, a concept that explains the origins of the term shielding constant (we note that shielding constants can also be negative). The values of the shielding constants for various elements differ widely, as do their variations with molecular structure. In general the variation of the shielding constants increases with increasing atomic number – ranging from 10–10² ppm for light atoms to $\approx 10^4$ ppm for Pb (see also Table 11). The range of shielding constants for a particular nucleus is relevant when we compare the computed and experimental spectra, for instance to assign the peaks in the NMR spectrum to specific nuclei. The required accuracy for the shielding of hydrogen may be ≈ 0.1 ppm, whereas for carbon or nitrogen nuclei it may be sufficient to obtain results accurate to ≈ 5 ppm.

Let us also briefly consider the shielding constants in paramagnetic molecules. In this case we have a total electron spin S which implies a set of $2S + 1$ degenerate spin states. These different states will be split when an external magnetic field is applied to the sample. The ensemble average of the projection of the spin operator

Table 11 Absolute shielding (in ppm) of selected nuclei in commonly used standard reference molecules^{a, b}

NMR standard		Primary source	Absolute shielding		Shift δ	
Nucleus, standard	Absolute shielding					
¹ H in TMS(l)	32.873	H ₂	26.2886	Sundholm and Gauss (1997)	6.584	Jackowski et al. (2010)
¹³ C in TMS(l)	186.37	CO	0.6	Raynes et al. 1989	185.77	Jackowski et al. (2000)
¹⁵ N in CH ₃ NO ₂ (l)	-135.8	NH ₃	264.54	Jameson et al. 1981	-400.34	Jameson et al. 1981
¹⁷ O in H ₂ O(l)	287.5	CO	-62.74	Wasylishen and Bryce (2002)	350.2	Wasylishen and Bryce (2002)

^aSee the references for experimental details and error bars

^bGasphase NMR or isolated molecule ab initio values for primary source, (l) = liquid

along the Z axis can be obtained in much the same way as the magnetization induced in a molecular sample containing unpaired electrons as:

$$\langle S_Z \rangle = -\frac{g_e}{2} \frac{S(S+1)}{3kT} B_Z. \quad (219)$$

The total magnetic moment of the electrons becomes directly proportional to the external magnetic field induction, and thus additional contributions to the nuclear shielding constant arise, since we need to consider new terms linear and bilinear in the nuclear and electronic magnetic moments. From the terms in the molecular Hamiltonian introduced in section “[Small Terms Due to the Vector Potential in the Hamiltonian](#)” we see that these are the hyperfine interactions involving the Fermi contact and spin–dipole operators, $\hat{H}^{K(\text{FC})}$ and $\hat{H}^{K(\text{SD})}$ given in Eqs. 51 and 52, respectively. Using the ensemble average for the spin–projection given in Eq. 219, we find that the contributions to the nuclear shielding constants due to these hyperfine interactions are (Rinkevicius et al. 2003):

$$\sigma_{K,\alpha\beta}^{\text{con,FC}} = \frac{4\mu_B^2}{3kT} S(S+1) \left\langle 0 \left| \hat{H}_{\alpha\beta}^{K(\text{FC})} \right| 0 \right\rangle, \quad (220)$$

$$\sigma_{K,\alpha\beta}^{\text{con,SD}} = \frac{4\mu_B^2}{3kT} S(S+1) \left\langle 0 \left| \hat{H}_{\alpha\beta}^{K(\text{SD})} \right| 0 \right\rangle. \quad (221)$$

These contributions are closely related to the hyperfine interactions observed in ESR spectroscopy (see section “[ESR Hyperfine Coupling Tensors](#)”). The above expressions have been derived assuming kT to be much larger than the separation

of the different electronic spin states. We have also for simplicity ignored higher-order corrections arising from the spin-orbit interaction, which for many open-shell species can be quite significant.

It is worth noticing that the so-called contact contributions to the nuclear shielding tensor, Eqs. 220 and 221, are temperature dependent, and thus behave differently than the orbital contributions. Experimentally, the contact contributions are often quite large and assumed to vary much more with molecular structure than the dia- and paramagnetic contributions, and thus these latter two terms are often assumed constant in different chemical species or taken to be identical to closely related closed-shell molecules. However, theoretical studies have shown that this is a rather crude approximation, and it may fail to correctly describe the total shifts due to the contact and orbital terms (Rinkevicius et al. 2003).

A direct comparison of calculated shielding constants with experiment is hampered by the difficulty in experimentally determining the absolute shielding constants calculated by theory. The lack of experimental data for absolute shielding constants has many origins: first, the shielding constant is defined as a correction to the direct interaction between the external magnetic field induction and the bare nuclear magnetic moment. However, it is impossible to experimentally determine this direct interaction. Moreover, determining the magnetic field induction of an NMR spectrometer with sufficient accuracy is difficult.

The shielding constant can be considered as a measure of the local field experienced by a nuclear magnetic moment due to the shielding of the external magnetic field by the electron density. The *relative* difference in the local magnetic fields around different nuclei provides similar information and for this reason experimental NMR studies report the *chemical shift* δ_K defined as:

$$\begin{aligned}\delta_K &= \frac{\sigma_K(\text{reference}) - \sigma_K(\text{sample})}{1 - \sigma_K(\text{reference})} \\ &\cong \sigma_K(\text{reference}) - \sigma_K(\text{sample})\end{aligned}\quad (222)$$

The trends for the absolute shielding are thus reversed to those for the chemical shift – that is, $\delta_K < \delta_L$ means that the nucleus K is more shielded than nucleus L . The chemical shifts are taken relative to the shielding in a selected molecule (see Table 11); for instance, liquid tetramethylsilane (TMS) is used as a standard reference for ^{13}C . Unfortunately, large molecules in solution or in condensed phase are chosen for many nuclei to define experimental chemical shifts. It is not yet possible to perform benchmark calculations for these standard systems; moreover, it is often not even possible to compute the shielding in these reference compounds at the same level of accuracy as for smaller molecules that are of interest.

There is a semi-experimental approach for determining absolute shielding scales, based on the relationship between the paramagnetic shielding constant and the nuclear spin-rotation constants, as indicated in Eq. 231 below. The absolute shielding constants have been used to obtain the magnetic dipole moments of bare nuclei from NMR spectra – in this case, it is not sufficient to measure the chemical shifts. Similarly to the electric quadrupole moments, improved values of the nuclear

magnetic dipole moments can now be determined applying increasingly accurate absolute shielding constants obtained in *ab initio* calculations for small molecules (Jaszuski and Jackowski 2008).

The use of perturbation-dependent basis sets is mandatory in theoretical studies of nuclear shielding constants. This is particularly important in order to ensure that the calculated shielding constants are independent of the choice of gauge origin. At the same time, for smaller basis sets, the use of perturbation-dependent basis sets improves basis set convergence compared to calculations only employing conventional basis sets. Moreover, the convergence may be significantly improved by applying property-optimized basis sets, such as the pcS-*n* basis sets optimized by Jensen for DFT calculations of the shielding constants (Jensen 2008).

For atomic systems, theory can determine rather accurately the absolute shielding constants. First, fairly large nonelectronic contributions such as the zero-point vibrational corrections are absent. Secondly, because of symmetry only the diamagnetic term – an expectation value, which is much easier to compute than an accurate value of the response function – contributes. For instance, for the hydrogen atom the nonrelativistic shielding constant can be calculated as:

$$\sigma^{\text{dia}} = \alpha_{\text{fs}}^2 \frac{e^2}{3m_e} \left\langle 0 \left| \frac{1}{r_{iK}} \right| 0 \right\rangle = 17.750454 \text{ ppm.} \quad (223)$$

In a similar manner, explicitly correlated calculation of the $1/r_{iK}$ expectation value leads to an absolute shielding constant for the helium atom of 59.9367794 ppm (Drake 2006) (the most accurate value obtained in a recent study including relativistic, quantum-electrodynamic, and nuclear mass effects is $\sigma_{\text{He}} = 59.96743$ ppm (Rudziński et al. 2009)). For all the rare-gas atoms the shielding constants have been computed using a four-component relativistic approach (Vaara and Pyykkö 2003). The calculations were performed within the Dirac–Hartree–Fock approximation and, if we assume that the electron correlation effects are small, they provide the shielding scale for the rare gas atoms.

The development of efficient methods – using perturbation-dependent basis sets and highly correlated wave functions for the calculation of nuclear magnetic shieldings – has allowed in practice to reach high accuracy also for shielding constants in molecules (see the reviews by Gauss and Stanton 2002 and Helgaker et al. 1999). For instance, the use of carefully chosen basis sets with the CCSD(T) method has enabled calculations of the equilibrium shielding constants with an accuracy of about 1–3 ppm for ^{13}C shieldings in a range of molecules (Auer et al. 2003). Most importantly, it has been demonstrated that systematic convergence of the results with improvement in the level of the calculation can be achieved. In order to obtain highly accurate shielding constants comparable with experiment, in addition to the effects of electron correlation, vibrational effects also have to be considered. The latter are obviously mandatory to estimate the isotopic substitution effects, since for a fixed molecular geometry the shielding constants are independent of the nuclear masses. The most accurate and reliable results are presently obtained

in this manner – within the CCSD(T) approach, applying GIAOs and taking into account the vibrational effects [see for instance a recent ^{19}F shielding study (Harding et al. 2008a)]. CCSD(T) shielding constants may be computed for increasingly large molecules, as shown by benzene calculations with basis sets including up to 474 functions (Harding et al. 2008b). Higher-order coupled-cluster methods have also been developed, but so far their application has been limited to small molecules (Kállay and Gauss 2004). At the HF or DFT level, shielding constants may nowadays be computed for molecules with more than 1000 atoms, with basis sets including over 10,000 functions (Kussmann and Ochsenfeld 2007; Ochsenfeld et al. 2004).

In practice, care should be exercised when using the data of Table 11 to analyze experimental chemical shifts and computed absolute shielding values. The differences between various levels of approximation in the theory are often difficult to foresee, and variations of the experimental conditions may also lead to unpredictable changes in the shielding. When a particular approximation is used in a sequence of calculations, a comparison of the calculated chemical shifts with the corresponding experimental data might be more appropriate. Moreover, whereas there is no guarantee that zero-point vibrational effects or changes due to intermolecular interactions will be identical for nuclei of interest in two molecules (or even within a single molecule), one may expect that these effects partly cancel out for two nuclei (Ruud et al. 2001), and thus that higher accuracy may be achieved for chemical shifts than for absolute shielding constants.

To obtain the shielding constants in ppm, the computed values in a.u. are multiplied by $10^6 \times \alpha_{\text{fs}}^2 \cong 53.25136$.

Shielding Derivatives

The nuclear shielding depends, as does the molecular energy itself, on additional sources of perturbations, such as electric and magnetic fields. Assuming these perturbations to be static, we may perform a Taylor expansion of the nuclear shielding tensor in the presence of the additional perturbations. As the simplest example, let us mention the perturbation-induced changes in the shielding in the presence of a static electric field. They can be described in terms of so-called shielding polarizabilities (a review was given by Raynes 1996), where for instance the first shielding polarizability is given by:

$$\left. \frac{\partial \sigma_K(\mathbf{F})}{\partial \mathbf{F}} \right|_{\mathbf{F}=\mathbf{0}} = \left. \frac{\partial^3 E(\mathbf{B}, \mathbf{m}_K, \mathbf{F})}{\partial \mathbf{B} \partial \mathbf{m}_K \partial \mathbf{F}} \right|_{\mathbf{B}, \mathbf{m}_K, \mathbf{F}=\mathbf{0}} = \left\langle \left\langle \hat{H}^{\text{B},K}; \boldsymbol{\mu} \right\rangle \right\rangle + \left\langle \left\langle \hat{H}^{\text{B}}; \hat{H}^{\text{K(PSO)}} \right\rangle, \boldsymbol{\mu} \right\rangle. \quad (224)$$

We note that there is a route to origin-independent shielding polarizabilities, as the combination of origin-independent linear response calculations of the shielding constant in the presence of a finite electric field can be used in a finite-field procedure. This allows for the calculation of origin-independent shielding polarizabilities using for instance London atomic orbitals (Rizzo and Gauss 2002).

Spin-Orbit Corrections to Nuclear Magnetic Shielding

The formalism presented above for calculating the nuclear shielding constants works well for molecules containing elements from the first two or three rows of the periodic table. However, when the molecule contains heavier nuclei, this formalism may fail to reproduce correctly even the qualitative features of all the shielding constants, also for light elements such as hydrogen or carbon in a molecule. A classic example of this failure of Eq. 213 is shown in Fig. 4, which illustrates the hydrogen shielding in the hydrogen halides. Whereas the hydrogen shielding observed experimentally increases strongly with increasing atomic number, only a small increase is observed when Eq. 213 is used to calculate the shielding constants. The reason for this failure can be traced to relativistic effects, and in the case of the hydrogen shielding in Fig. 4 to the effects of the spin-orbit interactions.

Some of the most important relativistic effects – such as the discussed spin-orbit contributions to the shielding of light elements bonded to heavy atoms – can be calculated using perturbation theory, beginning with the nonrelativistic treatment and considering the spin-orbit operator as an additional perturbation representing one of the relativistic corrections. The spin-orbit operator provides a coupling between the orbital magnetic moment of the electron, in the case of NMR induced by an external magnetic field induction, and either its own or another electron's spin-magnetic moment. This induced spin-magnetic moment may in turn interact with the nuclear magnetic moments through the Fermi contact (Eq. 51) or spin-dipolar (Eq. 52) mechanism. These contributions involve three perturbing

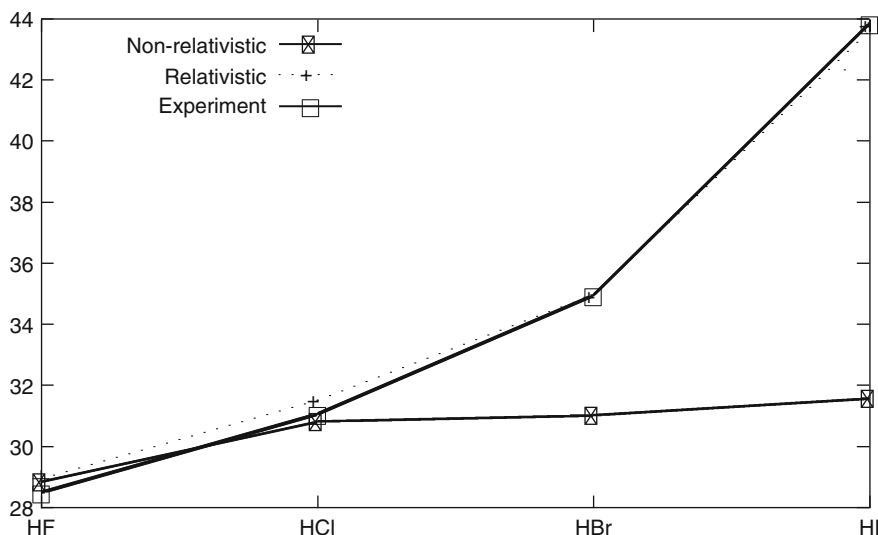


Fig. 4 The proton shielding constants (in ppm) in the HX series (X = F, Cl, Br, I) calculated with and without spin-orbit coupling contributions. Also included are the available experimental data. The theoretical data are CASSCF results from Vaara et al. (1998)

operators, and they are thus proportional to quadratic response functions such as $\langle\langle \hat{H}^{K(\text{FC})} + \hat{H}^{K(\text{SD})}; \hat{H}^{\text{SO}}, \hat{H}^{\text{B}} \rangle\rangle$.

In addition, in the presence of an external magnetic field there is a term in the spin-orbit operator which is proportional to the external magnetic field induction (we recall here the discussion of the general form of the spin-orbit operators in section “[Small Terms Due to the Vector Potential in the Hamiltonian](#),” Eqs. 58, 59, and 60). The operator $\hat{H}^{\text{SO,B(n)}}$ can in turn interact with the Fermi contact and spin-dipole operators, leading to second-order contributions to the shielding tensor which are proportional to triplet linear response functions (Vaara et al. 1999) of the form $\langle\langle \hat{H}^{K(\text{FC})} + \hat{H}^{K(\text{SD})}; \hat{H}^{\text{SO,B(n)}} \rangle\rangle$.

For the shielding constants of the heavy elements themselves, relativistic effects arising from the other terms in the Breit–Pauli Hamiltonian (see section “[The Molecular Breit–Pauli Hamiltonian](#)”) are in general found to be more important than the spin-orbit corrections (Manninen et al. 2003). In general, relativistic effects on the heavy-atom shielding are due to a large number of contributions (Fukui et al. 1996; Ruiz de Azúa et al. 2003) in contrast to the hydrogen shieldings in molecules containing heavy elements, which are dominated by the spin-orbit corrections. Although there have been several studies of scalar relativistic effects on shielding, yielding results in good agreement with four-component Dirac–Hartree–Fock methods, the Breit–Pauli Hamiltonian is not variationally bound and thus the perturbation theory approach may fail for heavy elements.

Nuclear Spin Rotation Constants

The interaction between the nuclear magnetic dipole and the effective magnetic field of a rotating molecule at that nucleus leads to a splitting in the rotational spectrum (Flygare 1974; Gauss et al. 1996). This splitting is usually expressed in terms of the nuclear spin \mathbf{I}_K and the total rotational angular momentum \mathbf{J} as

$$\Delta E_K = -\mathbf{I}_K^T \mathcal{M}_K \mathbf{J}, \quad (225)$$

where $\mathcal{M}_K = \mathcal{M}_K^{\text{el}} + \mathcal{M}_K^{\text{nucl}}$ is the nuclear spin–rotation tensor in the Born–Oppenheimer approximation. More precisely, the energy shift depends on the nuclear magnetic moment interacting with the magnetic field induced by the molecular rotation, so we shall proceed from the equivalent equation,

$$\Delta E_K = -\frac{2\pi}{\mu_{\text{N}} g_K} \mathbf{m}_K^T \mathcal{M}_K \mathbf{J}. \quad (226)$$

The electronic contribution $\mathcal{M}_K^{\text{el}}$ is a second-order property

$$\mathcal{M}_K^{\text{el}} = -\frac{\mu_{\text{N}} g_K}{2\pi} \left. \frac{\partial^2 E^{\text{el}}}{\partial \mathbf{m}_K \partial \mathbf{J}} \right|_{\mathbf{m}_K, \mathbf{J}=0}, \quad (227)$$

and thus it can be obtained from the linear response function $\mathbf{I}^{-1} \langle \langle \hat{H}^{K(\text{PSO})}; \hat{I}_K \rangle \rangle$, where \mathbf{I} is the moment-of-inertia tensor and the angular momentum is evaluated with respect to nucleus K . The nuclear contribution is

$$\mathcal{M}_K^{\text{nucI}} = \alpha_{\text{fs}}^2 \mu_{\text{NG}K} \sum_{L \neq K} Z_L \frac{R_{LK}^2 \mathbf{1} - \mathbf{R}_{LK} \mathbf{R}_{LK}^T}{R_{LK}^3} \mathbf{I}^{-1}. \quad (228)$$

The principal axes of the nuclear spin–rotation tensor are determined by the molecular moment of inertia. Indeed, the magnetic moment induced by the rotational motion is well defined only in the rotation frame. Spin–rotation constants are related to the difference between the total shielding (for instance obtained using LAOs) and the diamagnetic part of the shielding derived from a conventional calculation with \mathbf{R}_K as the global gauge origin, $\sigma_K^{\text{dia}}(\mathbf{R}_K)$,

$$\mathcal{M}_K^{\text{el}} = \mathcal{M}_K - \mathcal{M}_K^{\text{nucI}} = \frac{\hbar}{\mu_{\text{B}}} \frac{\mu_{\text{NG}K}}{2\pi} [\sigma_K - \sigma_K^{\text{dia}}(\mathbf{R}_K)] \mathbf{I}^{-1} \quad (229)$$

in analogy with our definition of the paramagnetic magnetizability, Eq. 149. Hence, $\mathcal{M}_K^{\text{el}}$ is related to the correspondingly defined paramagnetic shielding:

$$\sigma_K^{\text{para}}(\mathbf{R}_K) = \langle \langle -\mathbf{m}(\mathbf{R}_K); \hat{H}^{K(\text{PSO})} \rangle \rangle, \quad (230)$$

as (Flygare 1974):

$$\sigma_K^{\text{para}}(\mathbf{R}_K) = \frac{\mu_{\text{B}}}{\hbar} \frac{2\pi}{\mu_{\text{NG}K}} \mathcal{M}_K^{\text{el}} \mathbf{I}. \quad (231)$$

The additional contributions to the spin–rotation constants arising in paramagnetic molecules have been discussed by Minaev et al. (2003).

As the spin–rotation interaction splits lines in the rotational spectrum, the unit of the spin–rotation constant is kHz.

NMR Indirect Spin–Spin Coupling

The interaction of the magnetic dipoles of two nuclei, including also the classic dipolar interaction between the bare nuclei, can be determined (compare section “NMR Effective Spin Hamiltonian”) from perturbation theory as the second derivative of the energy

$$\left. \frac{\partial^2 E(\mathbf{m}_K, \mathbf{m}_L)}{\partial \mathbf{m}_K \partial \mathbf{m}_L} \right|_{\mathbf{m}_K=0, \mathbf{m}_L=0} = \mathbf{K}_{KL} + \mathbf{D}_{KL}, \quad (232)$$

where \mathbf{K}_{KL} is the reduced indirect spin–spin coupling tensor given as the derivative of $E^{\text{el}}(\mathbf{m}_K, \mathbf{m}_L)$, and \mathbf{D}_{KL} describes the direct interaction of two nuclear dipoles. The reduced spin–spin coupling constants, independent of the nuclear magnetogyric ratios, should be used (for instance) to compare the strength of the coupling between pairs of different nuclei, and to analyze relative accuracies of calculated constants.

In standard gas- or liquid-phase NMR spectra the interaction of magnetic dipoles leads to a splitting of the lines of the spectrum, described by the NMR indirect spin–spin coupling constant J_{KL} . This constant is the trace of the tensor (Eq. 211):

$$\mathbf{J}_{KL} = \frac{1}{h} \mu_{\text{N}} g_K \mu_{\text{N}} g_L \mathbf{K}_{KL}.$$

The total spin–spin coupling tensor includes (see, e.g., Vahtras et al. 1992b) the so-called diamagnetic spin-orbit contribution $-\langle 0 | \hat{H}^{K,L(\text{DSO})} | 0 \rangle$ – obtained as the expectation value of the corresponding operator (Eq. 49), and a sum of linear response terms $\langle \langle \hat{H}^{K(\text{PSO})}; \hat{H}^{L(\text{PSO})} \rangle \rangle$ and $\langle \langle \hat{H}^{K(\text{FC})} + \hat{H}^{K(\text{SD})}; \hat{H}^{L(\text{FC})} + \hat{H}^{L(\text{SD})} \rangle \rangle$.

The paramagnetic spin-orbit operator $\hat{H}^{K(\text{PSO})}$ (Eq. 46) does not depend on electron spin, whereas the Fermi contact $\hat{H}^{K(\text{FC})}$ and spin–dipole $\hat{H}^{K(\text{SD})}$ operators (Eqs. 51 and 52) are electron-spin dependent. Therefore, as shown for instance by the sum-over-states description, in the calculation of the PSO term only singlet excited states have to be considered. In contrast, the FC–FC, SD–SD, and mixed FC–SD contributions depend on the triplet excited states. This leads to computational difficulties when the reference state function is not stable with respect to triplet perturbations. For instance, when the HF approximation is used to describe the reference state, completely meaningless results may be obtained.

\mathbf{J}_{KL} is an asymmetric tensor, and the different terms (DSO, PSO, SD, FC, and mixed SD–FC) yield contributions of different symmetry. For example, the term $\langle \langle \hat{H}^{K(\text{FC})}; \hat{H}^{L(\text{FC})} \rangle \rangle$ is fully isotropic, whereas the terms $\langle \langle \hat{H}^{K(\text{FC})}; \hat{H}^{L(\text{SD})} \rangle \rangle$ are fully anisotropic. Therefore, if we are only interested in the isotropic value of the spin–spin coupling constant, we can express it as the sum of four contributions:

$$J_{KL} = J_{KL}^{\text{DSO}} + J_{KL}^{\text{PSO}} + J_{KL}^{\text{SD}} + J_{KL}^{\text{FC}}, \quad (233)$$

and we do not need to compute the SD–FC terms. They have to be included when all the tensor components are required (see Buckingham and Love (1970), Buckingham et al. (1982), Jameson (1987) for details).

In addition to NMR spectra, microwave spectra also may provide information on the spin–spin coupling (Vaara et al. 2002). In particular, for diatomic molecules accurate data can be obtained from molecular beam electric and magnetic resonance when the splitting of a single rotational–vibrational state can be observed. One of the parameters of this splitting (usually called c_4) is equal to the isotropic value of the indirect spin–spin coupling tensor, another one (c_3) depends both on the direct coupling ($\propto R_{KL}^{-3}$) and on the anisotropy of the tensor.

In contrast to the shielding constants, London orbitals are neither needed nor used in any calculation of spin–spin coupling constants. There is no dependence

Table 12 The Fermi contact contribution to the spin–spin coupling constant of HD (in Hz), calculated using the CCSD Polarization Propagator Approximation (Oddershede et al. 1988)

Basis set ^a	Added <i>s</i> -function exponent	J^{FC}
10s6p2d		38.98
11s6p2d	5000	39.24
12s6p2d	15,000	39.69
13s6p2d	50,000	39.77
14s6p2d	150,000	39.93
15s6p2d	500,000	39.94

^aGaussian-type functions. The effects of additional functions of *p*- and *d*-type symmetry are negligible

of the perturbing operators on the external magnetic field and thus no gauge origin problem.

The basis-set dependence of the Fermi contact term has been often a source of problems in the calculation of spin–spin coupling constants. To describe properly the response to the $\delta(\mathbf{r}_{iK})$ perturbation we need a basis set which is sufficiently flexible in the region of the nucleus. An illustration of the slow convergence of computed J^{FC} values for the simplest example – the HD molecule – is given in Table 12. We can see that practically only the last added function does not affect the result, whereas there are very few other properties where an *s*-type function with so high exponents as 150,000 would be needed for the H atom.

In general, two modifications of the standard energy-optimized basis sets are needed: the standard functions should be uncontracted and tight *s*, *p*, and *d* functions should be added. The *s*-type functions are of particular importance for the Fermi contact contribution, whereas other functions may be more relevant for the other contributions. Various basis sets constructed in this way have been successfully used in ab initio and DFT calculations; further improvement of convergence with the basis set extension is achieved reoptimizing the exponents and contraction coefficients (Benedikt et al. 2008; Jensen 2006).

In Table 13 we present, as an example, the different contributions to all the coupling constants in acetylene. The dominant role of the FC term is typical for many molecules, in particular for large coupling constants in organic compounds. The SD contribution is often small, and therefore (being most expensive to calculate) it is sometimes neglected or analyzed at a lower level of accuracy than the other terms. In some cases such an approach is, however, not possible. For instance, the FC, PSO, and SD terms, respectively, dominate for different coupling constants in ClF_3 (Jaszuński 2004).

To illustrate the accuracy of the results that may be achieved for small molecules, we present in Table 14 the values for all the coupling constants in acetylene. The comparison with experimental data clearly demonstrates that it is necessary to consider the effects of vibration and rotation (zero-point vibrations and temperature).

Finally, we note that J_{KL} is very sensitive to variations in the molecular geometry. It has been repeatedly observed that the best results are obtained when the same approximation (for instance DFT functional and basis set) is used for geometry

Table 13 C₂H₂, all the contributions to the spin-spin coupling constants (in Hz)^a

	¹ J(CC)	¹ J(CH)	² J(CH)	³ J(HH)
DSO	0.01	0.31	-1.35	-3.59
PSO	6.67	-0.84	5.58	4.80
SD	9.04	0.43	1.02	0.57
FC	172.34	247.40	47.77	9.43
Total	188.07	247.29	53.02	11.22

^aJaszuński and Ruud (2001), RASSCF-1/cc-pCVQZsu2 results

Table 14 C₂H₂, indirect spin-spin coupling constants (in Hz)

	¹ J(CC)	¹ J(CH)	² J(CH)	³ J(HH)
Experiment, measured ^a	174.78 ± 0.02	247.56 ± 0.05	50.14 ± 0.05	9.62 ± 0.05
Nuclear motion contributions ^b	-10.255	4.866	-3.703	-1.272
Equilibrium geometry, from experiment ^c	185.04	242.70	53.84	10.89
Equilibrium geometry, ab initio, MCSCF ^d	184.68	244.27	53.08	10.80

^aGasphase, Jackowski et al. (2000)

^bWigglesworth et al. (2000)

^cCalculated using the nuclear motion contributions given above

^dJaszuński and Ruud (2001). FC term – RASSCF-4; other contributions – RASSCF-1

optimization and for the following calculation of spin-spin coupling constants. We refer to the recent reviews discussing the calculation of spin-spin coupling constants (Aucar 2008; Helgaker et al. 2008; Krivdin and Contreras 2007) for a discussion of many practical issues and numerous examples of successful applications.

The unit of J is Hz; K is given in $\text{kg m}^{-2} \text{s}^{-2} \text{A}^{-2} = \text{N A}^{-2} \text{m}^{-3} = \text{T}^2 \text{J}^{-1}$ (Raynes 1992).

Electron Spin-Related Properties

Electron spin resonance (ESR) and NMR are in principle very similar from a theoretical point of view, involving excitations and de-excitations in the electron and nuclear spin space, respectively. Thus, the effective Hamiltonians have much the same appearance in ESR and NMR theory, and the structure of the interactions is also very similar from a purely theoretical point of view.

However, from a computational viewpoint, there is a large difference between electron and nuclear spin interactions. In the case of the nuclear spin interactions, the Born-Oppenheimer approximation allowed us to decouple the nuclear spin wave function from the electronic wave function. This is not possible for the electron spin; indeed, the change in the electron spin state directly affects also the spatial part of the total electronic wave function. This is in principle not a problem. In *spin-unrestricted* approaches, the α and β electrons are treated separately, and changes

of electron spin through perturbation theory expressions can quite straightforwardly be included. However, since the spatial part of the orbitals of the α and β spins may differ, the total electron spin is no longer necessarily a good quantum number (as required for the exact wave function), giving rise to what is often referred to as *spin contamination*. Spin contamination can be avoided forcing the spin to be a good quantum number, for instance by demanding that the spatial orbitals for the α and β electrons are the same and writing the wave function as a fixed linear combination of Slater determinants referred to as configuration-state functions (CSFs). We also note that many open-shell wave functions of low and intermediate spin have strong multireference character, putting severe constraints on the wave function models that can be used. An alternative approach to address low and intermediate spin states is through the use of broken symmetry states (Noodleman 1981).

We recall that the Hamiltonian, Eq. 2, does not involve the electron spin. Spin-restricted implementations of electronic structure methods, therefore, only optimize the wave function parameters of the same spin symmetry as the reference wave function. Since operators that change the spin symmetry are known not to contribute to the ground-state energy (at least in the nonrelativistic picture in which spin-orbit effects are ignored), such spin-breaking wave function parameters are left undefined. This has the consequence that average values of triplet-perturbing operators cannot be accurately calculated in spin-restricted approaches, and an alternative approach to circumvent this problem has been proposed (Fernandez et al. 1992). Furthermore, in the calculation of second-order properties, such as for instance spin–spin coupling constants discussed in section “NMR Indirect Spin–Spin Coupling,” triplet instabilities in the wave function may occur. We have discussed this problem in section “NMR Indirect Spin–Spin Coupling” and we will not dwell further on this here. We only note that these considerations apply to all properties discussed in the following sections.

Spin-Orbit Coupling Constants

The spin–orbit coupling constant is obtained from the single residue of a triplet linear response function:

$$\lim_{\omega \rightarrow \omega_f} \hbar (\omega - \omega_f) \langle \langle \hat{H}^{\text{SO}}; V^{\omega_f} \rangle \rangle_{\omega}, \quad (234)$$

where V^{ω_f} is an arbitrary operator (corresponding to the excitation vector). There are two terms in the spin-orbit interaction operator \hat{H}^{SO} (Eq. 55); the contribution to the spin-orbit coupling from the two-electron part is usually smaller than the one-electron contribution, though it is as large as 30% for the spin–orbit coupling constant in triplet O₂ (Vahtras et al. 1992c).

The two-electron spin-orbit integrals calculated using the operators in Eqs. 13 and 14 of the Breit–Pauli Hamiltonian have three spatial components and lower symmetry than conventional two-electron repulsion integrals, and the cost of

calculating these integrals is therefore very high. In order to reduce the costs of computations of spin-orbit effects, building on the observation that the one-electron contribution dominates, either scaled one-electron spin-orbit approximations are used (in which an effective nuclear charge is introduced) (Kosegi et al. 1995) or atomic mean-field approaches are utilized (Hess et al. 1996). In the latter approach, the two-electron interactions are included in a mean-field sense, in principle similar to the way the two-electron repulsion is included in a mean-field sense in Hartree–Fock calculations.

The spin-orbit coupling is the major intramolecular interaction mechanism responsible for intersystem crossings, such as singlet–triplet transitions. It should be kept in mind that, in contrast to, for example, external fields, \hat{H}^{SO} does not describe a perturbation which has a well-defined physical meaning and arbitrarily modified strength. It is our description of the system which provides first a set of singlet and triplet states and only next includes \hat{H}^{SO} and yields the values of the spin-orbit coupling constants. Indeed, within a relativistic framework, the spin-orbit interaction appears naturally and is included variationally in the optimization of the wave function. The spin-orbit effects often differ quite substantially in their behavior from the scalar relativistic effects (as for instance illustrated in the case of the hydrogen shielding constants in molecules containing heavy elements). It is therefore important, in order to get physical insight into the role of various relativistic effects, that spin-orbit effects can be addressed separately from other effects in a four-component framework, and an approach for achieving such a decomposition of the relativistic effects at the four-component level has been proposed (Visscher and Saue 2000).

The spin-orbit coupling constant is normally reported in cm^{-1} .

Phosphorescence

When the spin-orbit operator \hat{H}^{SO} is taken into account, the ground and excited state, which we assume here to be of singlet and triplet spin multiplicity, respectively, are no longer pure spin states – $|^1 0^{(0)}\rangle$ and $|^3 f^{(0)}\rangle$ – but are contaminated by, to first order in perturbation theory:

$$|^3 0^{(1)}\rangle = -\sum_{3_n} \frac{|^3 n^{(0)}\rangle \langle ^3 n^{(0)} | \hat{H}^{\text{SO}} | ^1 0^{(0)}\rangle}{\omega_n}, \quad (235)$$

$$|^1 f^{(1)}\rangle = \sum_{1_n} \frac{|^1 n^{(0)}\rangle \langle ^1 n^{(0)} | \hat{H}^{\text{SO}} | ^3 f^{(0)}\rangle}{\omega_f - \omega_n}. \quad (236)$$

Including these first-order corrections to the ground- and excited-state wave functions in the expression for the dipole transition moment, we find that the first-order nonvanishing contribution to the dipole transition moment between a singlet and a

triplet state may be written in terms of a residue of a quadratic response function (Olsen and Jørgensen 1995; Vahtras et al. 1992a):

$$\langle {}^1 0 | \boldsymbol{\mu} | {}^3 f \rangle^{(1)} = \lim_{\omega \rightarrow \omega_f} \hbar (\omega - \omega_f) \left\langle \left\langle \boldsymbol{\mu}; \hat{H}^{\text{SO}}, V^\omega \right\rangle \right\rangle_{0, \omega} / \langle {}^3 f | V^\omega | {}^1 0 \rangle, \quad (237)$$

where V^ω is an arbitrary triplet operator (determining the excitation vector) and ω matches the singlet-triplet excitation energy.

The phosphorescent radiative lifetime τ_k of the k th triplet spin component of state $|{}^3 f\rangle$ is obtained from the relation:

$$\frac{1}{\tau_k} = \frac{4\omega_f^3 \alpha_{\text{fs}}^3}{3\hbar} \sum_{\nu} \left| \langle {}^1 0 | \hat{\mu}_\nu | {}^3 f_k \rangle \right|^2, \quad (238)$$

and thus depends not only on the transition moment from the ground to the excited state, but also on the excitation energy. The transition moments may vary for different polarizations of the light, and the average phosphorescent lifetime is dominated by the shortest lifetime, corresponding to the polarization with the largest partial transition rate. The total transition rate is given in s^{-1} , the total lifetime being its inverse.

ESR Effective Spin Hamiltonian

The effective ESR Hamiltonian, including also contributions from the nuclear spins and nuclear quadrupole moments (the latter primarily observable in electron-nuclear double resonance – ENDOR – spectroscopy), may be written as (Harriman 1978):

$$\mathcal{H}^{\text{ESR}} = \mathcal{H}^{\text{ex}} + \mathcal{H}^{\text{SZ}} + \mathcal{H}^{\text{ZFS}} + \mathcal{H}^{\text{hf}} + \mathcal{H}^{\text{nZ}} + \mathcal{H}^{\text{Nq}}. \quad (239)$$

Like the NMR effective Hamiltonian, the ESR effective Hamiltonian contains no reference to the electronic structure of the molecule, and the different contributions to the ESR Hamiltonian thus describe phenomenologically the interactions present in the molecule.

The first term in Eq. 239, \mathcal{H}^{ex} , contributes to the overall energy of interaction between electronic spin states, but the energy differences are in general too large to be observed by ESR spectroscopy. This exchange contribution accounts for the difference in energy between states of different multiplicity, and the Hamiltonian

$$\mathcal{H}^{\text{ex}} = -2\mathcal{J} \mathbf{s}_1 \cdot \mathbf{s}_2, \quad (240)$$

is often referred to as the Heisenberg Hamiltonian. It is important in the analysis of the spin states in open-shell species. However, since to a large extent it defines an intrinsic relation between different electronic states of a molecule, we do not consider it to describe a molecular property in the sense we have defined, and we will, therefore, not discuss this contribution any further.

The second term in Eq. 239, \mathcal{H}^{sZ} , is the electron spin–Zeeman interaction describing the interaction between the spin-magnetic moment of the electron and the external magnetic field induction

$$\mathcal{H}^{sZ} = \mu_B \mathbf{S}^T \mathbf{g} \mathbf{B}, \quad (241)$$

where \mathbf{g} is the electron g tensor. We use here \mathbf{S} to indicate the total effective spin of the system. For a free electron, the strength of this interaction would be determined by Eq. 6 and thus be given by the free electron g factor ($g_e \approx 2.0023193$, in which the quantum-electrodynamic corrections to the electron g factor are included). However, unpaired electrons in a molecule experience a local magnetic field, arising from a partial shielding (or deshielding) due to the other electrons in the molecule, which leads to a shift relative to g_e , and therefore:

$$\mathbf{g} = g_e \mathbf{1} + \Delta \mathbf{g}. \quad (242)$$

\mathcal{H}^{ZFS} is the spin–spin interaction:

$$\mathcal{H}^{ZFS} = \mathbf{S}^T \mathbf{D} \mathbf{S}, \quad (243)$$

which gives rise to the zero-field splitting of the ESR spectrum, since it can be observed even without an external magnetic field. It is described by the dipole interaction between the spin-magnetic dipoles of two unpaired electrons, and is thus only present in states with more than one unpaired electron, such as for instance triplet states. This term will in general lead to a splitting of the different spin sublevels in the triplet state.

The electronic g tensor and the electron spin–spin interactions (if present) are in most cases the dominating contributions to the ESR spectrum. The interactions of the spin-magnetic moments of the electrons with the nuclear magnetic moments give rise to hyperfine interaction in the ESR spectrum described by \mathcal{H}^{hf} :

$$\mathcal{H}^{hf} = \sum_K (A_K^{fc} \mathbf{S}^T \cdot \mathbf{I}_K + \mathbf{S}^T \mathbf{A}_K^{sd} \mathbf{I}_K). \quad (244)$$

There are thus hyperfine interactions arising from all nonzero nuclear magnetic moments. We note from Eq. 244 that the hyperfine interactions in general can be divided into two contributions. The first term is purely isotropic and has the form of a contact interaction, whereas the second term A_K^{sd} corresponds to the dipolar interaction of the electron spin-magnetic moment and the nuclear magnetic moment, and is purely anisotropic. The isotropic hyperfine interaction measures the spin density at a given nucleus.

The last terms in Eq. 239 are in general not observable in ESR spectroscopy, because of the increased linewidths due to the presence of the unpaired electrons. The nuclear Zeeman interaction is given by:

$$\mathcal{H}^{nZ} = -\mu_N \sum_K g_K \mathbf{B}^T \cdot \mathbf{I}_K, \quad (245)$$

and is analogous to the interaction included in the NMR effective Hamiltonian Eq. 207, where the shift with respect to the nuclear g factor was described in terms of the shielding constant. The final term, the nuclear quadrupole interaction \mathcal{H}^{Nq} , also only involves the two nuclear spins and has already been discussed in section “Electric Field Gradient at the Nucleus, Nuclear Quadrupole Coupling Constant.”

ESR Electronic g -Tensor

In this section we concentrate on the part of the ESR effective Hamiltonian describing the interaction between the electron spin and the external magnetic field, Eq. 241. In the case of molecules with a single open-shell electronic doublet state, as well as high-spin radicals in the strong field limit, the energy difference between the eigenvalues of the effective ESR Hamiltonian can be written as

$$\Delta E = \mu_{\text{B}} \sqrt{\mathbf{B}^{\text{T}} \mathbf{G} \mathbf{B}}, \quad (246)$$

where we have introduced the so-called symmetric \mathbf{G} tensor, measured in experiment, and defined as $\mathbf{G} = \mathbf{g} \cdot \mathbf{g}^{\text{T}}$. The shift in the g factor may be anisotropic, and thus the \mathbf{g} tensor is a 3×3 nonsymmetric tensor. Both \mathbf{G} and \mathbf{g} tensors provide the same information and are thus suitable for analyzing EPR data. We emphasize that the effective Hamiltonian employed here requires the use of degenerate perturbation theory for a proper analysis, since we consider states that are degenerate in the absence of the perturbation. We do not discuss these aspects in detail here, referring instead to Löwdin and Goscinski (1999) for a general introduction to this topic, and to Rinkevicius et al. (2008) for the g tensor in particular.

Within degenerate perturbation theory at the nonrelativistic level, there are in principle two contributing terms arising from expectation values of the spin–Zeeman (Eq. 50) and the magnetic dipole operator (Eq. 44), respectively. The latter can be shown to be zero, and the effect of the spin–Zeeman operator is to recover the free-electron g factor, g_{e} . Thus, within a purely nonrelativistic picture, there would be no effect of the electronic structure of the molecule on the interaction between an external magnetic field and the magnetic moment of the unpaired electron.

Including relativistic effects to order α_{fs}^2 leads to three corrections to the electronic g tensor (we consider here only the shift Δg relative to the free-electron value):

$$\Delta \mathbf{g} = \Delta \mathbf{g}^{\text{SO}} + \Delta \mathbf{g}^{\text{RMC}} + \Delta \mathbf{g}^{\text{GC}} + \mathcal{O}(\alpha_{\text{fs}}^4). \quad (247)$$

Of these three contributions, the linear response contribution arising from the spin-orbit operator interacting with the orbital magnetic dipole operator is in most cases dominating, giving corrections to the g tensor of the form:

$$\Delta g_{\alpha\beta}^{\text{SO}} = \frac{2}{\langle S_Z \rangle} \left\langle \left\langle \hat{H}_{\alpha}^{\text{B}}, \hat{H}_{\beta}^{\text{SO}} \right\rangle \right\rangle, \quad (248)$$

where $\langle S_Z \rangle$ is the expectation value of the z component of the total electron spin. The relativistic mass correction Δg^{RMC} and gauge correction Δg^{GC} contributions are obtained as expectation values:

$$\Delta g_{\alpha\beta}^{\text{RMC}} = \frac{2}{\langle S_Z \rangle} \langle 0 | \hat{H}_{\alpha\beta}^{\text{SZ/KE}} | 0 \rangle, \quad (249)$$

$$\Delta g_{\alpha\beta}^{\text{GC}} = \frac{2}{\langle S_Z \rangle} \langle 0 | \hat{H}_{\alpha\beta}^{\text{GC}} | 0 \rangle. \quad (250)$$

Here, $\hat{H}^{\text{SZ/KE}}$ is the kinetic energy correction to the spin–Zeeman term given in the Breit–Pauli Hamiltonian (Eq. 10), whereas $\hat{H}^{\text{GC}} = \hat{H}^{\text{SO,B(1)}} + \hat{H}^{\text{SO,B(2)}}$ is the sum of the one- and two-electron magnetic-field corrections to the spin-orbit operator, Eqs. 59 and 60. Higher-order relativistic corrections have also been derived (Rinkevicius et al. 2008). We note, however, that the electron g tensor is a property for which a two- or four-component relativistic framework provides the most elegant expression, with the g tensor being reduced to a simple expectation value of the orbital magnetic dipole operator (see for instance van Lenthe et al. (1997)).

As for other open-shell properties, most calculations employ unrestricted approaches, in which conventional response-theory methodology can be used to calculate the relevant quantities if performed starting from a nonrelativistic wave function and using perturbation theory to include the relativistic corrections, as applied in Geurts et al. (1980), Noodleman and Baerends (1984). Spin-restricted approaches have also been applied, though it is important in this case to ensure that the triplet-perturbing operators acting on an open-shell reference state properly take all possible excitations into account (Rinkevicius et al. 2008). We also note that multireference CI and MCSCF wave functions have been used to calculate the g tensor of small molecules (see, e.g., Engström et al. (1998), Lushington and Grein (1996)).

The computed electronic g tensor will in general also be origin dependent in approximate calculations, since it depends on the external magnetic field. The origin dependence of the orbital magnetic dipole operator is canceled by the origin dependence of the gauge-correction term for exact states and wave functions. This means that in approximate calculations, care should be exercised to ensure gauge origin independence, for instance using London atomic orbitals (Ciofini et al. 2004). In general, however, electronic g tensors appear to be less sensitive than other properties to the choice of gauge origin (see for instance Kacprzak and Kaupp 2004), although exceptions have been observed.

Zero-Field Splitting

In degenerate perturbation theory, the ZFS of a spin-degenerate state ${}^{2S+1}\Psi_0$ is evaluated from the eigenvalues of the matrix

$$|\mathcal{H}_{ij}^{\text{ZFS}} - W\delta_{ij}| = 0, \quad (251)$$

where

$$\mathcal{H}_{ij}^{\text{ZFS}} = \frac{\langle 2S+1\Psi_0^i | \hat{H}^{\text{S,S}} | 2S+1\Psi_0^j \rangle - \sum_{n,\lambda} \sum_k \frac{\langle 2S+1\Psi_0^i | \hat{H}^{\text{SO}} | \lambda\Psi_n^k \rangle \langle \lambda\Psi_n^k | \hat{H}^{\text{SO}} | 2S+1\Psi_0^j \rangle}{\lambda E_n - E_0}}{1} \quad (252)$$

In these expressions, ${}^\lambda\Psi_n^k$ is the zeroth-order wave function, λ is the multiplicity of the state, S is the total spin quantum number, and the indices $i, j, k = M_S$ determine the projection of the total spin (Minaev et al. 2003; Vahtras et al. 2002). As the operator couples two electron spins, ZFS will only occur in molecules with at least two unpaired electrons.

As for many other properties, the zero-field splitting has two distinct contributions. The first contribution in Eq. 254 is an expectation value of the two-electron spin–spin dipole–dipole coupling operator,

$$\hat{H}^{\text{S,S}} = \frac{\alpha_{\text{fs}}^2}{2} \sum_{j \neq i} \left[\frac{\mathbf{m}_i \cdot \mathbf{m}_j}{r_{ij}^3} - \frac{3(\mathbf{m}_i \cdot \mathbf{r}_{ij})(\mathbf{r}_{ij} \cdot \mathbf{m}_j)}{r_{ij}^5} \right], \quad (253)$$

whereas the other term is the indirect interaction between the spin-magnetic moments of two electrons, mediated through the interaction between the orbital magnetic moments in the spin-orbit operator. For transition-metal complexes, the spin-orbit effects, often dominated by a few close-lying electronic states, can be included in a limited sum-over-states expansion (Bolvin 2006) or using ligand-field theory (Neese and Solomon 2003). The electron spin–spin interaction can, however, be dominant in molecules with weak spin-orbit interactions, such as triplet states of organic molecules (Vahtras et al. 2002).

We can represent the zero-field splitting as

$$\mathcal{H}^{\text{ZFS}} = \mathbf{S}^T \mathbf{D} \mathbf{S}, \quad (254)$$

and in a basis that diagonalizes \mathbf{D} , it is customary to write

$$\mathcal{H}^{\text{ZFS}} = D \left[S_z^2 - \frac{1}{3} S(S+1) \right] + E \left[S_x^2 - S_y^2 \right], \quad (255)$$

in which $D = D_{zz} - \frac{1}{2}(D_{xx} + D_{yy})$ and $E = \frac{1}{2}(D_{xx} - D_{yy})$, and there are thus only two independent parameters for the spin–spin coupling constants. There is in principle a contribution involving the total spin multiplied by the isotropic value of D , but since it does not contribute to the splitting of the electron spin levels, it does not affect the ESR spectrum.

There have not been too many studies, in particular using highly correlated wave functions, of the basis set and correlation requirements for the direct spin–spin

interaction to the zero-field splitting. For the spin-orbit contribution, the basis set requirements follow to a large extent those applicable for the spin-orbit correction to the g tensor. Otherwise, an accurate description of the relevant electronic excited states contributing to the spin-orbit contribution is required, which by itself may be a challenge in the case of transition-metal complexes (Bolvin 2006).

ESR Hyperfine Coupling Tensors

The ESR hyperfine splitting constants A_K appear in the ESR spin Hamiltonian Eq. 239 in the term \mathcal{H}^{hf} dependent on $\mathbf{S}^T \mathbf{A}_K \mathbf{I}_K$. This contribution couples the electron and nuclear magnetic moments, and thus provides very detailed information about the structure of radicals, in much the same manner as shielding constants do in NMR spectroscopy. To compute this correction, we thus need to select in the perturbation expansion all the terms bilinear in \mathbf{S} and \mathbf{m}_K .

Considering the expansion of the spin–Zeeman term in Eq. 40, when we insert the magnetic vector potential arising due to the nuclear magnetic moments, we find two contributions that are bilinear in the electron spin and nuclear magnetic moments, namely:

$$\mathbf{A}_K = \mathbf{A}_K^{\text{fc}} + \mathbf{A}_K^{\text{sd}}. \quad (256)$$

The purely isotropic contribution to the hyperfine interaction $\mathbf{A}_K^{\text{fc}} = A_K^{\text{fc}} \mathbf{1}$ and the purely anisotropic contribution \mathbf{A}_K^{sd} (Fernandez et al. 1997) in Eq. 256 are due to the Fermi contact (Eq. 51) and the spin–dipole (Eq. 52) operators, respectively. Both contributions can in principle be calculated as expectation values of the respective operators, using the spin-density matrix $\mathbf{D}^{\alpha-\beta}$ (i.e., the difference between the density matrices for α and β spin):

$$A_K^{\text{fc}} = -\frac{1}{2} g_K g_e \mu_B \mu_N \langle S_Z \rangle^{-1} \sum_{\mu\nu} \langle \phi_\mu | \hat{H}^{K(\text{FC})} | \phi_\nu \rangle D_{\mu\nu}^{\alpha-\beta}, \quad (257)$$

$$A_K^{\text{sd}} = -\frac{1}{2} g_K g_e \mu_B \mu_N \langle S_Z \rangle^{-1} \sum_{\mu\nu} \langle \phi_\mu | \hat{H}^{K(\text{SD})} | \phi_\nu \rangle D_{\mu\nu}^{\alpha-\beta}. \quad (258)$$

There are also contributions of the order α_{fs}^4 – a linear response contribution $\langle\langle \hat{H}^{\text{SO}}; \hat{H}^{K(\text{PSO})} \rangle\rangle$ due to relativistic effects caused by the spin-orbit operator and expectation values $\langle 0 | \hat{H}^{\text{SO},K(1)} | 0 \rangle$ and $\langle 0 | \hat{H}^{\text{SO},K(2)} | 0 \rangle$, Eqs. 62 and 63 (see Neese (2003), Arbuznikov et al. (2004)).

The ESR hyperfine coupling is determined by triplet perturbations. Thus, in principle one should use an unrestricted wave function to describe the reference state. However, it is also possible to use a spin-restricted wave function (Fernandez et al. 1992) and take into account the triplet nature of the perturbation in the definition of the response. Within such a (e.g., SCF or MCSCF) restricted-unrestricted approach,

first-order properties are given as the sum of the usual expectation value term and a response correction that takes into account the change of the wave function induced by the perturbation (of the type $\langle 0^{(1)} | H^0 | 0 \rangle$). This restricted-unrestricted approach has also been extended to restricted Kohn–Sham density functional theory (Rinkevicius et al. 2004).

In general, the most severe basis set requirements come from the isotropic hyperfine interactions, the anisotropic contribution in general being well converged with basis sets that perform well for the isotropic contribution. We mention in particular that there are basis sets specifically designed for calculations of hyperfine coupling constants (EPR-III) using density functional theory (Barone 1995). In general, for main-group elements, DFT does not display a particularly strong sensitivity on the choice of exchange–correlation functional, and provides results in quite good agreement with CCSD(T) (Perera et al. 1994b).

Even though the majority of calculations of hyperfine coupling constants are performed using unrestricted approaches, the importance of spin contamination and its effect on the hyperfine coupling constants remains somewhat unclear, in particular in DFT calculations. Whereas restricted and unrestricted approaches give more or less identical results for simple radicals, the situation is less clear in the case of transition metal compounds, and a few examples of the differences between restricted and unrestricted approaches are collected in Table 15, based on the results reported by Rinkevicius et al. (2004). We note that the differences between restricted and unrestricted approaches in general are small for the same functional, but that

Table 15 Calculated isotropic hyperfine coupling constants (in MHz) of transition metal compounds and their dependence on the exchange–correlation functionals; results taken from Rinkevicius et al. (2004)

Molecule	Isotope	BP86	BLYP	B3LYP	UBP86	UBLYP	UB3LYP	Expt.
TiO	⁴⁷ Ti	−245.0	−252.7	−248.6	−251.0	−257.5	−252.8	−241.0(60)
	¹⁷ O	−11.5	−11.3	−7.9	−8.2	−8.2	−4.9	n.a. ^a
VN	⁵¹ V	1321.8	1367.5	1317.4	1393.5	1432.6	1388.9	1311.8
	¹⁴ N	8.7	8.2	4.7	6.2	6.0	3.2	n.a.
VO	⁵¹ V	780.2	808.4	791.2	821.0	847.8	829.5	778.0(2)
	¹⁷ O	−6.9	−6.0	−2.5	−3.1	−2.7	1.1	0(4)
MnO	⁵⁵ Mn	507.8	527.9	508.6	526.8	543.1	521.8	479.9 (100)
	¹⁷ O	−5.6	−7.0	−9.4	−5.4	−6.6	−8.0	n.a.
MnH	⁵⁵ Mn	355.8	360.7	209.7	380.	380.0	331.8	279.4(12)
	¹ H	24.7	34.3	28.1	25.8	35.6	28.0	20.7 (39)
TiF ₃	⁴⁷ Ti	−127.5	−131.4	−158.2	−216.6	−218.0	−192.2	−184.4(4), −177.1(4)
	¹⁹ F	17.1	18.8	1.6	5.0	8.7	−5.6	8.3(4), 8.0(4)
MnO ₃	⁵⁵ Mn	1262.2	1267.6	1476.2	2009.3	2042.4	1735.5	1613(6)
	¹⁷ O	−3.9	−6.2	−6.4	−4.1	−5.1	2.6	n.a.

^aNot available

these differences become significant when spin contamination can be a problem, for instance for TiF_3 and MnO_3 .

The hyperfine coupling constants are usually reported in MHz; the contributions evaluated in a.u. need to be multiplied by $g_K g_e \mu_B \mu_N$.

Conclusions

We end this discussion of molecular properties with some additional comments related both to theory and to experiment. First of all, we have mainly discussed nonrelativistic methods, since these are still dominant in practical applications. The approach starting with a nonrelativistic description of the molecular system and introducing relativistic effects as a correction/perturbation has obviously some limitations and cannot be applied in highly accurate studies of molecules containing heavy elements. There are different approaches that start from a relativistic description of the unperturbed molecule (using one-, two-, or four-component wave functions), and we refer for a discussion of relativistic methods applied in the theory of molecular properties to Reiher and Wolf (2009) and Saue (2002). There is an increasing number of successful applications of these methods, for example, in the study of NMR shielding constants (Autschbach 2007; Kaupp 2004; Maldonado and Aucar 2009) and spin–spin coupling constants (Autschbach 2008).

Secondly, we have not discussed in any detail the effects of nuclear motion. Methods used to calculate these vibrational corrections, for both zero-point vibrational effects and temperature effects, have been described elsewhere in this book. There are, however, other effects that should also be considered. We have not discussed the role of the purely vibrational contributions to molecular (electric) properties (Bishop 1990), which in certain cases can be as large as the electronic contributions (Kirtman et al. 2000). Moreover, in conformationally flexible molecules one has to consider the effects of large nuclear motions. For instance, for a proper comparison with experiment, it may not be sufficient to perform an *ab initio* calculation for a single molecular structure. In experiment one will always observe the average value of the different thermally accessible isomers (rotamers, conformers), and in order to allow for a direct comparison with these experimental observations, a Boltzmann average of the properties of these isomers must be computed. This is particularly important when the properties of the isomers are very different, possibly even differing in sign (Pecul et al. 2004).

Finally, most experimental data are not obtained in the gas phase. Therefore, in order to have a faithful comparison with experiment, the effects of the solvent must be considered. There is an increasing number of studies of molecular properties in the presence of a solvent; we refer to a recent review for more details (Mennucci and Cammi 2007). Molecular dynamics or Monte Carlo procedures can be used when it is necessary to describe the nuclear dynamics in more detail, for example, for a better description of solvent effects or when considering sophisticated isomerization processes (Böhme et al. 2007; Møgelhøj et al. 2008, 2009).

The theory of molecular properties has for a long time provided a precise description of the interactions of a molecule with external and internal fields. In the last years, progress in the development and implementation of computational methods has enabled increasingly accurate calculations of such properties, often yielding results more accurate than experimental data. We note here that a molecular property is always defined as a quantity characterizing an isolated molecule, in a particular state, subject to particular perturbation(s) – conditions not easily achieved in experiment.

Theoretical calculations may thus provide insight into the electronic origins of the behavior of a molecule in the presence of electromagnetic fields, and help disentangle the contributions of solvents, vibrational effects, molecular rotation, and other small perturbations to the experimental observation. By combining the information obtainable in experiment with the tools of analysis offered by quantum-chemical calculations, a better understanding of the interactions between molecules and electromagnetic fields can be achieved, which in turn may facilitate for instance the design of molecules and functional materials with specific responses to external perturbations.

Acknowledgments We are grateful to Dr. Stephan Sauer and to Dr. Olav Vahtras for many helpful comments. We are indebted to Dr. Anna Kaczmarek-Kędziera for help in the editing of the manuscript.

Bibliography

- Abragam, A. (1961). *The principles of nuclear magnetic resonance*. Oxford: Oxford University Press.
- Anet, F. A. L., O'Leary, D. J., Wade, C. G., & Johnson, R. D. (1990). NMR relaxation by the antisymmetric component of the shielding tensor: A longer transverse than longitudinal relaxation time. *Chemical Physics Letters*, *171*, 401.
- Arbuznikov, A. V., Vaara, J., & Kaupp, M. (2004). Relativistic spin-orbit effects on hyperfine coupling tensors by density-functional theory. *Journal of Chemical Physics*, *120*, 2127.
- Aucar, G. A. (2008). Understanding NMR J-couplings by the theory of polarization propagators. *Concepts in Magnetic Resonance Part A*, *32*, 88.
- Auer, A. A., Gauss, J., & Stanton, J. F. (2003). Quantitative prediction of gas-phase ^{13}C nuclear magnetic shielding constants. *Journal of Chemical Physics*, *118*, 10407.
- Autschbach, J. (2007). Density functional theory applied to calculating optical and spectroscopic properties of metal complexes: NMR and optical activity. *Coordination Chemistry Reviews*, *251*, 1796.
- Autschbach, J. (2008). Two-component relativistic hybrid density functional computations of nuclear spin–spin coupling tensors using Slater-type basis sets and density-fitting techniques. *Journal of Chemical Physics*, *129*, 094105. Erratum: *130*, 209901 (2009).
- Bak, K. L., Gauss, J., Helgaker, T., Jørgensen, P., & Olsen, J. (2000). The accuracy of molecular dipole moments in standard electronic structure calculations. *Chemical Physics Letters*, *319*, 563.
- Baker, J., Buckingham, A. D., Fowler, P. W., Steiner, E., Lazzarotti, P., & Zanasi, R. (1989). The electrostatic model of field gradients at nuclei – an application to hydrogen-bonded complexes of HCl. *Journal of the Chemical Society, Faraday Transactions 2*, *85*, 901.

- Barone, V. (1995). Structure, magnetic properties and reactivities of open-shell species from density functional and self-consistent hybrid methods. In D. P. Chong (Ed.), *Recent advances in density functional methods* (Vol. 1, p. 287). Singapore: World Scientific.
- Barron, L. D. (2004). *Molecular light scattering and optical activity*. Cambridge: Cambridge University Press.
- Barron, L. D., & Vrbancich, J. (1984). Magneto-chiral birefringence and dichroism. *Molecular Physics*, 51, 715.
- Benedikt, U., Auer, A. A., & Jensen, F. (2008). Optimization of augmentation functions for correlated calculations of spin-spin coupling constants and related properties. *Journal of Chemical Physics*, 129, 064111.
- Benkova, Z., Sadlej, A. J., Oakes, R. E., & Bell, S. E. J. (2005). Reduced-size polarized basis sets for calculations of molecular electric properties. I. The basis set generation. *Journal of Computational Chemistry*, 26, 145.
- Bethe, H., & Salpeter, E. (1957). *Quantum mechanics of one- and two-electron atoms*. New York: Academic.
- Bishop, D. M. (1990). Molecular vibrational and rotational motion in static and dynamic electric fields. *Reviews of Modern Physics*, 62, 343.
- Bishop, D. M. (1995). Dispersion formula for the average first hyperpolarizability β . *Journal of Chemical Physics*, 95, 5489.
- Bishop, D. M., & Pipin, J. (1989). Improved dynamic hyperpolarizabilities and field-gradient polarizabilities for helium. *Journal of Chemical Physics*, 91, 3549.
- Bogaard, M. P., & Orr, B. J. (1975). Electric dipole polarizabilities of atoms and molecules. In A. D. Buckingham (Ed.), *International review of science, physical chemistry, molecular structure, and properties, series 2* (Vol. 2, p. 149). London: Butterworths.
- Böhm, M. C., Ramirez, R., & Schulte, J. (2007). On the influence of vibrational modes and intramolecular isomerization processes on the NMR parameters of bullvalene: A Feynman path integral – Ab initio investigation. *Chemical Physics*, 342, 1.
- Bolvin, H. (2006). An alternative approach to the g-matrix: Theory and applications. *ChemPhysChem*, 7, 1575.
- Bonin, K. D., & Kresin, V. V. (1997). *Electric-dipole polarizabilities of atoms, molecules and clusters*. Singapore: World Scientific.
- Born, M. (1918). Elektronentheorie des natürlichen optischen Drehungsvermögens isotroper und anisotroper Flüssigkeiten. *Annalen der Physik*, 55, 177.
- Buckingham, A. D. (1959). Direct method of measuring molecular quadrupole moments. *Journal of Chemical Physics*, 30, 1580.
- Buckingham, A. D. (1967). Permanent and induced molecular moments and long-range intermolecular forces. *Advances in Chemical Physics*, 12, 107.
- Buckingham, A. D., & Disch, R. L. (1963). The quadrupole moment of the carbon dioxide molecule. *Proceedings of the Royal Society A*, 273, 275.
- Buckingham, A. D., & Longuet-Higgins, H. C. (1968). The quadrupole moments of dipolar molecules. *Molecular Physics*, 14, 63.
- Buckingham, A. D., & Love, I. (1970). Theory of the anisotropy of nuclear spin coupling. *Journal of Magnetic Resonance*, 2, 338.
- Buckingham, A. D., & Pople, J. A. (1956). A theory of magnetic double refraction. *Proceedings of the Physical Society. Section B*, 69, 1133.
- Buckingham, A. D., & Stephens, P. J. (1966). Magnetic optical activity. *Annual Review of Physical Chemistry*, 17, 399.
- Buckingham, A. D., Pyykkö, P., Robert, J. B., & Wiesenfeld, L. (1982). Symmetry rules for the indirect nuclear spin-spin coupling tensor revisited. *Molecular Physics*, 46, 177.
- Champagne, B., Perpète, E. A., Jacquemin, D., van Gisbergen, S. J. A., Baerends, E. J., Soubra-Ghaoui, C., Robins, K. A., & Kirtman, B. (2000). Assessment of conventional density functional schemes for computing the dipole moment and (hyper)polarizabilities of push-pull π -conjugated systems. *Journal of Physical Chemistry A*, 104, 4755.

- Christiansen, O., Koch, H., & Jørgensen, P. (1995). The second-order approximate coupled cluster singles and doubles model CC2. *Chemical Physics Letters*, 243, 409.
- Christiansen, O., Jørgensen, P., & Hättig, C. (1998). Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy. *International Journal of Quantum Chemistry*, 68, 1.
- Christiansen, O., Coriani, S., Gauss, J., Hättig, C., Jørgensen, P., Pawłowski, F., & Rizzo, A. (2006). Accurate nonlinear optical properties for small molecules. In M. G. Papadopoulos, A. J. Sadlej, & J. Leszczynski (Eds.), *Non-linear optical properties of matter: From molecules to condensed phases* (p. 51). Dordrecht: Springer.
- Ciofini, I., Adamo, C., & Barone, V. (2004). Complete structural and magnetic characterization of biological radicals in solution by an integrated quantum mechanical approach: Glycyl radical as a case study. *Journal of Chemical Physics*, 121, 6710.
- Coriani, S., Hättig, C., & Rizzo, A. (1999a). The electric-field-gradient-induced birefringence of helium, neon, argon, and SF₆. *Journal of Chemical Physics*, 111, 7828.
- Coriani, S., Jørgensen, P., Rizzo, A., Ruud, K., & Olsen, J. (1999b). Ab initio determinations of magnetic circular dichroism. *Chemical Physics Letters*, 300, 61.
- Craig, D. P., & Thirunamachandran, T. (1984). *Molecular quantum electrodynamics. An introduction to radiation molecule interaction*. Mineola: Dover.
- Cronstrand, P., Luo, Y., Norman, P., & Agren, H. (2003). Ab initio calculations of three-photon absorption. *Chemical Physics Letters*, 375, 233.
- Cybalski, S. M., & Bishop, D. M. (1994). Calculations of magnetic properties. V. Electron-correlated hypermagnetizabilities (Cotton–Mouton effect) for H₂, N₂, HF, and CO. *Journal of Chemical Physics*, 101, 424.
- De Boni, L., Toro, C., & Hernández, F. E. (2008). Synchronized double L-scan technique for the simultaneous measurement of polarization-dependent two-photon absorption in chiral molecules. *Optics Letters*, 33, 2958.
- Drake, G. W. F. (2006). High precision calculations for helium. In G. W. F. Drake (Ed.), *Springer handbook of atomic, molecular and optical physics* (p. 199). New York: Springer.
- Dyall, K. G., & Faegri, K., Jr. (2007). *Introduction to relativistic quantum chemistry*. USA: Oxford University Press.
- Enevoldsen, T., Oddershede, J., & Sauer, S. P. A. (1998). Correlated calculations of indirect nuclear spin-spin coupling constants using second-order polarization propagator approximations: SOPPA and SOPPA(CCSD). *Theoretical Chemistry Accounts*, 100, 275.
- Engström, M., Minaev, B., Vahtras, O., & Agren, H. (1998). Linear response calculations of electronic g-factors and spin-rotational coupling constants for diatomic molecules with a triplet ground state. *Chemical Physics*, 237, 149.
- Faraday, M. (1846a). XLIX. Experimental researches in electricity. Nineteenth series. *Philosophical Magazine*, 28, 294.
- Faraday, M. (1846b). Experimental researches in electricity. Nineteenth series. *Philosophical Transactions of the Royal Society of London*, 136, 1.
- Fernandez, B., Jørgensen, P., Byberg, J., Olsen, J., Helgaker, T., & Jensen, H. J. A. (1992). Spin polarization in restricted electronic structure theory: Multiconfiguration self-consistent-field calculations of hyperfine coupling constants. *Journal of Chemical Physics*, 97, 3412.
- Fernandez, B., Christiansen, O., Jørgensen, P., Byberg, J., Gauss, J., & Ruud, K. (1997). Hyperfine and nuclear quadrupole coupling in chlorine and fluorine dioxides. *Journal of Chemical Physics*, 106, 1847.
- Flygare, W. H. (1974). Magnetic interactions in molecules and an analysis of molecular electronic charge distribution from magnetic parameters. *Chemical Reviews*, 74, 653.
- Fowler, P. W., & Buckingham, A. D. (1989). The magnetic hyperpolarizability anisotropy of some two-electron systems. *Molecular Physics*, 67, 681.
- Fowler, P. W., Lazzarotti, P., Steiner, E., & Zanasi, R. (1989). The theory of Sternheimer shielding in molecules in external fields. *Chemical Physics*, 133, 221.

- Fowler, P. W., Hunt, K. L. C., Kelly, H. M., & Sadlej, A. J. (1994). Multipole polarizabilities of the helium atom and collision-induced polarizabilities of pairs containing He or H atoms. *Journal of Chemical Physics*, *100*, 2932.
- Fukui, H., Baba, T., & Inomata, H. (1996). Calculation of nuclear magnetic shieldings. X. Relativistic effects. *Journal of Chemical Physics*, *105*, 3175. Erratum: *106*, 2987 (1997).
- Gauss, J., & Stanton, J. F. (1995). Coupled-cluster calculations of nuclear magnetic resonance chemical shifts. *Journal of Chemical Physics*, *103*, 3561.
- Gauss, J., & Stanton, J. F. (2002). Electron-correlated approaches for the calculation of NMR chemical shifts. *Advances in Chemical Physics*, *123*, 355.
- Gauss, J., Ruud, K., & Helgaker, T. (1996). Perturbation-dependent atomic orbitals for the calculation of spin-rotation constants and rotational g tensors. *Journal of Chemical Physics*, *105*, 2804.
- Gauss, J., Ruud, K., & Kállay, M. (2007). Gauge-origin independent calculation of magnetizabilities and rotational g tensors at the coupled-cluster level. *Journal of Chemical Physics*, *127*, 074101.
- Geurts, P. J. M., Bouten, P. C. P., & van der Avoird, A. (1980). Hartree–Fock–Slater–LCAO calculations on the Cu(II) bis(dithiocarbamate) complex; Magnetic coupling parameters and optical spectrum. *Journal of Chemical Physics*, *73*, 1306.
- Giorgio, E., Viglione, R., Zanasi, R., & Rosini, C. (2004). Ab initio calculation of optical rotatory dispersion (ORD) curves: A simple and reliable approach to the assignment of the molecular absolute configuration. *Journal of the American Chemical Society*, *126*, 12968.
- Hansen, A. E., & Bouman, T. D. (1980). Natural chiroptical spectroscopy: Theory and computations. *Advances in Chemical Physics*, *44*, 545.
- Harding, M. E., Lenhart, M., Auer, A. A., & Gauss, J. (2008a). Quantitative prediction of gas-phase ^{19}F nuclear magnetic shielding constants. *Journal of Chemical Physics*, *128*, 244111.
- Harding, M. E., Metzroth, T., Gauss, J., & Auer, A. A. (2008b). Parallel calculation of CCSD and CCSD(T) analytic first and second derivatives. *Journal of Chemical Theory and Computation*, *4*, 64.
- Harriman, J. E. (1978). *Theoretical foundations of electron spin resonance*. New York: Academic.
- Hättig, C., & Jørgensen, P. (1998). Dispersion coefficients for first hyperpolarizabilities using coupled cluster quadratic response theory. *Theoretical Chemistry Accounts*, *100*, 230.
- Hättig, C., Christiansen, O., & Jørgensen, P. (1998a). Multiphoton transition moments and absorption cross sections in coupled cluster response theory employing variational transition moment functionals. *Journal of Chemical Physics*, *108*, 8331.
- Hättig, C., Christiansen, O., & Jørgensen, P. (1998b). Coupled cluster response calculations of two-photon transition probability rate constants for helium, neon and argon. *Journal of Chemical Physics*, *108*, 8355.
- Helgaker, T., Ruud, K., Bak, K. L., Jørgensen, P., & Olsen, J. (1994). Vibrational Raman optical-activity calculations using London atomic orbitals. *Faraday Discussions*, *99*, 165.
- Helgaker, T., Jaszunski, M., & Ruud, K. (1997). Ab initio calculation of the NMR shielding and indirect spin–spin coupling constants of fluoroethylene. *Molecular Physics*, *91*, 881.
- Helgaker, T., Jaszunski, M., & Ruud, K. (1999). Ab initio methods for the calculation of NMR shielding and indirect spin–spin coupling constants. *Chemical Reviews*, *99*, 293.
- Helgaker, T., Jørgensen, P., & Olsen, J. (2000). *Molecular electronic-structure theory*. Chichester: Wiley.
- Helgaker, T., Jaszunski, M., & Pecul, M. (2008). The quantum-chemical calculation of NMR indirect spin-spin coupling constants. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *53*, 249.
- Hess, B. A., Marian, C. M., Wahlgren, U., & Gropen, O. (1996). A mean-field spin-orbit method applicable to correlated wavefunctions. *Chemical Physics Letters*, *251*, 365.
- Jackowski, K., Wilczek, M., Pecul, M., & Sadlej, J. (2000). Nuclear magnetic shielding and spin-spin coupling of 1,2- ^{13}C -enriched acetylene in gaseous mixtures with xenon and carbon dioxide. *Journal of Physical Chemistry A*, *104*, 5955. Erratum: *104*, 9806 (2000).

- Jackowski, K., Makulski, W., & Wasylishen, R. E. (2011). To be published.
- Jackson, D. A. (1998). *Classical electrodynamics* (3rd ed.). New York: Wiley.
- Jameson, C. J. (1987). Spin–spin coupling. In J. Mason (Ed.), *Multinuclear NMR* (p. 89). New York: Plenum Press.
- Jameson, C. J., Jameson, A. K., Oppunggu, D., Wille, S., Burrell, P. M., & Mason, J. (1981). ¹⁵N nuclear magnetic shielding scale from gas phase studies. *Journal of Chemical Physics*, *74*, 81.
- Jamieson, M. J. (1991). A time-dependent Hartree–Fock study of dispersion in the Cotton–Mouton effect for the helium isoelectronic sequence. *Chemical Physics Letters*, *183*, 9.
- Jansík, B., Rizzo, A., & Agren, H. (2005). Response theory calculations of two-photon circular dichroism. *Chemical Physics Letters*, *414*, 461.
- Jansík, B., Rizzo, A., & Agren, H. (2007). Ab initio study of the two-photon circular dichroism in chiral natural amino acids. *The Journal of Physical Chemistry B*, *111*, 446. Erratum: *111*, 2409 (2007).
- Jaszuński, M. (2004). Ab initio study of the shielding and spin–spin coupling constants in CIF₃, PF₃ and PF₅. *Chemical Physics Letters*, *385*, 122.
- Jaszuński, M., & Jackowski, K. (2008). Nuclear magnetic dipole moments from NMR spectra – quantum chemistry and experiment. In S. G. Karshenboim (Ed.), *Precision physics of simple atoms and molecules* (Lecture notes in physics, Vol. 745, p. 233). Berlin/Heidelberg: Springer.
- Jaszuński, M., & Ruud, K. (2001). Spin–spin coupling constants in C₂H₂. *Chemical Physics Letters*, *336*, 473.
- Jaszuński, M., Jørgensen, P., Rizzo, A., Ruud, K., & Helgaker, T. (1994). MCSCF calculations of Verdet constants. *Chemical Physics Letters*, *222*, 263.
- Jensen, F. (2006). The basis set convergence of spin–spin coupling constants calculated by density functional methods. *Journal of Chemical Theory and Computation*, *2*, 1360.
- Jensen, F. (2008). Basis set convergence of nuclear magnetic shielding constants calculated by density functional methods. *Journal of Chemical Theory and Computation*, *4*, 719.
- Kacprzak, S., & Kaupp, M. (2004). Electronic g-tensors of semiquinones in photosynthetic reaction centers. A density functional study. *The Journal of Physical Chemistry B*, *108*, 2464.
- Kállay, M., & Gauss, J. (2004). Analytic second derivatives for general coupled-cluster and configuration-interaction models. *Journal of Chemical Physics*, *120*, 6841.
- Kalugin, N. K., Kleindienst, P., & Wagnière, G. H. (1999). The magnetochiral birefringence in diamagnetic solutions and in uniaxial crystals. *Chemical Physics*, *248*, 105.
- Kaupp, M. (2004). Relativistic effects on NMR chemical shifts. In P. Schwerdtfeger (Ed.), *Relativistic electronic structure theory. Part 2. Applications* (p. 552). Amsterdam: Elsevier.
- Keal, T. W., & Tozer, D. J. (2003). The exchange–correlation potential in Kohn–Sham nuclear magnetic resonance shielding calculations. *Journal of Chemical Physics*, *119*, 3015.
- Kellö, V., & Sadlej, A. J. (1998). The quadrupole moment of the ³⁹K and ⁴¹K nuclei from microwave data for KF and KCl. *Chemical Physics Letters*, *292*, 403.
- Kendall, R. A., Dunning, T. H., Jr., & Harrison, R. J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *Journal of Chemical Physics*, *96*, 6796.
- Kerr, J. (1875a). XL. A new relation between electricity and light: Dielectric media birefringent. *Philosophical Magazine*, *50*, 337.
- Kerr, J. (1875b). LIV. A new relation between electricity and light: Dielectric media birefringent (second paper). *Philosophical Magazine*, *50*, 446.
- Kielich, S. (1972). General molecular theory and electric field effects in isotropic dielectrics. In M. Davies (Ed.), *Specialist periodical report, dielectric and related molecular processes* (Vol. 1, p. 192). London: Chemical Society.
- Kirpekar, S., Oddershede, J., & Jensen, H. J. A. (1995). Relativistic corrections to molecular dynamic dipole polarizabilities. *Journal of Chemical Physics*, *103*, 2983.
- Kirtman, B., Champagne, B., & Luis, J. M. (2000). Efficient treatment of the effect of vibrations on electrical, magnetic, and spectroscopic properties. *Journal of Computational Chemistry*, *21*, 1572.

- Kleinman, D. A. (1972). Nonlinear dielectric polarization in optical media. *Physical Review*, *126*, 1977.
- Koch, H., Christiansen, O., Jørgensen, P., Sánchez de Merás, A. M. J., & Helgaker, T. (1997). The CC3 model: An iterative coupled cluster approach including connected triples. *Journal of Chemical Physics*, *106*, 1808.
- Kosegi, S., Gordon, M. S., Schmidt, M. W., & Matsunaga, N. (1995). Main group effective nuclear charges for spin-orbit calculations. *Journal of Physical Chemistry*, *99*, 12764.
- Krivdin, L. B., & Contreras, R. H. (2007). Recent advances in theoretical calculations of indirect spin-spin coupling constants. In *Annual reports on NMR spectroscopy* (Vol. 61, p. 133).
- Kukulich, S. G. (1969). Measurement of the molecular g values in H₂O and D₂O and hyperfine structure in H₂O. *Journal of Chemical Physics*, *50*, 3751.
- Kussmann, J., & Ochsenfeld, C. (2007). Linear-scaling method for calculating nuclear magnetic resonance chemical shifts using gauge-including atomic orbitals within Hartree-Fock and density-functional theory. *Journal of Chemical Physics*, *127*, 054103.
- Kutzelnigg, W. (1980). Theory of magnetic-susceptibilities and NMR chemical shifts in terms of localized quantities. *Israel Journal of Chemistry*, *19*, 193.
- Łach, G., Jeziorski, B., & Szalewicz, K. (2004). Radiative corrections to the polarizability of helium. *Physical Review Letters*, *92*, 233001.
- Langevin, P. (1910). Sur les biréfringences électrique et magnétique. *Radium, Paris*, *7*, 249.
- Langhoff, P. W., & Karplus, M. (1970). Application of Padé approximants to dispersion force and optical polarizability computations. In G. A. Baker Jr. & J. L. Gammel (Eds.), *The Padé approximant in theoretical physics* (pp. 41–97). New York: Academic.
- Langhoff, P. W., Gordon, R. G., & Karplus, M. (1971). Comparisons of dispersion force bounding methods with applications to anisotropic interactions. *Journal of Chemical Physics*, *55*, 2126.
- Lazzeretti, P., & Zanasi, R. (1996). Molecular magnetic properties via formal annihilation of paramagnetic contribution to electronic current density. *International Journal of Quantum Chemistry*, *60*, 249.
- Lazzeretti, P., Malagoli, M., & Zanasi, R. (1994). Computational approach to molecular magnetic properties by continuous transformation of the origin of the current density. *Chemical Physics Letters*, *220*, 299.
- Löwdin, P.-O., & Goscinski, O. (1999). Studies in perturbation theory. XIV. Treatment of constants of motion, degeneracies and symmetry properties by means of multidimensional partitioning. *International Journal of Quantum Chemistry*, *5*, 685.
- Lushington, G. H., & Grein, F. (1996). The electronic g -tensor of MgF: A comparison of ROHF and MRD-CI level results. *International Journal of Quantum Chemistry*, *60*, 1679.
- Lutnæs, O. B., Teale, A. M., Helgaker, T., Tozer, D. J., Ruud, K., & Gauss, J. (2009). Benchmarking density-functional-theory calculations of rotational g tensors and magnetizabilities using accurate coupled-cluster calculations. *Journal of Chemical Physics*, *131*, 144104.
- Maldonado, A. F., & Aucaer, G. A. (2009). The UKB prescription and the heavy atom effects on the nuclear magnetic shielding of vicinal heavy atoms. *Physical Chemistry Chemical Physics*, *11*, 5615.
- Manninen, P., Lantto, P., Vaara, J., & Ruud, K. (2003). Perturbational ab initio calculations of relativistic contributions to nuclear magnetic resonance shielding tensors. *Journal of Chemical Physics*, *119*, 2623.
- Marchesan, D., Coriani, S., Forzato, C., Nitti, P., Pitacco, G., & Ruud, K. (2005). Optical rotation calculation of a highly flexible molecule: The case of paraconic acid. *Journal of Physical Chemistry A*, *109*, 1449.
- Mason, J. (1993). Conventions for the reporting of nuclear magnetic shielding (or shift) tensors suggested by participants in the NATO ARW on NMR shielding constants at the University of Maryland, College Park, July 1992. *Solid State Nuclear Magnetic Resonance*, *2*, 285.
- McClain, W. M. (1974). Two-photon molecular spectroscopy. *Accounts of Chemical Research*, *7*, 129.
- McWeeny, R. (1992). *Methods of molecular quantum mechanics* (2nd ed.). London: Academic.

- Mennucci, B., & Cammi, R. (2007). *Continuum solvation models in chemical physics: From theory to applications*. Chichester: Wiley.
- Michl, J., & Thulstrup, E. W. (1995). *Spectroscopy with polarized light*. Weinheim: VCH.
- Minaev, B., Loboda, O., Rinkevicius, Z., Vahtras, O., & Agren, H. (2003). Fine and hyperfine structure in three low-lying $^3\Sigma^+$ states of molecular hydrogen. *Molecular Physics*, *101*, 2335.
- Misquitta, A. J., Podeszwa, R., Jeziorski, B., & Szalewicz, K. (2005). Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional calculations. *Journal of Chemical Physics*, *123*, 214103.
- Møgelhøj, A., Aidas, K., Mikkelsen, K. V., & Kongsted, J. (2008). Solvent effects on the nitrogen NMR shielding and nuclear quadrupole coupling constants in 1-methyltriazoles. *Chemical Physics Letters*, *460*, 129.
- Møgelhøj, A., Aidas, K., Mikkelsen, K. V., Sauer, S. P. A., & Kongsted, J. (2009). Prediction of spin-spin coupling constants in solution based on combined density functional theory/molecular mechanics. *Journal of Chemical Physics*, *130*, 134508.
- Moss, R. E. (1973). *Advanced molecular quantum mechanics*. London: Chapman and Hall.
- Müller, T., Wiberg, K. B., & Vaccaro, P. H. (2000). Cavity ring-down polarimetry (CRDP): A new scheme for probing circular birefringence and circular dichroism in the gas phase. *Journal of Physical Chemistry A*, *104*, 5959.
- Neese, F. (2003). Metal and ligand hyperfine couplings in transition metal complexes: The effect of spin-orbit coupling as studied by coupled perturbed Kohn-Sham theory. *Journal of Chemical Physics*, *118*, 3939.
- Neese, F., & Solomon, E. I. (2003). Interpretation and calculation of spin-Hamiltonian parameters in transition metal complexes. In J. S. Miller & M. Drillon (Eds.), *Magnetism: Molecules to materials IV* (p. 345). Weinheim: Wiley.
- Nielsen, E. S., Jørgensen, P., & Oddershede, J. (1980). Transition moments and dynamic polarizabilities in a second order polarization propagator approach. *Journal of Chemical Physics*, *73*, 6238.
- Noodleman, L. (1981). Valence bond description of antiferromagnetic coupling in transition metal dimers. *Journal of Chemical Physics*, *74*, 5737.
- Noodleman, L., & Baerends, E. J. (1984). Electronic structure, magnetic properties, ESR, and optical spectra for 2-iron ferredoxin models by LCAO- $X\alpha$ valence bond theory. *Journal of the American Chemical Society*, *106*, 2316.
- Norman, P., Jiemchoorj, A., & Sernelius, B. E. (2003). Polarization propagator calculations of the polarizability tensor at imaginary frequencies and long-range interactions for the noble gases and n-alkanes. *Journal of Chemical Physics*, *118*, 9167.
- Ochsenfeld, C., Kussmann, J., & Koziol, F. (2004). Ab initio NMR spectra for molecular systems with a thousand and more atoms: A linear-scaling method. *Angewandte Chemie International Edition*, *43*, 4485.
- Oddershede, J., Geertsen, J., & Scuseria, G. E. (1988). Nuclear spin-spin coupling constant of hydrogen molecule with deuterium (HD). *Journal of Physical Chemistry*, *92*, 3056.
- Olsen, J., & Jørgensen, P. (1985). Linear and nonlinear response functions for an exact state and for an MCSCF state. *Journal of Chemical Physics*, *82*, 3235.
- Olsen, J., & Jørgensen, P. (1995). Time-dependent response theory with applications to self-consistent field and multiconfigurational self-consistent field wave functions. In D. R. Yarkony (Ed.), *Modern electronic structure theory* (p. 857). Singapore: World Scientific.
- Parkinson, W. A., & Oddershede, J. (1997). Response function analysis of magnetic optical rotation. *International Journal of Quantum Chemistry*, *64*, 599.
- Paterson, M. J., Christiansen, O., Pawłowski, F., Jørgensen, P., Hättig, C., Helgaker, T., & Sałek, P. (2006). Benchmarking two-photon absorption with CC3 quadratic response theory, and comparison with density-functional response theory. *Journal of Chemical Physics*, *124*, 054322.
- Pawłowski, F., Jørgensen, P., & Hättig, C. (2004). Gauge invariance of oscillator strengths in the approximate coupled cluster triples model CC3. *Chemical Physics Letters*, *389*, 413.

- Pecul, M., Ruud, K., Rizzo, A., & Helgaker, T. (2004). Conformational effects on the optical rotation of alanine and proline. *Journal of Physical Chemistry A*, *108*, 4269.
- Pedersen, T. B., Koch, H., Boman, L., & Sánchez de Merás, A. M. J. (2004). Origin invariant calculation of optical rotation without recourse to London orbitals. *Chemical Physics Letters*, *393*, 319.
- Perera, S. A., Sekino, H., & Bartlett, R. J. (1994a). Coupled-cluster calculations of indirect nuclear coupling constants: The importance of non-Fermi contact contributions. *Journal of Chemical Physics*, *101*, 2186.
- Perera, S. A., Watts, J. D., & Bartlett, R. J. (1994b). A theoretical study of hyperfine coupling constants. *Journal of Chemical Physics*, *100*, 1425.
- Pluta, T., & Sadlej, A. J. (1998). HyPol basis sets for high-level-correlated calculations of electric dipole hyperpolarizabilities. *Chemical Physics Letters*, *297*, 391.
- Podeszwa, R., & Szalewicz, K. (2008). Physical origins of interactions in dimers of polycyclic aromatic hydrocarbons. *Physical Chemistry Chemical Physics*, *10*, 2735.
- Raynes, W. T. (1992). Letter to editor. *Magnetic Resonance in Chemistry*, *30*, 686.
- Raynes, W. T. (1996). Electric field effects on shielding constants. In D. M. Grant & R. K. Harris (Eds.), *Encyclopaedia of NMR* (p. 1846). New York: Wiley.
- Raynes, W. T., McVay, R., & Wright, S. J. (1989). An improved ^{13}C nuclear shielding scale. *Journal of the Chemical Society, Faraday Transactions 2*, *85*, 759.
- Reiher, M., & Wolf, A. (2009). *Relativistic quantum chemistry*. Weinheim: Wiley.
- Rinkevicius, Z., Vaara, J., Telyatnyk, L., & Vahtras, O. (2003). Calculations of nuclear magnetic shielding in paramagnetic molecules. *Journal of Chemical Physics*, *118*, 2550.
- Rinkevicius, Z., Telyatnyk, L., Vahtras, O., & Agren, H. (2004). Density functional theory for hyperfine coupling constants with the restricted-unrestricted approach. *Journal of Chemical Physics*, *121*, 7614.
- Rinkevicius, Z., de Almeda, K. J., Oprea, C. I., Vahtras, O., Agren, H., & Ruud, K. (2008). Degenerate perturbation theory for electronic g tensors: Leading-order relativistic effects. *Journal of Chemical Theory and Computation*, *4*, 1810.
- Rizzo, A., & Coriani, S. (2005). Birefringences: A Challenge for both theory and experiment. *Advances in Quantum Chemistry*, *50*, 143.
- Rizzo, A., & Gauss, J. (2002). Shielding polarizabilities calculated at the coupled-cluster singles and doubles level augmented by a perturbative treatment of triple excitations. *Journal of Chemical Physics*, *116*, 869.
- Rizzo, C., Rizzo, A., & Bishop, D. M. (1997). The Cotton–Mouton effect in gases: Experiment and theory. *International Reviews in Physical Chemistry*, *16*, 81.
- Rizzo, A., Jansík, B., Pedersen, T. B., & Agren, H. (2006). Origin invariant approaches to the calculation of two-photon circular dichroism. *Journal of Chemical Physics*, *125*, 064113.
- Rizzo, A., Frediani, L., & Ruud, K. (2007). An ab initio investigation of the Buckingham birefringence of furan, thiophene, and selenophene in cyclohexane solution. *Journal of Chemical Physics*, *127*, 164321.
- Rudziński, A., Puchalski, M., & Pachucki, K. (2009). Relativistic, QED, and nuclear mass effects in the magnetic shielding of ^3He . *Journal of Chemical Physics*, *130*, 244102.
- Ruiz de Azúa, M. C., Melo, J. I., & Giribet, C. G. (2003). Orbital contributions to relativistic corrections of the NMR nuclear magnetic shielding tensor originated in scalar field-dependent operators. *Molecular Physics*, *101*, 3103.
- Ruud, K., & Helgaker, T. (1997). The magnetizability, rotational g tensor, and quadrupole moment of PF_3 revisited. *Chemical Physics Letters*, *264*, 17.
- Ruud, K., Helgaker, T., Bak, K. L., Jørgensen, P., & Olsen, J. (1995). Accurate magnetizabilities of the isoelectronic series BeH^- , BH , and CH^+ . The MCSCF-GIAO approach. *Chemical Physics*, *195*, 157.
- Ruud, K., Vaara, J., Lounila, J., & Helgaker, T. (1998). Vibrationally averaged magnetizabilities and rotational g tensors of the water molecule. *Chemical Physics Letters*, *297*, 467.

- Ruud, K., Taylor, P. R., & Jaszuński, M. (2000). Comment on "On the magnetic susceptibility of fluorine". *Journal of Physical Chemistry A*, *104*, 168.
- Ruud, K., Astrand, P.-O., & Taylor, P. R. (2001a). Zero-point vibrational effects on proton shieldings: Functional-group contributions from ab initio calculations. *Journal of the American Chemical Society*, *123*, 4826.
- Ruud, K., Taylor, P. R., & Astrand, P.-O. (2001b). Zero-point vibrational effects on optical rotation. *Chemical Physics Letters*, *337*, 217.
- Ruud, K., Stephens, P. J., Devlin, F. J., Taylor, P. R., Cheeseman, J. R., & Frisch, M. J. (2003). Coupled-cluster calculations of optical rotation. *Chemical Physics Letters*, *373*, 606.
- Sadlej, A. J. (1988). Medium-size polarized basis sets for high-level correlated calculations of molecular electric properties. *Collection of Czechoslovak Chemical Communications*, *53*, 1995.
- Sasagane, K., Aiga, F., & Itoh, R. (1993). Higher-order response theory based on the quasienergy derivatives: The derivation of the frequency-dependent polarizabilities and hyperpolarizabilities. *Journal of Chemical Physics*, *99*, 3738.
- Saue, T. (2002). Post Dirac-Fock-methods – Properties. In P. Schwerdtfeger (Ed.), *Relativistic electronic structure theory. Part I. Fundamentals* (p. 332). Amsterdam: Elsevier.
- Sauer, S. P. A. (1997). Second-order polarization propagator approximation with coupled-cluster singles and doubles amplitudes – SOPPA(CCSD): The polarizability and hyperpolarizability of Li^- . *Journal of Physics B: Atomic, Molecular and Optical Physics*, *30*, 3773.
- Sauer, S. P. A., Enevoldsen, T., & Oddershede, J. (1993). Paramagnetism of closed shell diatomic hydrides with six valence electrons. *Journal of Chemical Physics*, *98*, 9748.
- Sauer, S. P. A., Jensen, H. J. A., & Ogilvie, J. F. (2005). Quantum-chemical calculations of radial functions for rotational and vibrational g factors, electric dipolar moment and adiabatic corrections to the potential energy for analysis of spectra of HeH^+ . *Advances in Quantum Chemistry*, *48*, 319.
- Schatz, P. N., & McCaffery, A. J. (1969). Faraday effect. *Quarterly Review of the Chemical Society*, *23*, 552.
- Schindler, M., & Kutzelnigg, W. (1982). Theory of magnetic susceptibilities and NMR chemical shifts in terms of localized quantities. II. Application to some simple molecules. *Journal of Chemical Physics*, *76*, 1919.
- Seth, M., Ziegler, T., Banerjee, A., Autschbach, J., van Gisbergen, S. J. A., & Baerends, E. J. (2004). Calculation of the A term of magnetic circular dichroism based on time dependent-density functional theory I. Formulation and implementation. *Journal of Chemical Physics*, *120*, 10942.
- Shelton, D. P., & Rice, J. E. (1994). Measurements and calculations of the hyperpolarizabilities of atoms and small molecules in the gas phase. *Chemical Reviews*, *94*, 3.
- Smith, S. A., Palke, W. E., & Gerig, J. T. (1992). The Hamiltonians of NMR. Part I. *Concepts in Magnetic Resonance*, *4*, 107.
- Snyder, P. A., Atanasova, S., & Hansen, R. W. C. (2004). Ethylene. Experimental evidence for new assignments of electronic transitions in the $\pi \rightarrow \pi^*$ energy region. Absorption and magnetic circular dichroism measurements with synchrotron radiation. *Journal of Physical Chemistry A*, *108*, 4194.
- Solheim, H., Ruud, K., Coriani, S., & Norman, P. (2008). Complex polarization propagator calculations of magnetic circular dichroism spectra. *Journal of Chemical Physics*, *128*, 094103.
- Stanton, J. F., & Bartlett, R. J. (1993). The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *Journal of Chemical Physics*, *98*, 7029.
- Stanton, J. F., Gauss, J., & Christiansen, O. (2001). Equilibrium geometries of cyclic SiC_3 isomers. *Journal of Chemical Physics*, *114*, 2993.
- Stephens, P. J. (1976). Magnetic circular dichroism. *Advances in Chemical Physics*, *35*, 197.
- Stephens, P. J., Devlin, F. J., Cheeseman, J. R., & Frisch, M. J. (2001). Calculation of optical rotation using density functional theory. *Journal of Physical Chemistry A*, *105*, 5356.

- Sundholm, D., & Gauss, J. (1997). Isotope and temperature effects on nuclear magnetic shieldings and spin-rotation constants calculated at the coupled-cluster level. *Molecular Physics*, *92*, 1007.
- Tellgren, E. I., Soncini, A., & Helgaker, T. (2008). Nonperturbative ab initio calculations in strong magnetic fields using London orbitals. *Journal of Chemical Physics*, *129*, 154114.
- Thorvaldsen, A. J., Ruud, K., Rizzo, A., & Coriani, S. (2008). Analytical calculations of frequency-dependent hypermagnetizabilities and Cotton–Mouton constants using London atomic orbitals. *Journal of Chemical Physics*, *129*, 164110.
- Tinoco, I. (1975). Two-photon circular dichroism. *Journal of Chemical Physics*, *62*, 1006.
- Tozer, D. J., & Handy, N. C. (1998). Improving virtual Kohn–Sham orbitals and eigenvalues: Application to excitation energies and static polarizabilities. *Journal of Chemical Physics*, *109*, 10180.
- Vaara, J., & Pyykkö, P. (2003). Relativistic, nearly basis-set-limit nuclear magnetic shielding constants of the rare gases He–Rn: A way to absolute nuclear magnetic resonance shielding scales. *Journal of Chemical Physics*, *118*, 2973.
- Vaara, J., Ruud, K., Vahtras, O., Agren, H., & Jokisaari, J. (1998). Quadratic response calculations of the electronic spin-orbit contribution to nuclear shielding tensors. *Journal of Chemical Physics*, *109*, 1212.
- Vaara, J., Ruud, K., & Vahtras, O. (1999). Second- and third-order spin-orbit contributions to nuclear shielding tensors. *Journal of Chemical Physics*, *111*, 2900.
- Vaara, J., Jokisaari, J., Wasylishen, R. E., & Bryce, D. L. (2002). Spin–spin coupling tensors as determined by experiment and computational chemistry. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *41*, 233.
- Vahtras, O., Agren, H., Jørgensen, P., Jensen, H. J. A., Helgaker, T., & Olsen, J. (1992a). Multiconfigurational quadratic response functions for singlet and triplet perturbations: The phosphorescence lifetime of formaldehyde. *Journal of Chemical Physics*, *97*, 9178.
- Vahtras, O., Agren, H., Jørgensen, P., Jensen, H. J. A., Padkjær, S. B., & Helgaker, T. (1992b). Indirect nuclear spin–spin coupling constants from multiconfiguration linear response theory. *Journal of Chemical Physics*, *96*, 6120.
- Vahtras, O., Agren, H., Jørgensen, P., Jensen, H. J. A., Helgaker, T., & Olsen, J. (1992c). Spin–orbit coupling constants in a multiconfiguration linear response approach. *Journal of Chemical Physics*, *96*, 2118.
- Vahtras, O., Loboda, O., Minaev, B., Agren, H., & Ruud, K. (2002). Ab initio calculations of zero-field splitting parameters. *Chemical Physics*, *279*, 133.
- van Lenthe, E., Wormer, P. E. S., & van der Avoird, A. (1997). Density functional calculations of molecular *g*-tensors in the zero-order regular approximation for relativistic effects. *Journal of Chemical Physics*, *107*, 2488.
- Verhoeven, J., & Dymanus, A. (1970). Magnetic properties and molecular quadrupole tensor of the water molecule by beam-maser Zeeman spectroscopy. *Journal of Chemical Physics*, *52*, 3222.
- Visscher, L., & Saue, T. (2000). Approximate relativistic electronic structure methods based on the quaternion modified Dirac equation. *Journal of Chemical Physics*, *113*, 3996.
- Wasylishen, R. E., & Bryce, D. L. (2002). A revised experimental absolute magnetic shielding scale for oxygen. *Journal of Chemical Physics*, *117*, 10061.
- Wiberg, K. B., Wang, Y. G., Wilson, S. M., Vaccaro, P. H., & Cheeseman, J. R. (2006). Sum-over-states calculation of the specific rotations of some substituted oxiranes, chloropropionitrile, ethane, and norbornenone. *Journal of Physical Chemistry A*, *110*, 13995.
- Wigglesworth, R. D., Raynes, W. T., Kirpekar, S., Oddershede, J., & Sauer, S. P. A. (2000). Nuclear spin-spin coupling in the acetylene isotopomers calculated from ab initio correlated surfaces for 1J(C, H), 1J(C, C), 2J(C, H), and 3J(H, H). *Journal of Chemical Physics*, *112*, 3735. Erratum: *114*, 9192 (2001).
- Willeits, A., Rice, J., Burland, D. M., & Shelton, D. P. (1992). Problems in the comparison of theoretical and experimental hyperpolarizabilities. *Journal of Chemical Physics*, *97*, 7590.

- Woon, D. E., & Dunning, T. H., Jr. (1994). Gaussian basis sets for use in correlated molecular calculations. IV. Calculation of static electrical response properties. *Journal of Chemical Physics*, *100*, 2975.
- Yanai, T., Tew, D. P., & Handy, N. C. (2004). A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chemical Physics Letters*, *393*, 51.
- Zanasi, R., Lazzeretti, P., Malagoli, M., & Piccinini, F. (1995). Molecular magnetic properties within continuous transformations of origin of the current density. *Journal of Chemical Physics*, *102*, 7150.
- Zuber, G., & Hug, W. (2004). Rarefied basis sets for the calculation of optical tensors. 1. The importance of gradients on hydrogen atoms for the Raman scattering tensor. *Journal of Physical Chemistry A*, *108*, 2108.

Weak Intermolecular Interactions: A Supermolecular Approach

13

Mark Waller and Stefan Grimme

Contents

Introduction	594
Hydrogen-Bonded Complexes	596
Aromatic $\pi \dots \pi$ Stacking	596
Other Interaction Types	598
Interaction Energy	598
Methods	599
Coupled Cluster	599
Quantum Monte Carlo	600
Møller–Plesset Perturbation theory	600
Spin Component Scaled MP2	601
Hartree–Fock	602
Density Functional Theory	602
SemiEmpirical Methods	605
Molecular Mechanics	606
Basis Sets	607
Counterpoise Correction	608
Benchmark Sets	609
Performance Considerations	610
Concluding Remarks and Recommendations	613
Bibliography	614

Abstract

Weak intermolecular interactions, which are ubiquitous in biological and materials chemistry, are fast becoming more routinely and accurately investigated owing to the increased performance of computational methods being actively developed. A vast array of pragmatic methods have been proposed using

M. Waller (✉) • S. Grimme

Organisch-Chemisches Institut, Westfälische Wilhelms-Universität Münster, Münster, Germany
e-mail: m.waller@uni-muenster.de; grimmes@uni-muenster.de

empirical, semi-empirical, density functional theory, and ab initio approaches, which all serve to widen the scope of feasible problems. Especially for the calculation of the important London dispersion interactions, significant progress has been achieved. Herein, we present a general overview on a number of illustrative strategies used to routinely investigate structures and energies of such systems. The composition and advantages/disadvantages of different benchmark sets, which have been found to be of crucial importance in assessing such a wide range of methods is discussed. Finally, a number of experience-based perspectives are provided in relation to the scaling and accuracy of the “more popular” methods used when investigating non-covalent interactions.

A present trend in quantum chemistry is on cheap and reliable methods that effectively solve present-day problems in biological and materials chemistry. Quantum chemistry now confidently looks beyond small polyatomic molecules and toward large supramolecular complexes; this represents an area on the cutting edge of simulation sciences.

This chapter deals with weak intermolecular (non-covalent) interactions between molecules in the gas phase. These interactions are essential for the quantitative description and understanding of complex molecular aggregates in physics (e.g., surface science), chemistry, and molecular biology. The same interactions also occur in an intramolecular fashion between atoms or groups in one molecule. One of the big advantages of the supermolecular approach described herein is that it can handle both situations on an equal footing. Just for convenience and due to space limitations, we will consider here only intermolecular cases (complexes of at least two molecules). The reader should, however, keep in mind that much of what we are saying about quantum chemical methods similarly holds for the quantum chemical simulation of protein folding.

The following chapter is a pragmatic overview on “current” methods that are useful in obtaining reliable data from quantum chemical calculations, with a strong focus on methods used (and developed) primarily to study such non-covalent interactions. Weak intermolecular interactions in the solid or solution phase are almost completely neglected here, this is by no means a reflection on their importance, rather a way of restricting the scope of this chapter to a particular stream of research. A thorough description of the underlying theory of molecular interactions is presented in Volume I written by Alston Misquita. Only a succinct overview of weak intermolecular interactions is given below to “*set the scene.*”

Introduction

The individual stabilization of weak interactions is to a good approximation additive and is, therefore, extremely important. Weak molecular interactions are characterized by stabilization energies of typically less than $\Delta E < -30$ kJ/mol (Suresh and Naik 2000) per pair of interacting atoms (fragments). Therefore, in normal physiochemical environments, there is a continuous formation and breaking

of such molecular interactions. The resultant combined interaction energies can be responsible for some substantial effects on, e.g., molecular conformation and reactivity. However, one should keep in mind that due to additive property, complexes made out of about 100 atoms can have non-covalent (gas phase) interaction energies more resembling that of covalent bonds (i.e., 300–500 kJ/mol).

In chemistry and biochemistry, weak interactions are a hot topic of interest, especially within the emerging areas of molecular recognition (Desiraju and Steiner 1999), self-assembly (Hyla-Kryspin et al. 2004), supramolecular chemistry, and general host–guest interactions (Arunan and Gutowsky 1993; Desiraju and Steiner 1999; Hyla-Kryspin et al. 2004; Kim et al. 2000; Sharma et al. 2009; Sinnokrot et al. 2002; Steed et al. 1979; Tsuzuki and Luthi 2004; Tsuzuki et al. 1994). Weak interactions are fundamental in many aspects of biology with the quintessential example being the elegant double helix of Weak intermolecular interactionsDNA, resulting from a subtle interplay of weak interactions (Antony and Grimme 2008), see Fig. 1.

Furthermore, the secondary and tertiary structure of proteins, membrane structures, and complex intracellular particles such as ribosomes are all maintained by a variety of weak interactions (Wong 2009). Structure-based drug design is therefore dependent upon an understanding of the geometric and energetic aspects of these abundant interactions. Since pharmaceutical agents often bind to biological targets under the control of weak interactions, deconvolution of such host-receptor interactions into separate quantifiable interactions enables drug designers to make subtle modifications and then predict the consequences using modeling techniques.

Fig. 1 Deoxyribonucleic acid (DNA) highlighting the intrastrand interaction (*blue bracket*) between the two ball-and-stick base pairs, and the interstrand interaction (*red bracket*) which is between the lower ball-and-stick base and the adjacent base rendered as tubes

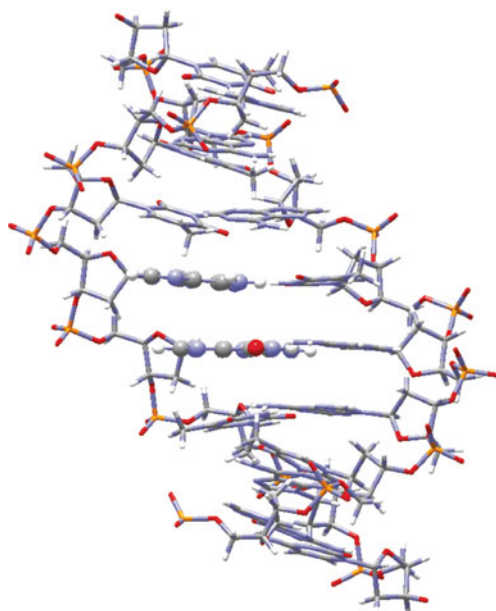
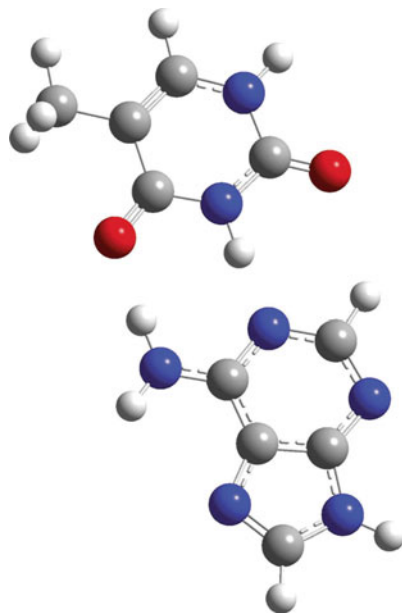


Fig. 2 The Watson and Crick (interstrand) adenine-thymine hydrogen-bonded base pair



Hydrogen-Bonded Complexes

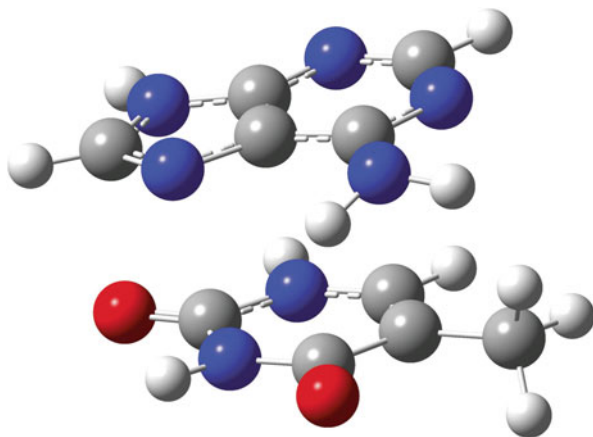
The electron distribution in covalent bonds is frequently polarized due to electronegativity differences of the atoms involved. The lone pair(s) on electron-rich atoms such as oxygen or nitrogen are attractive targets for terminal electron-poor atoms such as hydrogen on adjacent molecules. A hydrogen bond develops wherever a LP-bearing electronegative atom (acceptor) and a hydrogen atom that is bound covalently to another electronegative atom (donor) are sufficiently near, see Fig. 2.

Hydrogen bonds are typically strong if the atoms (donor, hydrogen, and acceptor) are collinear. The strength of their interactions reduces drastically for H-bonds which depart significantly from collinearity. The length of hydrogen bonds depends on the type of the constituent groups, but is typically 2.5–3.5 Å (between donor and acceptor nuclei) (Desiraju and Steiner 1999). Molecules may develop both inter- and intramolecular hydrogen bonds. So-called weak H-bonding, e.g., between C–H and fluorine atoms (with interactions in the range 3–10 kJ/mol) are a current topic of research (Hyla-Kryspin et al. 2004).

Aromatic $\pi \dots \pi$ Stacking

When aromatic rings are oriented favorably, they may develop a stabilizing interaction known as “ $\pi \dots \pi$ stacking” (i.e., involving the π electrons). The $\pi \dots \pi$ interaction is known to influence reactions, stabilize complexes, and influence structure. It follows that the estimation of the energetic and structural features

Fig. 3 The (intrastrand) adenine–thymine stacked base pair



of these interactions would be extremely useful in understanding many important chemical phenomena (Fig. 3).

A myriad of experimental and theoretical methods have been employed to investigate $\pi \dots \pi$ -stacking interactions. State-of-the-art electronic structure methods such as Møller–Plesset perturbation and coupled-cluster methods (see below) show that dispersive forces play the primary stabilizing role in π -stacked complexes (Arunan and Gutowsky 1993; Kim et al. 2000; Sinnokrot et al. 2002; Steed et al. 1979; Tsuzuki and Luthi 2004; Tsuzuki et al. 1994), while the electrostatic forces are responsible for details of the geometrical arrangements. Dispersion is a result of electron correlation, therefore methods that ignore electron correlation cannot be used at all and those that make drastic approximations to it must be carefully validated. The challenge of adequately describing the dispersion interaction is a major focus of the following chapter.

The non-covalent interactions in biological systems are of significant interest, e.g., in purine riboswitches binding sites (Sharma et al. 2009). The interactions of benzene and polycyclic aromatic hydrocarbons (PAHs) can be extended to real-world applications such as investigating the interaction of stacked graphene-nucleobase complexes (Antony and Grimme 2008). Concave buckyball catchers (Wong 2009) were investigated using sophisticated computational methods (e.g., DFT-D and SCS-MP2) explained later within this chapter. Fullerenes and carbon nanotubes (CNT) (Dappe et al. 2009), self-assembled of molecular tweezers formed by discrete biomolecular π -stacks (Lohr et al. 2009), and C_{60} -derived nanobaskets (dos Santos et al. 2009) were all investigated in 2009 using methods directly developed to model dispersion in an efficient manner. An interesting question was raised by Grimme, asking whether “special $\pi \dots \pi$ -stacking interactions really exist?” (Grimme 2008).

Other Interaction Types

The “weak” interaction energy discussed in detail below can be decomposed into various physically meaningful components, i.e., attractive electrostatic, polarization, and dispersion interactions and the repulsive Pauli exchange term. The sum of dispersion and Pauli repulsion is often called van der Waals interaction. In the various types of non-covalent interactions, these contributions are contained to a different degree. For example, there is a continuous transition from a purely dispersion bound van der Waals complex like the methane dimer to the ammonia dimer which already has partial H-bonding character to the water dimer as the H-bonding prototype complex. The typical H-bond has strong electrostatic and polarization contributions but only small (but non-negligible) dispersion forces. Pauli-repulsion is always present and mainly determined by the distance (wave function overlap) between the fragments. The benzene dimer on the other hand would be unbound when one could switch off the dispersion term. From the above, it is clear that only an accurate account of all these physically important contributions will lead to a consistent description of non-covalent interactions in general.

Interaction Energy

The interaction energy can be calculated with any of the methods outlined within this chapter, to a varying degree of accuracy. Adopting a “*Supramolecular approach*,” the interaction energy is defined as

$$\Delta E^{\text{int}} = E(AB) - E(A) - E(B), \quad (1)$$

where $E(A)$, $E(B)$, and $E(AB)$ are the electronic energies of the monomers A , B , and of the complex AB , respectively. This concept of interaction energy can be extended to a general multi-body regime:

$$\Delta E^{\text{int}} = E(A_1, A_2, A_3, \dots, A_N) - \sum_{i=1}^N E(A_i). \quad (2)$$

When investigating multi-body effects, the question of cooperative and anti-cooperative interactions can become an additional consideration of the utmost importance.

When one adopts a supermolecular approach using ab initio methods, one has to be mindful of the strengths, and limitations, of a given method:

- (a) Is a particular method size-consistent?
- (b) Does a particular method provide an adequate description of electron correlation?
- (c) Is the use of a particular finite basis set sufficient?

- (d) Does the method inherently have problems in the description of the electronic structure of *A* or *B*?

Within this chapter, we shall seek to answer such questions in a general overarching fashion.

An alternative to the supermolecular approach is symmetry adapted perturbation theory (SAPT) that computes the interaction energy directly (Chipman and Hirschfelder 1980; Jeziorski et al. 1978, 1980). The SAPT method has been shown to be suitable for studying weak intermolecular interactions and, in general, provides results similar to the current “gold standard” CCSD(T) (see below). Although this method has been instrumental to the understanding of weak intermolecular interactions, we have omitted details to rein in the scope of the chapter.

Methods

A general overview of the methods, in particular those developed to particularly address weak intermolecular interactions, will be given. The in-depth theoretical details of the underlying methods are given in other parts of this handbook and shall not be restated here in any depth. This should ensure that readers of different levels (and background) could gain some insight into how such calculations may be performed.

Ab initio methods are naturally more computationally expensive than empirical or semiempirical methods. However, they have the obvious advantage that no parameters need to be derived beforehand, which is often laborious and time consuming. The advantage of semiempirical methods are obviously speed, compared to their fully ab initio brethren. A discussion on describing weak intermolecular interactions would be obviously incomplete without an overview of the classical methods.

A strong disclaimer is that we cannot possibly cover all methods derived to compute weak intermolecular interactions, and the methods that we choose to cover shall inevitably date rather quickly. Neither of these considerations is avoidable within such a large and dynamic field of research, and this should hardly be taken as a negative.

Coupled Cluster

Coupled cluster with singles and doubles excitations (CCSD) is a size-consistent post-HF electron correlation method. The wavefunction, Ψ , in coupled cluster theory is formulated in terms of a cluster (exponential) expansion including the single and double excitation operators \hat{T}_1 and \hat{T}_2 . The effect of triple excitations (T) is calculated with perturbation theory.

CCSD(T) has become, and remains, the “gold-standard” for studying weak intermolecular interactions in the gas phase. The accuracy of this method in general

surpasses that of experimental determination and marks the limit of what is feasible using current computing resources.

Quantum Monte Carlo

Quantum Monte Carlo (QMC) effectively solves the many-body problem by a random walk through the electronic configuration space; it has been shown to be a promising method in quantum chemistry. One of the major advantages of QMC is the ability to perform massively parallel calculations, which can effectively increase the scope of what is computationally tractable by distributing the work over hundreds or even thousands of processors. QMC is a general method and, therefore, also has been applied recently to the computation of non-covalent interactions (e.g., the S22 data set) (Korth et al. 2008).

Møller–Plesset Perturbation theory

Møller–Plesset second-order perturbation theory (MP2) is a common method used in computational chemistry to include electron correlation as an extension to Hartree–Fock (HF) theory which neglects Coulomb correlation and thus also misses all dispersion effects. The perturbation is the difference between the Fock-operator and the exact electronic Hamiltonian.

In general, MP2 is a substantial improvement upon Hartree–Fock for all types of non-covalent interactions. The largest advantage of using second-order perturbation is, however, the inclusion of dispersion interactions. The method in general overestimates such an interaction; however, a number of strategies have been developed to address this problem. For example, Hobza et al. was the first to show that MP2/6-31G(0.25)* (i.e., with spatially expanded polarization functions) works well in a number of cases due to error compensation (Hobza and Šponer 1996).

Local (truncated) correlation methods, for example, LMP2, have the advantage of superior scaling behavior of the computation time with system size. In addition, they remove to large extent the basis set superposition error (explained below in section “CounterpoiseCorrection”). Such local methods have been applied to weak intermolecular interactions in combination with density fitting approximations (Goll et al. 2008; Hill and Platts 2008).

A further way to correct for the overestimation of MP2 for dispersion bound complexes is MP2.5 (Pitonák et al. 2009). This is a sum of MP2/CBS (complete basis set limit) interaction energies and a scaled third-order energy contribution obtained in small or medium size basis sets. MP2.5 results agree very closely with the estimated CCSD(T)/CBS interaction energies for weak intermolecular interactions. In particular, a very balanced treatment of hydrogen-bonded compared to stacked complexes is achieved with MP2.5. The main advantage of the approach is that it employs only a single empirical parameter and is thus based on two rigorously defined, asymptotically correct ab initio methods, MP2 and MP3.

The method is an accurate and computationally feasible alternative to CCSD(T) for the computation of the properties of various kinds of non-covalently bound systems.

Spin Component Scaled MP2

The total MP2 correlation energy is partitioned into parallel- and antiparallel-spin components that may be separately scaled. The two parameters p_s and p_t (scaling factors), whose values can be justified by basic theoretical arguments, were optimized on a thermochemical benchmark set (Grimme 2003).

$$E^{\text{corr}}(\text{SCS-MP2}) = p_s E_{\uparrow\downarrow} + p_t E_{\uparrow\uparrow+\downarrow\downarrow}. \quad (3)$$

It was shown, that the new method performs significantly better than standard MP2. Significant improvements are especially observed for cases which are usually known as MP2 pitfalls, while cases already described well with MP2 remain almost unchanged. Also, for difficult systems including strong (nondynamical) correlation effects, the improved MP2 method clearly outperforms DFT(B3LYP) and yields results of QCISD or sometimes QCISD(T) quality (QCISD is a variant of CCSD). The uniformity with which the new method improves upon MP2, thereby rectifying many of its problems and suggests it as a valuable quantum chemical method for the investigation of weak intermolecular interactions.

In an effort to simplify and reduce the scaling of the SCS-MP2 method, Head-Gordon has proposed a scaled opposite-spin (SOS) method. The SOS-MP2 method completely neglects the antiparallel-spin contribution to the MP2 energy and scales the parallel contribution by a factor of 1.3 (Jung et al. 2004). The Head-Gordon group has also developed a method, termed modified opposite-spin (MOS), in an attempt to provide a better description of long-range interactions (Lochan et al. 2005). The SCS(MI)-MP2 (MI: molecular interaction) was specifically designed to treat intermolecular interactions (Distasio and Head-Gordon 2007).

SCSN-MP2 is another parameterization of the spin scaling parameters which completely neglects the contribution from antiparallel-spin electron pairs to the MP2 energy while scaling the parallel contribution by 1.76 (Hill and Platts 2007). These spin-component scaled for nucleobases (SCSN) parameters were obtained by minimizing, with respect to SCS parameters, the RMS interaction energy error relative to the best available literature values, over a set of ten stacked nucleic acid base pairs. The applicability of this scaling to a wide variety of non-covalent interactions is verified through evaluation of a larger set of model complexes, including those dominated by dispersion and electrostatics.

One disadvantage of all SCS variants (except when the scaling factors add to a value of two) is that asymptotically (for large interfragment distances) the correct dispersion energy is not obtained. Although this is of less importance in many complexes, one should keep this in mind when very extended systems are considered.

Hartree–Fock

Hartree–Fock theory is a well-established method in computational chemistry. Therefore, no details on the method or implementation are given here. It suffices to say that Hartree–Fock theory is adequately accurate for hydrogen-bonding interactions (with, however, significant underbinding) in organic first-row compounds. Hartree–Fock, however, fails spectacularly for dispersion bound complexes. A dispersion corrected Hartree–Fock method (HFD) was proposed (Hepburn et al. 1975) and has seen some applications. Its main further disadvantage is the systematic overestimation of dipole moments and underestimation of the fragment polarizability. These two issues are much better described by DFT (at the same cost) so that Hartree–Fock can be considered as being more or less completely outdated.

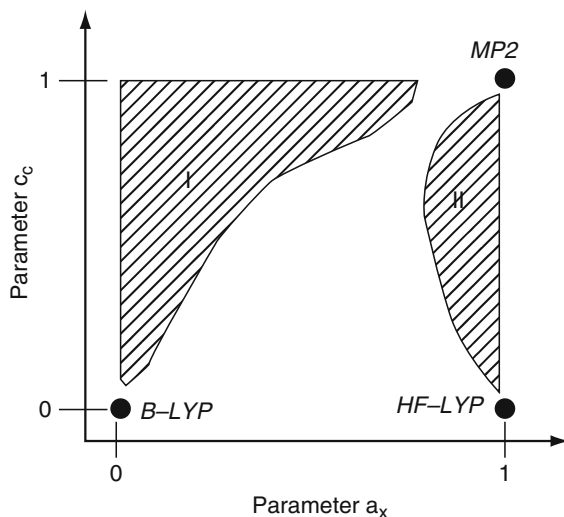
Density Functional Theory

Density functional theory (DFT) has gained huge popularity due to its ability to provide good accuracy, and due to its favorable scaling, in a timely fashion. Although it has been shown, and is commonly known, that an exact exchange correlation functional is possible, it remains ever elusive. To make matters worse, “traditional” (Kohn–Sham) DFT in present approximations does not provide a correct description of the dispersion interactions.

Density functional theory still remains “plagued” by an abundance of functionals developed with a plethora of design goals. Some grouping of functionals was proposed by Perdew and Schmidt (2001) using the “Jacob’s ladder” metaphor. In this metaphor, functionals may be grouped according to their formulation:

- First-rung functionals are only dependent upon the density and are known as local density approximation (LDA). Although LDA does overestimate the interaction energies in some dispersion bound systems (but for the wrong reasons), its accuracy can only be described as sporadic, and therefore its application in intermolecular interactions is NOT recommended.
- Second-rung functionals utilize in addition the gradient of the density, known as generalized gradient approximation (GGA). They typically show repulsive behavior for dispersion bound complexes similar to Hartree–Fock.
- Third-rung functionals further include the kinetic energy density or Laplacian (second derivative) of the density. These functionals show only marginal improvements over their second-rung counterparts.
- Fourth-rung functionals include Hartree–Fock exchange and are known as hybrid-GGAs. A number of hybrid functionals have been intentionally parameterized using datasets that include dispersion bound complexes, e.g., M05 (Zhao et al. 2005, 2006) and M06 (Zhao and Truhlar 2006a, b, 2008) (and variants thereof) from Truhlar and coworkers. The BHandH (Becke 1993) has been shown to give reasonable answers for particular problems of interest, but should be used with extreme caution (Swart et al. 2007; Waller et al. 2006).

Fig. 4 Schematic description of possible methods arising from different combinations of HF exchange and PT2 mixing parameters. The *dashed areas* I and II can be excluded by physical reasoning. B-LYP is a second-rung GGA and in HF-LYP the correlation energy is mixed with Hartree-Fock



- Fifth-rung functionals include the occupied and virtual orbitals simultaneously and are known as double hybrid functionals. They offer a superior description of long range-dispersion, while maintaining some quantum mechanical rigor. By their very nature, the current implementations are fundamentally semiempirical. A general *ansatz* for a combination of Kohn-sham DFT and perturbation theory is based on the following expression for the exchange-correlation energy E_{xc} :

$$E_{xc} = (1 - a_x) E_x^{\text{GGA}} + a_x E_x^{\text{HF}} + (1 - c_c) E_c^{\text{GGA}} + c_c E_c^{\text{PT2}}. \quad (4)$$

If $c_c = 0$ and $a_x \neq 0$, then a fourth-rung functional; If $c_c = 0$ and $a_x = 0$, then a second-rung functional is recovered, see Fig. 4.

The details of many “garden variety” density functionals are not discussed in this chapter, rather only recently developed schemes that specifically address shortcomings of density functionals in accurately describing weak intermolecular interactions.

DFT-D

An empirical method to account for dispersive interactions in practical calculations with the density functional theory (termed DFT-D) (Grimme 2006) has been tested for a wide variety of molecular complexes. As in previous schemes, the dispersive energy is described by a damped interatomic potential of the form $C_6 R^{-6}$. The use of pure (non-hybrid), gradient-corrected density functionals (e.g., BLYP or PBE), together with the resolution-of-the-identity (RI) approximation for the Coulomb integrals, allows for an efficient computation of large systems (as is often required for supramolecular chemistry).

By using a global scaling factor for the atomic C_6 coefficients, the functional dependence of the results could be strongly reduced. The double counting of correlation effects for strongly bound complexes is found to be insignificant if steep damping functions are employed. For stacked aromatic systems and the important base pairs, the DFT-D(BLYP) model seems to be even superior to standard MP2 treatments (which are known to systematically overbind). The good results obtained, in a variety of diverse examples, suggest that the DFT-D approach is a practical tool for describing weak intermolecular interactions. Furthermore, the DFT-D data may be used to calibrate much simpler (e.g., force-field) potentials. Alternatively, the optimized structures may be used as input for more accurate ab initio calculations of the interaction energies (Grimme 2006b).

The dispersive correction term in DFT-D is given by

$$E_{\text{disp}} = -s_6 \sum_{i=1}^{N_{\text{at}}-1} \sum_{j=i+1}^{N_{\text{at}}} \frac{C_6^{ij}}{R_{ij}^6} f_{\text{dmp}}(R_{ij}), \quad (5)$$

where N is the number of atoms in the system, C_6^{ij} denotes the dispersion coefficient for atom pair ij , s_6 is a global scaling factor (see below), and R_{ij}^6 is the interatomic distance. The damping function is

$$f_{\text{dmp}}(R) = \frac{1}{1 + e^{-\alpha(R/R_0-1)}}, \quad (6)$$

where R_0 is the sum of atomic respective van der Waals (cut-off) radii, see Fig. 5.

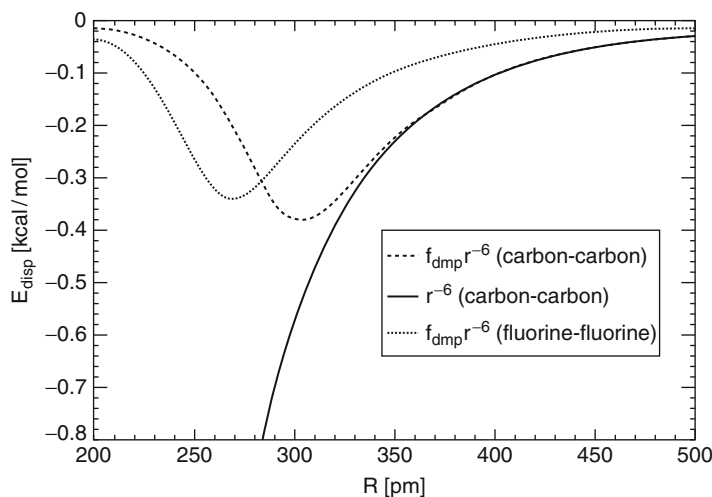


Fig. 5 Dispersion correction (Eq. 5) for two carbon and two fluorine atoms ($s_6 = 1.0$) separated by a distance R . The *solid line* shows the undamped potential for comparison

This simple approach has recently been improved regarding accuracy, less empiricism (the most important parameters R_0 and C_6 are computed *ab initio*), and general applicability to most elements of the periodic table (Grimme et al. 2010). An important change in this so-called DFT-D3 method is that the C_6 dispersion coefficients are dependent on the molecular structure which accounts for subtle effects, e.g., the hybridization state of an atom changes.

The DFT-D method does not only work for molecular complexes and intramolecular dispersion effects but is rather general, and Sauer and coworkers extended this correction to periodic systems (Kerber et al. 2008).

Range Separated and Dispersion Functionals

An alternative method is the nonlocal van der Waals density functional (vdW-DF) of Langreth and Lundqvist and coworkers (Andersson et al. 1996; Lundqvist et al. 1995). In vdW-DF, the nonlocal correlation is calculated explicitly in a non-empirical manner. It is typically used with standard GGAs like PBE or revPBE which introduces some inconsistency, as this results in an unbalanced treatment of computed interaction energies and equilibrium distances. The overall quality of vdW-DF results for typical non-covalent complexes is not better than with DFT-D but requires substantially more computation time.

The range-separated hybrid (RSH+MP2) (Ángyán et al. 2005; Gerber and Ángyán 2007) employs MP2 to account for the dispersion interaction, HF for long-range exchange, and, finally, LDA for short-range exchange and correlation. The method provides promising results for rare gas complexes but has not thoroughly been tested on standard benchmark sets (see section “[BenchmarkSets](#)”). The basic idea is similar to that of double-hybrid functionals.

Dispersion Calibrated Effective Potentials

A Dispersion-Corrected Atom-Centered Potential (DCACP) is a novel approach whereby the “missing” dispersion effect in traditional DFT is introduced using a suitably engineered effective core potential (ECP) (von Lilienfeld et al. 2004). The fitted atom–electron potentials in the DCACP method have been shown to perform well for weak intermolecular interactions (Arey et al. 2009; Tapavicza et al. 2007). Importantly, for practical purposes, the parameterized DCACPs are somewhat transferable to atoms in a wide range of different chemical environments (Lin and Rothlisberger 2008; Lin et al. 2007a, b; von Lilienfeld et al. 2005). This approach, however, has the same disadvantage as standard DFT-D in that the atomic parameters do not change with the electronic environment.

SemiEmpirical Methods

Traditionally, it has been generally believed that semiempirical methods are not particularly well suited to hydrogen bonding problems. In order to overcome these

limitations, the PM6 method has been re-parameterized and additional empirical terms were added to increase the accuracy of this semiempirical hamiltonian for describing weak intermolecular interactions (PM6-DH) (Řezáč et al. 2009) and the second generation corrections scheme PM6-DH2 (Korth et al. 2010). Such empirical corrections for the semiempirical methods brings about a new frontier in large simulations.

Another approach was proposed by Tuttle and Thiel (2008), starting with the OM x family of methods and adopting the dispersion correction of Grimme (2004). This particular combination was referred to as OM x -D. The authors found it necessary to rescale the empirical s_6 parameter (see above). Similarly, Hillier et al. (McNamara and Hillier 2007) proposed the AM1-D and PM3-D to alleviate some of the shortcomings of some more traditional semiempirical Hamiltonians.

The RM1_{BH} method (Feng et al. 2009) was constructed by adding Gaussian functions to the core–core repulsion items in the original AM1 formula. It was demonstrated that RM1_{BH} provides good agreement with the values of both high-level calculations and experiments for the binding energies of biological hydrogen-bonded systems.

Somewhat removed is the density functional based tight-binding method (DFTB), which is based on a second-order expansion of the Kohn-Sham total energy, employing a self-consistent redistribution of Mulliken charges (SCC-DFTB) (Elstner et al. 1998). It also employs dispersion corrections similar to the DFT-D method and has successfully been applied, e.g., to nuclear-base stacking problems (Elstner et al. 2001).

Molecular Mechanics

The empirically fitted force-fields such as CHARMM (Brooks et al. 1983), AMBER (Case et al. 2008), and GROMOS (Scott et al. 1999) perhaps perform better than one might originally imagine. Most importantly for the focus of this chapter are the noncovalent interactions, as naturally the interaction energy of a supramolecular complex should be dominated by the non-covalent terms. Therefore, a good description of such noncovalent interactions requires a balanced description of the electrostatic and van der Waals energies. Typically standard force-fields utilize the Coulomb's law for the electrostatic part and do not consider polarization. By judicious adjustment of the partial charges, one may strive to effectively model hydrogen-bonded interactions. Furthermore, one recognizes that the expression for the van der Waals interaction is normally represented as a Lennard-Jones potential. The r^{-12} term describing the Pauli repulsion and the r^{-6} term describing the long-range attractive dispersion forces. The vdW parameters may also be fitted to reproduce experimentally derived structures.

Basis Sets

Most *ab initio* and DFT schemes employ the LCAO method. The molecular orbitals are expanded using an orbital basis set – a collection of mathematical functions used to model the spatial variation of the orbitals. Historically, Slater-type orbitals were used due to their geometric similarity to the hydrogenic orbitals. They also remain a common choice in semiempirical MO theory. However, they do have a significant drawback in that there are no general and numerically stable solutions for the many-center integrals involving Slater-type orbitals. Consequently, most practitioners of quantum chemistry turned to Gaussian type orbitals.

Larger basis sets give rise to more accurate representations of the MOs. Practical considerations have dictated the need for optimization of basis sets to best describe molecular properties within a sensible computational time. This has led to the development of various basis sets each with differing characteristics reflecting their intended applications. The use of such optimized finite basis sets does produce reasonably accurate geometries (as the particular choice of a basis set does not strongly influence computed structures). However, energies are much more basis set dependent.

One may use two contracted GTOs per atomic orbital, as in the so-called double-zeta basis sets (Lochan et al. 2005), as this provides much greater flexibility. Split-valence basis sets partition the atomic orbitals into core and valence regions. The core AO's are assigned a minimal basis, while the valence orbitals are described at the double-zeta level.

The Pople style basis sets (Binkley et al. 1980; Dobbs and Hehre 1986; 1987a, b; Gordon et al. 1982; Hehre et al. 1969; Pietro et al. 1982) have received widespread use in the computational community. They may be described as: $X - YZg$ where X is the number of primitive Gaussians used to describe the core atomic orbitals. The valence orbitals are composed of a linear combination of Y and Z primitive Gaussian functions, respectively. Split-valence triple- and quadruple-zeta basis sets are also used, denoted as $X - YZWg$, $X - YZVWg$, etc.

The correlation consistent (cc) basis sets devised by Dunning (1989) were designed to converge systematically to the complete-basis-set (CBS) limit when used in conjunction with extrapolation techniques, see below. They may be described as: cc-pVNZ where $N = D, T, Q, 5, 6, \dots$ ($D =$ double zeta, $T =$ triple zeta, etc.).

Alternatively, the Ahlrichs style split-valence basis sets (Schaefer et al. 1992) are also ubiquitous in computational chemistry. Practically, extended AO basis sets of polarized TZV or QZV quality should be employed. Because the Ahlrichs basis sets have been carefully optimized variationally, the basis set superposition error (described below) is much smaller than with Pople or Dunning sets of about the same size. These expansions are consistently available for all elements of the periodic table.

Diffuse functions, denoted in Pople-type sets by a plus sign, +, and in Dunning-type sets by “aug” (abbreviation of “augmented”), are shallow Gaussian basis functions with large exponents. The diffuse functions are essential for the “tail” portion of the atomic orbitals, which are distant from the atomic nuclei. The diffuse basis functions are important when considering intermolecular complexes (a) to effectively span the intermolecular region and (b) to accurately describe the fragment polarizabilities. So-called polarization functions are also necessary for weak interactions to improve the flexibility of the wave function. These are functions with higher angular momentum than in the occupied ground state atomic orbitals.

The RI approximation is a method to approximate four-index two electron integrals without having any significant effect on the accuracy of results (Eichkorn et al. 1995). It requires specific auxiliary basis functions, which are available for many standard basis sets. The method is most efficient in GGA and MP2 calculations where it saves a factor of 10–20 in computation time for larger systems. It is also available in hybrid-GGA and Hartree–Fock treatments, but the savings are less (about a factor of 2–3). RI (also called density fitting) is strongly recommended in all treatments of larger complexes and is available in many major quantum chemistry codes.

Extrapolation schemes are employed to relieve the errors associated by using a truncated basis set. It is most often used to improve the computed correlation energy but also Hartree–Fock or DFT energies can be extrapolated. A number of schemes have been proposed and routinely used within the literature, a brief survey of some of the more popular schemes in chronological order: Martin (1996) and later Martin and Taylor (1997), Helgaker et al. (1998), Truhlar (1998), and Lee (2005) correction scheme. In the mostly employed standard scheme, two computations with basis sets of systematically improved quality (e.g., Dunning sets with $X = 2$, $Y = 3$ or $X = 3$, $Y = 4$) are required. The correlation energy is obtained by

$$E_{XY}^{\infty} = \frac{E_X^{\text{corr}} X^3 - E_Y^{\text{corr}} Y^3}{X^3 - Y^3}. \quad (7)$$

Basis set extrapolation also reduces considerably the basis set superposition error (see below).

Counterpoise Correction

The basis set superposition error (BSSE) arises as a direct result of basis set truncation. This truncation for a dimer or complex is less severe as it is for the monomer as the dimer basis set is the union of the two-monomer basis sets. This leads to an artificial energy lowering of the dimer compared to the monomers and leads to a too low (negative) interaction energy. The so-called *counterpoise correction* (CP) is the standard procedure for calculating an interaction energy using a supermolecular approach. The counterpoise corrected interaction energy is

$$\Delta E^{\text{CP}} = E_{AB}(AB) - E_{AB}(A) - E_{AB}(B), \quad (8)$$

where $E_{AB}(A)$ and $E_{AB}(B)$ denote the total energies of monomers A and B , respectively, computed with the dimer basis set of AB , i.e., in the calculation of monomer A including the basis set of the B monomer (but neglecting the nuclear charges and electrons of B). In this way, the basis set for each monomer becomes of about the same size as for the dimer thereby creating a more balanced description.

The counterpoise correction typically overestimates the BSSE since the monomer basis set is enhanced not only by empty orbitals of the other fragment, but also by orbitals occupied by electrons of the other monomer molecule which are excluded by the Pauli principle. Thus, if CP-corrected and uncorrected interaction energies are plotted as function of basis set size, they approach from above and below, respectively, the true interaction energy at the complete basis set (CBS) limit. CP corrections are mandatory for all double-zeta calculations and with MP2 or CCSD(T) also for triple-zeta basis treatments. In triple-zeta basis set (e.g., cc-pVTZ or TZVPP) DFT calculations, the BSSE is typically less than 5–10 % of the interaction energy which makes the laborious CP correction unnecessary. If sets of valence quadruple-zeta are used, it seems as if the error of the CP procedure is often similar to the (uncorrected) BSSE, but this is system-dependent and more definite conclusions about this issue requires further work.

Benchmark Sets

The importance of validating and assessing the relative strengths, and pitfalls, of a given theoretical method is often achieved using validation against standardized data sets. This circumvents the need to meta-analyze inhomogeneous studies (different basis sets, density functionals, etc.) of diverse molecular systems. Computational chemistry should ideally be predictive, and by having a clear and widely accepted set of references, one can preempt the magnitude of expected errors.

The development of force fields, semiempirical parameters, or even indeed density functionals, are reliant upon such carefully constructed datasets. What is needed is a good balance of diversity and relevance, when creating such a benchmark set. A benchmark set should ideally be able to expose systematic errors for different methods.

The BEDGB dataset (Řezáč et al. 2008) contains a number of subsets; S22 (Jurecčka et al. 2006) has been used extensively in benchmarking weak intermolecular interactions, see Fig. 6; S26 (Riley and Hobza 2007) augments four extra molecules into the S22 dataset; Small Halogen Bonding Complexes (Riley and Hobza 2008) contains six complexes calculated at the CCSD(T)/aug-cc-pVTZ CP level of theory; JSCH-2005 (Jurecčka et al. 2006) is a larger benchmark of non-covalent complexes; Set of five small peptides (Valdes et al. 2008) contains aromatic side chains and includes the P26 set; SCAI (Berka et al. 2009) – representative interactions of amino acid side chains.

The Minnesota non-covalent interaction databases were compiled by Truhlar et al. Benchmark Database of Noncovalent Interactions (2006): HB6/04 (Zhao et al. 2006; Zhao and Truhlar 2005a, b) is a hydrogen bond database that consists of

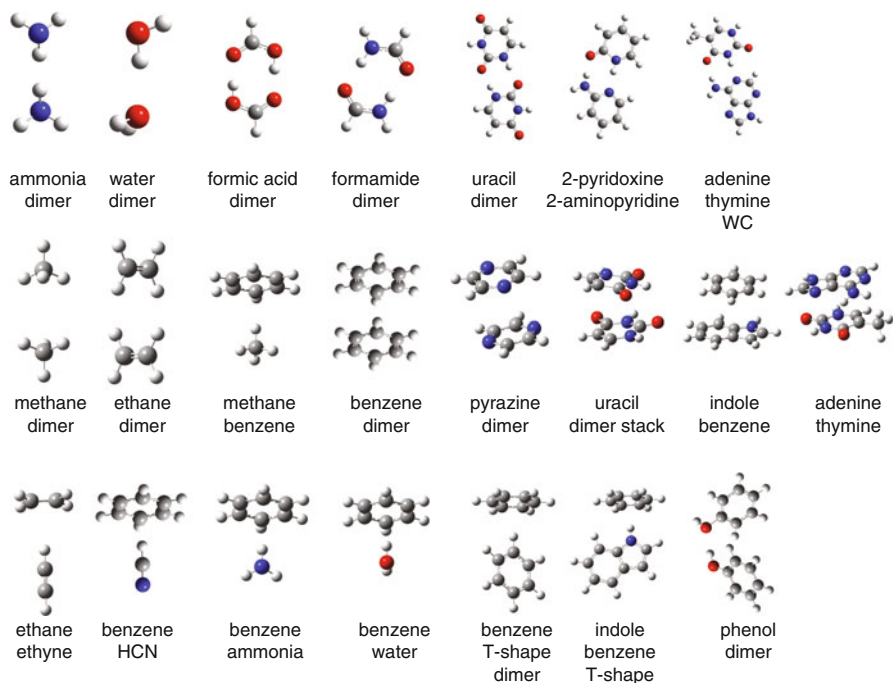


Fig. 6 The S22 dataset

the equilibrium binding energies of six hydrogen bonding dimers; DI6/04 (Zhao et al. 2006; Zhao and Truhlar 2005a, b,) database contains the binding energies of six dipole interaction complexes; WI7/05 (Zhao et al. 2005; 2006; Zhao and Truhlar 2005a) database consists of the binding energies of seven weak interaction complexes; PPS5/05 (Zhao et al. 2005; Zhao and Truhlar 2005a) database consists of binding energies of five $\pi \dots \pi$ -stacking complexes; CT7/04 (Zhao et al. 2005, 2006; Zhao and Truhlar 2005b) database consists of binding energies of seven charge transfer complexes.

When a collection of the different constituents of the Minnesota non-covalent interaction databases is made (HB6/04, DI6/04, WI7/05, PPS5/05, and CT7/04) the superset contains 31 complexes with many dimers in common with the S22 dataset. As these two databases are rather similar, for the remaining chapter, we will focus on the S22 dataset, see Fig. 5, to assess performance.

Performance Considerations

The methods were, more or less, presented above in order of decreasing accuracy. The higher-level ab initio methods are more accurate than the lower methods, as expected. The density functionals tend to roughly follow the Jacob's ladder

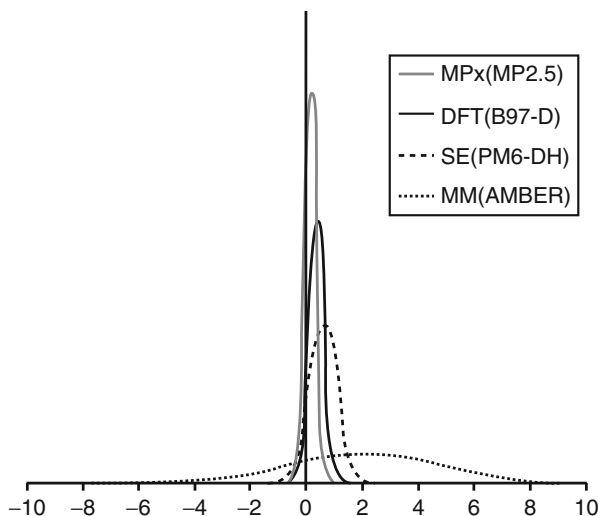


Fig. 7 The normal error distribution (in kcal/mol) for Møller-Plesset (MP), density functional theory (DFT), semiempirical (SE), molecular mechanics (MM)

analogy in terms of accuracy for the S22 dataset. Less accurate, for the S22, are the semiempirical methods with the empirical ones performing even worse. In order to provide a visual indication of the relevant methods Fig. 7 displays plots of the normal error distributions. One can immediately note that most methods are shifted to the right of the y-axis with CCSD(T) being the reference values. The higher-level methods again tend to be more narrow in their error distribution.

The dramatic gains in accuracy when one investigates weak intermolecular interactions at the DFT and MP2 levels of theory (with and without corrections) are displayed in Fig. 8.

Timing considerations (typically scaling) are important when one considers computational efficiency. Diminishing returns does occur with ab initio methods, and one must remain mindful of the desired accuracy and computational resources on hand. Such a delicate balancing act is crucial for practitioners of computational chemistry.

The computation time $t(cpu)$ required to execute a given quantum chemical method to obtain the energy follows a power law with respect to the number of electrons in the system, i.e.,

$$t(cpu) = a(N_{el})^b, \quad (9)$$

where a and b are constants which are characteristic for each method and its implementation. While b determines how fast the computation time increases with increasing system size (often abbreviated by the term “order n ”, i.e., an “order n ” method means $b = n$), the prefactor a measures how efficiently the given

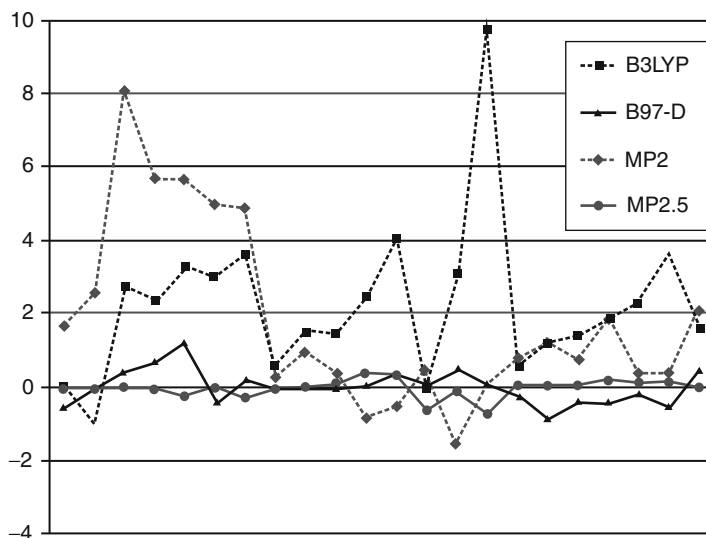


Fig. 8 The individual errors (in kcal/mol) for (a) DFT (b) DFT-D (c) MP2 (d) MP2.5 across the entries of the S22 dataset

Table 1 Computation times (in seconds on one Intel CoreDuo T9300 2.50 GHz CPU) for two non-covalently bound complexes with different quantum chemical methods

Method	Formal scaling exponent b	Water dimer ($N_{el} = 20$)	Benzene dimer ($N_{el} = 84$)
AM1	3	0.5	0.6
DFT(semi-local)	3	12.7	78.3
Hartree–Fock	4	5.6 (10.8)	278
RI-MP2	5	7.2 (12.5)	301
CCSD	6	11.7 (47.7)	17,284
CCSD(T)	7	14.1 (72.9)	25,810

Except for AM1, all calculations have been performed with a small cc-pVDZ (oraug-cc-pVDZ in parentheses) AO basis set and using the ORCA programpackage

combination of hard- and software performs the task. Although in general methods with small b are preferred, it is important to mention that to some extent a large value of b can be compensated by a small a so that formally “expensive” methods can sometimes be a better choice. Because the prefactor is very strongly influenced by many technical parameters, we present in Table 1 the computation times (single-point energy only) for two examples ordered according to formal scaling behavior.

It is seen (Table 1) that a realistic evaluation of the performance requires consideration of not too small systems and basis sets. The aug-cc-pVDZ results for the water dimer and the cc-pVDZ values for the benzene dimer, however, nicely illustrate the steep increase of the computation time with the level of sophistication. While RI-MP2 is not much more costly than HF (and even DFT), higher-level correlation methods (in particular CCSD(T)) already for medium-sized systems

lead to a substantial increase in computation time. The timings for the two systems also allow to check if the formal scaling laws hold. For example, for the water and benzene dimers, the system size increases by about a factor of four which translates for $b = 4$ to a computation time ratio of $4^4 = 256$ which is worse than the actually observed factor of about 50. The timings also show that already for a medium-sized system like the benzene dimer, CCSD(T) calculations with reasonable basis sets (at least aug-cc-pVDZ) are prohibitive in routine applications.

Concluding Remarks and Recommendations

The aim of developing a diverse set of well-balanced methods that can treat weak and strong inter- and intramolecular interactions continues to be a dominant area of computational chemistry. The ubiquitous nature of the weak intermolecular interactions in “state-of-the-art” materials and biomedical sciences ensures that work in this direction shall continue unabated.

The nature of fundamental intermolecular interactions were briefly introduced; a general impression of the relative strengths and weaknesses of different methods; basis sets; basis set extrapolation schemes; counterpoise correction schemes; and benchmark sets was presented. The chapter has concluded with overarching summary for the expected accuracy of the particular methods.

We finally want to make some general recommendations how “weak interactions” should be treated by quantum chemical methods:

1. Whenever possible, simple, approximate methods that are unavoidable in many practical applications should be checked or benchmarked against large (or complete) basis set CCSD(T) results. Often one can identify the main interactions and construct suitable model systems for this task. If the one-particle basis set issues and the basis set superposition error are carefully considered, CCSD(T) errors are $<5\%$ of the true interaction energies.
2. The simpler CCSD and MP3 *ab initio* methods are much less reliable (but still computationally expensive) and should not be used. In the “wave function world,” MP2 can (with some faults) also be recommended because the cost/performance ratio is rather good. MP2 should not be used for unsaturated or more strongly correlated (metallic) cases for which SCS-MP2 is a reliable alternative. If the systems are not too large (<50 atoms), MP2.5 can be used as well.
3. An interesting *ab initio* alternative when the system can be separated into chemically meaningful fragments is symmetry adapted perturbation theory (SAPT) that is complementary to the here-described supermolecular approach. Its main advantage is that it is free of BSSE and that it provides a natural partitioning of the interaction energy into physically meaningful components.
4. When the issue of London dispersion interactions is carefully considered with DFT, this yields a similar or often even better performance than MP2. Of the several approaches in this area, the DFT-D method has proven as an accurate

and robust computational tool. The main advantage of DFT-based approaches is that they can be applied also in electronically complicated situations (e.g., metal-containing systems) when MP n -based methods fail.

5. The biggest problem in DFT is the choice of the functional approximation. In many cases, computationally cheap (meta-)GGAs (e.g., BLYP, PBE, or TPSS) can be recommended. Such functionals should not be used when the self-interaction error (e.g., in charged open shell systems) plays a role. Then, hybrid functionals are required which also have less tendency for over-polarization. The currently highest level of approximation in DFT is represented by double-hybrid functionals (e.g., B2PLYP) that also perform very well for non-covalent interactions.
6. An important performance issue, that is unfortunately occasionally overlooked, is the tremendous computational savings gained by applying the resolution of the identity (RI) approximation when evaluating the two-electron integrals. This approximation, in combination with the development of well-performing “auxiliary” basis sets, has dramatically opened up new frontiers in large-scale calculations. We do not wish to comment on the accuracy of the various RI approximations, but we have experienced particularly good performance with the implementations in TURBOMOLE, ORCA, MOLPRO and Q-CHEM.
7. Semiempirical methods (and this also holds for many forcefields) mainly suffer from a poor description of the electrostatic interactions which, e.g., is demonstrated by their notoriously bad performance for hydrogen-bonded systems (for which even the simplest GGAs work very well). Because also their description of polarization (induction) contributions to binding is deficient, these methods can only be recommended for nonpolar systems or as parts in ONIOM-type approaches to describe the outer regions of a quantum chemical system.

Bibliography

- Andersson, Y., Langreth, D. C., & Lundqvist, B. I. (1996). van der Waals interactions in density-functional theory. *Physical Review Letters*, *76*, 102.
- AÁngyán, J. G., Gerber, I. C., Savin, A., & Toulouse, J. (2005). van der Waals forces in density functional theory: Perturbational long-range electron-interaction corrections. *Physical Review A*, *72*, 012510.
- Antony, J., & Grimme, S. (2008). Structures and interaction energies of stacked graphene–nucleobase complexes. *Physical Chemistry Chemical Physics*, *10*, 2722–2729.
- Arey, J. S., Aeberhard, P. C., Lin, I.-C., & Rothlisberger, U. (2009). Hydrogen bonding described using dispersion-corrected density functional theory. *The Journal of Physical Chemistry. B*, *113*, 4726–4732.
- Arunan, E., & Gutowsky, H. S. (1993). The rotational spectrum, structure and dynamics of a benzene dimer. *Journal of Chemical Physics*, *98*, 4294.
- Becke, D. (1993). A new mixing of Hartree–Fock and local density-functional theories. *Journal of Chemical Physics*, *98*, 1372–1377.

- Benchmark Database of Noncovalent Interactions. (2006). http://comp.chem.umn.edu/database_noncov/noncovalent.htm. Accessed 13 February 2010.
- Berka, K., Laskowski, R., Riley, K. E., Hobza, P., & Vondrasek, J. (2009). Representative amino acid side chain interactions in proteins. A comparison of highly accurate correlated ab initio quantum chemical and empirical potential procedures. *Journal of Chemical Theory and Computation*, 5, 982–992.
- Binkley, J. S., Pople, J. A., & Hehre, W. J. (1980). Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *Journal of the American Chemical Society*, 102, 939–947.
- Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, 3, 187–217.
- Case, D. A., Darden, T. A., Cheatham, T. E. I., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, M., Walker, R. C., Zhang, W., Merz, K. M., Jr., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., F. Wong, K., Paesani, F., Vanicek, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., & Kollman, P. A. (2008). *AMBER 10*. San Francisco: University of California.
- Chipman, D. M., & Hirschfelder, J. O. (1980). On symmetry in the polarization expansion for intermolecular forces. *Journal of Chemical Physics*, 73, 5164.
- Dappe, Y. J., Ortega, J., & Flores, F. (2009). Intermolecular interaction in density functional theory: Application to carbon nanotubes and fullerenes. *Physical Review B*, 79, 165409.
- Desiraju, G. R., & Steiner, T. (1999). *The weak hydrogen bond: In structural chemistry and biology* (IUCr monographs on crystallography, Vol. 9). Oxford: Oxford University Press/International Union of Crystallography Oxford.
- Distasio, R. A., Jr., & Head-Gordon, M. (2007). Optimized spin-component scaled second-order Møller-Plesset perturbation theory for intermolecular interaction energies. *Molecular Physics*, 105, 1073–1083.
- Dobbs, K. D., & Hehre, W. J. (1986). Molecular orbital theory of the properties of inorganic and organometallic compounds 4. Extended basis sets for third-and fourth-row, main-group elements. *Journal of Computational Chemistry*, 7, 359–378.
- Dobbs, K. D., & Hehre, W. J. (1987a). Molecular orbital theory of the properties of inorganic and organometallic compounds 5. Extended basis sets for first-row transition metals. *Journal of Computational Chemistry*, 8, 861.
- Dobbs, K. D., & Hehre, W. J. (1987b). Molecular orbital theory of the properties of inorganic and organometallic compounds. 6. Extended basis sets for second-row transition metals. *Journal of Computational Chemistry*, 8, 880–893.
- dos Santos, S. G., Pires, M. S., Lemos, V., Freire, V. N., Caetano, E. W. S., Galvão, D. S., Sato, F., & Albuquerque, E. L. (2009). C₆₀-derived nanobaskets: Stability, vibrational signatures, and molecular trapping. *Nanotechnology*, 20, 395701.
- Dunning, T. H., Jr. (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *Journal of Chemical Physics*, 90, 1007–1023.
- Eichkorn, K., Treutler, O., Öhm, H., Häser, M., & Ahlrichs, R. (1995). Auxiliary basis sets to approximate Coulomb potentials. *Chemical Physics Letters*, 240, 283–290.
- Elstner, M., Porezag, D., Jungnickel, G., Elsner, J., Haugk, M., Frauenheim, T., Suhai, S., & Seifert, G. (1998). Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Physical Review B*, 58, 7260.
- Elstner, M., Hobza, P., Frauenheim, T., Suhai, S., & Kaxiras, E. (2001). Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory based treatment. *Journal of Chemical Physics*, 114, 5149.
- Feng, F., Wang, H., Fang, W., & Yu, J. (2009). Can semiempirical quantum models calculate the binding energy of hydrogen bonding for biological systems? *Journal of Theoretical and Computational Chemistry*, 4, 691–711.

- Gerber, I. C., & Ángyán, J. G. (2007). London dispersion forces by range-separated hybrid density functional with second order perturbational corrections: The case of rare gas complexes. *Journal of Chemical Physics*, *126*, 044103.
- Goll, E., Leininger, T., Manby, F., Mitrushchenkov, A., Werner, H.-J., & Stoll, H. (2008). Local and density fitting approximations within the short-range/long-range hybrid scheme: Application to large non-bonded complexes. *Physical Chemistry Chemical Physics*, *10*, 3353.
- Gordon, M. S., Binkley, J. S., Pople, J. A., Pietro, W. J., & Hehre, W. J. (1982). Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements. *Journal of the American Chemical Society*, *104*, 2797–2803.
- Grimme, S. (2003). Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *Journal of Chemical Physics*, *118*, 9095.
- Grimme, S. (2004). Accurate description of van der Waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry*, *25*, 1463–1473.
- Grimme, S. (2006a). Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry*, *27*, 1787–1799.
- Grimme, S. (2006b). Semiempirical hybrid density functional with perturbative second-order correlation. *Journal of Chemical Physics*, *124*, 034108–034116.
- Grimme, S. (2008). Do special noncovalent π – π stacking interactions really exist? *Angewandte Chemie, International Edition*, *47*, 3430–3434.
- Grimme, S., Antony, J., Ehrlich, S., & Krieg, H. (2010). A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *Journal of Chemical Physics*, *132*, 154104.
- Halkier, A., Helgaker, T., Jørgensen, P., Klopper, W., Koch, H., Olsen, J., & Wilson, A. K. (1998). Basis-set convergence in correlated calculations on Ne. *Chemical Physics Letters*, *286*, 243–252.
- Hehre, W. J., Stewart, R. F., & Pople, J. A. (1969). Self-consistent molecular-orbital methods. I. use of Gaussian expansions of Slater-type atomic orbitals. *Journal of Chemical Physics*, *51*, 2657.
- Hepburn, J., Scoles, G., & Penco, R. (1975). A simple but reliable method for the prediction of intermolecular potentials. *Chemical Physics Letters*, *36*, 451.
- Hill, J. G., & Platts, J. A. (2007). Spin-component scaling methods for weak and stacking interactions. *Journal of Chemical Theory and Computation*, *3*, 80–85.
- Hill, J. G., & Platts, J. A. (2008). Calculating stacking interactions in nucleic acid base-pair steps using spin-component scaling and local second order Møller–Plesset perturbation theory. *Physical Chemistry Chemical Physics*, *10*, 2785–2791.
- Hobza, P., & Šponer, J. (1996). MP2 and CCSD(T) calculations on H-bonded and stacked formamide...formamide and formamidine...formamidine dimers. *Journal of Molecular Structure (THEOCHEM)*, *388*, 115–120.
- Hyla-Kryspin, I., Haufe, G., & Grimme, S. (2004). Weak hydrogen bridges: A systematic theoretical study on the nature and strength of C–H...F–C interactions. *Chemistry - A European Journal*, *10*, 3411–3422.
- Jeziorski, B., Chalasinski, G., & Szalewicz, K. (1978). Symmetry forcing and convergence properties of perturbation expansions for molecular interaction energies. *International Journal of Quantum Chemistry*, *14*, 271.
- Jeziorski, B., Schwalm, W. A., & Szalewicz, K. (1980). Analytic continuation in exchange perturbation theory. *Journal of Chemical Physics*, *73*, 6215.
- Jung, Y., Lochan, R. C., Dutoi, A. D., & Head-Gordon, M. (2004). Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method. *Journal of Chemical Physics*, *121*, 9793–9802.
- Jurečka, P., Šponer, J., Černý, J., & Hobza, P. (2006). Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Physical Chemistry Chemical Physics*, *8*, 1985–1993.

- Kerber, T., Sierka, M., & Sauer, J. (2008). Application of semiempirical long-range dispersion corrections to periodic systems in density functional theory. *Journal of Computational Chemistry*, *29*, 2088–2097.
- Kim, K. S., Tarakeshwar, P., & Lee, J. Y. (2000). Molecular clusters of π -systems: Theoretical studies of structures, spectra, and origin of interaction energies. *Chemical Reviews*, *100*, 4145.
- Korth, M., Lüchow, A., & Grimme, S. (2008). Toward the exact solution of the electronic Schrödinger equation for noncovalent molecular interactions: Worldwide distributed quantum Monte Carlo calculations. *The Journal of Physical Chemistry. A*, *112*, 2104–2109.
- Korth, M., Pitoňák, M., Řezáč, J., & Hobza, P. (2010). A transferable H-bonding correction for semiempirical quantum-chemical methods. *Journal of Chemical Theory and Computation*, *6*, 344–352.
- Lee, J. S. (2005). Accurate *ab initio* determination of binding energies for rare-gas dimers by basis set extrapolation. *Theoretical Chemistry Accounts*, *113*, 87–94.
- Lin, I.-C., & Rothlisberger, U. (2008). Describing weak interactions of biomolecules with dispersion-corrected density functional theory. *Physical Chemistry Chemical Physics*, *10*, 2730.
- Lin, I.-C., Coutinho-Neto, M. D., Felsenheimer, C., von Lilienfeld, O. A., Tavernelli, I., & Rothlisberger, U. (2007a). Library of dispersion-corrected atom-centered potentials for generalized gradient approximation functionals: Elements H, C, N, O, He, Ne, Ar, and Kr. *Physical Review B*, *75*, 205131.
- Lin, I.-C., von Lilienfeld, O. A., Coutinho-Neto, M. D., Tavernelli, I., & Rothlisberger, U. (2007b). Predicting noncovalent interactions between aromatic biomolecules with London-Dispersion-Corrected DFT. *The Journal of Physical Chemistry. B*, *111*, 14346.
- Lochan, R. C., Jung, Y., & Head-Gordon, M. (2005). Scaled opposite spin second order Møller-Plesset theory with improved physical description of long-range dispersion interactions. *The Journal of Physical Chemistry. A*, *109*, 7598–7605.
- Lohr, A., Grüne, M., & Würthner, F. (2009). Self-assembly of Bis(merocyanine) Tweezers into discrete bimolecular π -stacks. *Chemistry - A European Journal*, *15*, 3691–3705.
- Lundqvist, B. I., Andersson, Y., Shao, H., Chan, S., & Langreth, D. C. (1995). Density functional theory including van der Waals forces. *International Journal of Quantum Chemistry*, *56*, 247–255.
- Martin, J. M. L. (1996). *Ab initio* total atomization energies of small molecules – towards the basis set limit. *Chemical Physics Letters*, *259*, 669.
- Martin, J. M. L., & Taylor, P. R. (1997). Benchmark quality total atomization energies of small polyatomic molecules. *Journal of Chemical Physics*, *106*, 8620.
- McNamara, J. P., & Hillier, I. H. (2007). Semi-empirical molecular orbital methods including dispersion corrections for the accurate prediction of the full range of intermolecular interactions in biomolecules. *Physical Chemistry Chemical Physics*, *9*, 2362.
- Perdew, J. P., & Schmidt, K. (2001). Jacob's Ladder of density functional approximations for the exchange-correlation energy. In V. Van-Doren, C. V. Alsenoy, & P. Geerlings (Eds.), *Density functional theory and its applications to materials* (p. 1). New York: American Institute of Physics Press.
- Pietro, W. J., Francl, M. M., Hehre, W. J., Defrees, D. J., Pople, J. A., & Binkley, J. S. (1982). Self-consistent molecular orbital methods. 24. Supplemented small split-valence basis sets for second-row elements. *Journal of the American Chemical Society*, *104*, 5039–5048.
- Pitoňák, M., Neogrady, P., Cerný, J., Grimme, S., & Hobza, P. (2009). Scaled MP3 non-covalent interaction energies agree closely with accurate CCSD(T) benchmark data. *ChemPhysChem*, *10*, 282–289.
- RŘezáč, J., Jurečka, P., Riley, K. E., CČerný, J., Valdes, H., Pluháčková, K., Berka, K., RŘezáč, T., Pitoňák, M., Vondrášek, J., & Hobza, P. (2008). Quantum chemical benchmark energy and geometry database for molecular clusters and complex molecular systems (www.begdb.com): A users manual and examples. *Collection of Czechoslovak Chemical Communications*, *73*, 1261–1270. <http://www.begdb.com/>.

- Řezáč, J., Fanfrlik, J., Salahub, D., & Hobza, P. (2009). Semiempirical quantum chemical PM6 method augmented by dispersion and H-bonding correction terms reliably describes various types of noncovalent complexes. *Journal of Chemical Theory and Computation*, *5*, 1749–1760.
- Riley, K. E., & Hobza, P. (2007). Assessment of the MP2 method, along with several basis sets, for the computation of interaction energies of biologically relevant hydrogen bonded and dispersion bound complexes. *The Journal of Physical Chemistry. A*, *111*, 8257–8263.
- Riley, K., & Hobza, P. (2008). Investigations into the nature of halogen bonding including symmetry adapted perturbation theory analyses. *Journal of Chemical Theory and Computation*, *4*, 232–242.
- Schaefer, A., Horn, H., & Ahlrichs, R. (1992). Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *Journal of Chemical Physics*, *97*, 2571–2577.
- Scott, W. R. P., Hünenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., Torda, A. E., Huber, T., Krüger, P., & van Gunsteren, W. F. (1999). The GROMOS biomolecular simulation program package. *The Journal of Physical Chemistry. A*, *103*, 3596–3607.
- Sharma, P., Sharma, S., Chawla, M., & Mitra, A. (2009). Modeling the noncovalent interactions at the metabolite binding site in purine riboswitches. *Journal of Molecular Modeling*, *15*, 633–649.
- Sinnokrot, M. O., Valeev, E. F., & Sherrill, C. D. (2002). Estimates of the Ab Initio limit for π - π interactions: The benzene dimer. *Journal of the American Chemical Society*, *124*, 10887.
- Steed, J. M., Dixon, T. A., & Klemperer, W. (1979). Molecular beam studies of benzene dimer, hexafluorobenzene dimer, and benzene-hexafluorobenzene. *Journal of Chemical Physics*, *70*, 4940.
- Suresh, S. J., & Naik, V. M. (2000). Hydrogen bond thermodynamic properties of water from dielectric constant data. *Journal of Chemical Physics*, *113*, 9272.
- Swart, M., van der Wijst, T., Guerra, C. F., & Bickelhaupt, F. M. (2007). π - π stacking tackled with density functional theory. *Journal of Molecular Modeling*, *13*, 1245–1257.
- Tapavicza, E., Lin, I.-C., von Lilienfeld, O. A., Tavernelli, I., Coutinho-Neto, M. D., & Rothlisberger, U. (2007). Weakly bonded complexes of aliphatic and aromatic carbon compounds described with dispersion corrected density functional theory. *Journal of Chemical Theory and Computation*, *3*, 1673–1679.
- Truhlar, D. G. (1998). Basis-set extrapolation. *Chemical Physics Letters*, *294*, 45–48.
- Tsuzuki, S., & Luthi, H. P. (2004). Interaction energies of van der Waals and hydrogen bonded systems calculated using density functional theory: Assessing the PW91 model. *Journal of Chemical Physics*, *114*, 3949.
- Tsuzuki, S., Uchimaru, T., & Tanabe, K. (1994). Basis set effects on the intermolecular interaction of hydrocarbon molecules obtained by an *ab initio* molecular orbital method: Evaluation of dispersion energy. *Journal of Molecular Structure (THEOCHEM)*, *307*, 107.
- Tuttle, T., & Thiel, W. (2008). OM χ -D: Semiempirical methods with orthogonalization and dispersion corrections. Implementation and biochemical application. *Physical Chemistry Chemical Physics*, *10*, 2159.
- Valdes, H., Pluháčková, K., Pitoňák, M., Řezáč, J., & Hobza, P. (2008). Benchmark database on isolated small peptides containing an aromatic side chain: Comparison between wave function and density functional theory methods and empirical force field. *Physical Chemistry Chemical Physics*, *10*, 2747–2757.
- von Lilienfeld, O. A., Tavernelli, I., Röthlisberger, U., & Sebastiani, D. (2004). Optimization of effective atom centered potentials for London dispersion forces in density functional theory. *Physical Review Letters*, *93*, 153004.
- von Lilienfeld, O. A., Tavernelli, I., Rothlisberger, U., & Sebastiani, D. (2005). Performance of optimized atom-centered potentials for weakly bonded systems using density functional theory. *Physical Review B*, *71*, 195119.
- Waller, M. P., Robertazzi, A., Platts, J. A., Hibbs, D. E., & Williams, P. A. (2006). Hybrid density functional theory for π -stacking interactions: Application to benzenes, pyridines, and DNA bases. *Journal of Computational Chemistry*, *27*, 491–504.

- Wong, B. (2009). Noncovalent interactions in supramolecular complexes: A study on corannulene and the double concave buckycatcher. *Journal of Computational Chemistry*, *30*, 51–56.
- Zhao, Y., Schultz, N. E., & Truhlar, D. G. (2005). Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions. *Journal of Chemical Physics*, *123*, 161103.
- Zhao, Y., Schultz, N. E., & Truhlar, D. G. (2006). Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *Journal of Chemical Theory and Computation*, *2*, 364.
- Zhao, Y., & Truhlar, D. G. (2005a). Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions. *The Journal of Physical Chemistry. A*, *109*, 5656.
- Zhao, Y., & Truhlar, D. G. (2005b). Benchmark databases for nonbonded interactions and their use to test density functional theory. *Journal of Chemical Theory and Computation*, *1*, 415.
- Zhao, Y., & Truhlar, D. G. (2006a). A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *Journal of Chemical Physics*, *125*, 194101.
- Zhao, Y., & Truhlar, D. G. (2006b). Density functional for spectroscopy: No long-range self-interaction error, good performance for Rydberg and charge-transfer states, and better performance on average than B3LYP for ground states. *The Journal of Physical Chemistry. A*, *110*, 13126.
- Zhao, Y., & Truhlar, D. G. (2008). The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: Two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theoretical Chemistry Accounts*, *120*, 215–241.

John D. Watts

Contents

Introduction	622
General Aspects of Thermochemical Calculations	623
Calculating the ΔH , ΔS , and ΔG for a Chemical Reaction	623
Preliminary Comments on Accuracy	624
Examples	626
Case Study	626
Thermochemistry of Species with F–O Bonds	632
Calculating the Heat of Formation of H_2SO_2	634
Bibliography	637

Abstract

This chapter provides an introduction to the calculation of thermochemical data for chemical reactions using quantum chemical methods. The basic procedure is first described, namely, obtaining molecular structures and electronic energies of reactants and products, followed by vibrational frequency calculations and evaluation of thermal corrections. Since it is harder to obtain a given accuracy for some types of reactions than others, some discussion is provided on classes of reactions (e.g., isodesmic reactions) for which a given accuracy is easier to achieve than for a general reaction. Three examples illustrate different aspects of thermochemical calculations. The first example, the formation of ammonia from its elements, illustrates a variety of basis set and correlation effects on calculated data. The second example is concerned with calculations on small fluorine-oxygen species and a systematic side-by-side comparison of coupled-cluster and density-functional methods, including the use of isodesmic reactions.

J.D. Watts (✉)

Department of Chemistry and Biochemistry, Jackson State University, Jackson, MS, USA

e-mail: john.d.watts@jsums.edu

The third example describes the use of high-level coupled-cluster calculations to predict the standard enthalpy of formation of $S(OH)_2$.

Introduction

The term “chemical reactions” obviously covers a wide range of phenomena. The purpose of this chapter is to address some of the basic issues that one faces when using quantum chemical methods to calculate thermodynamic properties of chemical reactions. The focus here is more on the electronic problem than the nuclear problem, i.e., more on the potential energy surface than on motion thereon. The emphasis also is on gas-phase processes.

Fundamentally, the quantities to be calculated in thermochemical studies are the changes in enthalpy, entropy, and Gibbs free energy for a chemical reaction. From the standard Gibbs free energy change, the equilibrium constant may be obtained. At the most basic level, thermochemical data must be calculated for the reactants and products. This necessitates the calculation of minimum energy structures for all reactant and product species, followed by the calculation of harmonic vibrational frequencies, and thermal corrections. Very often one will seek to improve the results by high-level single-point energy calculations at the geometries obtained at lower-level methods, and using thermal corrections from a lower-level method. Another point is that the ultimate goal in a thermochemical study is often not the ΔH , ΔS , or ΔG of the reaction(s) directly studied. Rather, these quantities for one reaction are to be combined with data for another reaction, leading ultimately, for example, to an estimate of the standard heat of formation for a particular species. Sometimes experimental data may be incorporated in some way, and one often has some freedom in choosing which reactions to employ. The choices one makes can have a significant effect on the accuracy of the final answer, and hence the overall success of the project.

It is appropriate to mention a few monographs at this point. These provide a wide range of information on studying the thermochemistry of reactions by quantum chemical methods. The first of these is by Hehre et al. (1986). It summarizes a lot of studies using the theoretical-model chemistry approach pioneered by John A. Pople, as well as providing information on background theory and what have come to be called the Pople-type basis sets. Although this book is almost 25 years old, it is still a useful source. Hehre's book (Hehre 2003), which is provided as part of the Spartan software, is in some respects an update of (Hehre et al. 1986), providing new data, including data from various density-functional theory (DFT) methods. A different type of book is that by Helgaker et al. (2000). Most of it is devoted to a thorough and advanced presentation of the theory behind quantum chemical methods, but it also presents some numerical results on enthalpy changes using very large basis sets. A volume edited by Irikura and Frirup (1998) covers a range of issues in computational thermochemistry. Finally, a textbook by Lewars (2003) is very readable and has useful material on thermochemistry.

The plan of this chapter is as follows. The next section first addresses some general aspects of performing calculations and some points to be considered when

planning a project that will yield thermochemical data. Afterwards, it looks at some points concerning the accuracy of results obtained. The third and final section is split into a discussion of three types of studies that involve thermochemical calculations. The first examines in some detail the formation of (gaseous) ammonia from its elements. Various effects of basis set and method are illustrated. The second is concerned with obtaining accurate heats of formation on fluorine-oxygen species, especially using DFT methodology. The third is concerned with obtaining an accurate heat of formation of H_2SO_2 , for which experimental thermochemical data are lacking.

General Aspects of Thermochemical Calculations

Calculating the ΔH , ΔS , and ΔG for a Chemical Reaction

We consider a general gas-phase chemical reaction $aA(g) + bB(g) \rightarrow cC(g) + dD(g)$. In quantum chemical calculations, the enthalpy change for this reaction is calculated from the enthalpies of the species involved:

$$\Delta H^0 = cH_C + dH_D - [aH_A + bH_B].$$

This equation differs from the calculation of ΔH^0 from standard enthalpies of formation (ΔH_f^0), although it is equivalent to it. We return to the issue of obtaining ΔH_f^0 in theoretical methods later. At first sight, the notion of the enthalpy of a substance (H_A , H_B , H_C , and H_D) might be confusing: there is after all no such thing as an “absolute enthalpy.” The quantities H_A , H_B , H_C , and H_D are relative enthalpies, but they are enthalpies of substances *relative to the so-called quantum chemical standard state*. The quantum chemical standard state consists of infinitely separated electrons and nuclei of a substance. For example, the quantum chemical enthalpy of the water molecule is the enthalpy relative to an O nucleus, two protons, and ten electrons, all infinitely separated. The quantum chemical standard state is the zero of the potential energy in the usual quantum chemical Hamiltonian. To make a parallel with experimental thermochemistry, the ΔH_f^0 of a substance is its molar enthalpy relative to the enthalpy of its elements in their standard states.

Next we consider how the enthalpies are calculated. The quantum chemical enthalpy of a substance has electronic, nuclear, and PV contributions. The nuclear contribution consists of translational, rotational, and vibrational contributions. Hence:

$$H_A = E(A)_{\text{elec}} + E(A)_{\text{trans}} + E(A)_{\text{rot}} + E(A)_{\text{vib}} + PV.$$

The electronic contribution is the quantum chemical energy, that is, the energy of the stationary, nonrotating, non-vibrating system relative to the quantum chemical standard state; it is sometimes referred to as the bottom-of-the-well energy. The translational energy is the classical kinetic energy of a gas, and just depends on temperature. The rotational energy depends on the molecular structure, and the vibrational energy depends on the vibrational levels of the molecule. The

rotational and vibrational contributions depend on the temperature. In practice, the rotational and vibrational contributions are usually evaluated using the rigid-rotor and harmonic oscillator models. Equations for the various contributions are given in many places, for example, the textbook by McQuarrie and Simon (1999). A technical note on the Gaussian Web site by Ochterski (2000) is also recommended. Along with the equations, a sample output from the Gaussian software is explained. Calculating the enthalpy of a species thus requires a geometry optimization to obtain the molecular structure and a vibrational frequency calculation. The electronic energy is the energy at the optimized geometry.

The first step is to select a particular method and basis set that will be used to optimize the geometry and obtain vibrational frequencies. Several factors will be considered in this choice:

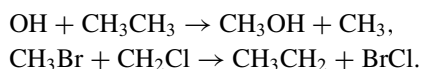
1. The method and basis set selected should give a satisfactory geometry and vibrational frequencies that are not *qualitatively* different from the true results.
2. The required computational resources should be available for the geometry optimization and frequency calculation. This is not a trivial matter: a geometry optimization involves several force calculations, and each force calculation typically requires 2–3 times the time required for an energy calculation. In addition to more processor time, further resources may be needed, for example, additional disk space. Even more resources are needed for frequency calculations. If vibrational frequencies are calculated by finite differences of gradients, calculations at lower symmetries than the molecule's symmetry are needed, and these require substantially more resources.
3. Ideally the method and basis set will be capable of providing reasonably accurate electronic energy differences, that is, $\Delta E_{\text{elec}} = cE(C)_{\text{elec}} + dE(D)_{\text{elec}} - [aE(A)_{\text{elec}} + bE(B)_{\text{elec}}]$. In practice, however, it is very common to use a higher-level method and/or basis set to provide an improved estimate for ΔE_{elec} by performing “single-point” energy calculations at the geometries obtained with a smaller basis set. ΔE_{elec} is the major contribution to ΔH , and it is more sensitive to quantum chemical method and basis set than are the other components of H . In conclusion, then, the method chosen for the geometry and frequency calculations is frequently not the method used for refined ΔE_{elec} calculations, especially in high-accuracy work. In fact, in high-accuracy work, it is also common to employ extrapolation techniques to estimate higher-level corrections.

Preliminary Comments on Accuracy

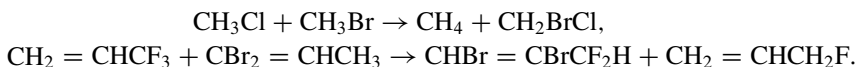
Naturally, one of the key questions concerns the level of theory needed to obtain a particular accuracy. There is no simple answer, but a few general comments are appropriate at this point (Hehre 2003; Hehre et al. 1986). The more similar are the reactants and products, the easier it should be to obtain accurate results. Similar in this sense means having similar chemical environments and bonding patterns in reactants and products. If the reactants and products are “similar” it

might be anticipated that the errors in the calculations on reactants and products are similar and to a significant degree cancel out. On the other hand, when there is less similarity, such cancelation is not expected, and higher-level calculations are likely to be needed for the same accuracy. For example, in the reaction $\text{N}_2 + 3\text{H}_2 \rightarrow 2\text{NH}_3$, the reactants and products are rather dissimilar: the reactants have an N–N triple bond and an H–H single bond, while the products have N–H single bonds. Another view is that the hybridization states of the N in reactants and products are different. Accurately calculating the ΔH^0 for the above “simple” reaction in fact is not at all an easy task. Both large basis sets and a high-level treatment of electron correlation are needed. Some numerical data to illustrate these points will be presented. It has long been known (Hehre et al. 1986) that it is especially difficult to obtain accurate results for reactions that have different numbers of unpaired electrons in reactants and products, for example, homolytic dissociation reactions.

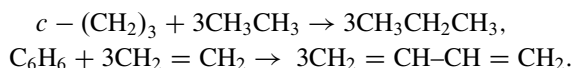
Several classes of reactions have been defined that display some degree of similarity in reactants and products. These include isogyric, isodesmic, and homodesmotic reactions. In isogyric reactions (Pople et al. 1983), the numbers of pairs of electrons in reactants and products is the same (and hence so is the number of unpaired electrons). For example, the following are isogyric processes:



Isogyric reactions constitute a very large group of reactions and it should not be assumed that it is easy to calculate the ΔH^0 of a general isogyric reaction. However, overall it tends to be easier to obtain an accurate ΔH^0 for an isogyric reaction than for a non-isogyric reaction. Any reaction that involves closed-shell reactants and products, such as the formation of ammonia from elements, is isogyric. Isodesmic and homodesmotic reactions are more restricted. In isodesmic reactions (Hehre et al. 1970), the number of bonds of each type and the number of lone pairs of each type are the same in reactants and products. For example:



Homodesmotic reactions (George et al. 1975) are a subset of isodesmic reactions. For organic compounds, the reactants and products have equal numbers of carbon atoms with the same hybridization. In addition, there are equal numbers of carbon atoms with each number of C–H bonds in reactants and products. Two examples, taken from a IUPAC document (Minkin 1999), are as follows:

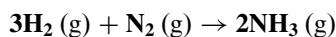


It is expected and has been amply verified (Hehre 2003; Hehre et al. 1986) that calculating ΔH for isodesmic and homodesmotic reactions is less demanding than for reactions in which reactants and products are less similar. Of course, this is of no apparent help if the target reaction is not one of these types. However, in a significant number of cases, it is possible to take advantage of (say) an isodesmic reaction in calculating a particular quantity of interest, such as a heat of formation.

A term commonly used in theoretical thermochemistry is “chemical accuracy.” This normally means an accuracy of 1 kcal mol^{-1} or about 4 kJ mol^{-1} . Achieving this accuracy is certainly not easy, and a lot of effort has been expended in trying to do so for reactions of relatively small molecules.

Examples

Case Study



In this section we discuss the calculation of the ΔH^0 of this reaction by different methods and basis sets. Although many other choices could have been made, this example illustrates many important points that are likely to apply to other cases. We begin by considering the calculation of the change in electronic energies, ΔE_{elec} . In fact, an experimental value of ΔE (the elec subscript is omitted hereafter in this section) can be derived by starting from the observed ΔH^0 and using known molecular structures and vibrational frequencies to obtain an “experimental” ΔE . Thus, we shall be assessing the ability of theoretical methods to calculate the experimental ΔE . Hehre et al. estimated this value to be $-37 \text{ kcal mol}^{-1}$ (Hehre et al. 1986, p. 283), but the later estimate of $-39.5 \text{ kcal mol}^{-1}$ (Helgaker et al. 2000, p. 866) by Helgaker et al. is probably more reliable. The ΔH^0 is $-22.1 \text{ kcal mol}^{-1}$.

Small Basis Set Study

We begin with results of calculations at geometries obtained with the Hartree–Fock (HF) method and the $6\text{-}31\text{G}^*$ basis set. In Table 1 we report ΔE values from the HF method and several correlated methods. The correlated methods used are second- and third-order perturbation theory (MP2 and MP3), fourth-order perturbation theory including single, double, and quadruple excitations (MP4SDQ), fourth-order perturbation theory including single, double, triple, and quadruple excitations (MP4), coupled-cluster singles-and-doubles (CCSD), and CCSD with noniterative treatment of triple excitations (CCSD(T)). Two well-known DFT treatments, namely, the BLYP and B3LYP functionals, are also used. Five basis sets are used. These are the $6\text{-}31\text{G}^*$ basis set and four extensions thereof, namely, $6\text{-}31\text{G}^*$, $6\text{-}31 + \text{G}^*$, $6\text{-}31 + \text{G}^*$, and $6\text{-}31++\text{G}^{**}$. The $6\text{-}31\text{G}^*$ basis set is a double-zeta valence basis set augmented with a set of d polarization functions on Li and heavier atoms. The $6\text{-}31 + \text{G}^*$ set adds a set of diffuse s and p functions to Li and heavier atoms to the $6\text{-}31\text{G}^*$ basis set. The $6\text{-}31\text{G}^{**}$ set includes a set of p polarization

Table 1 Calculated ΔE values for $3\text{H}_2(\text{g}) + \text{N}_2(\text{g}) \rightarrow 2\text{NH}_3(\text{g})$ (units are kcal mol^{-1})

	6-31G*	6-31G**	6-31 + G*	6-31 + G**	6-31++G**
HF	-27.8	-33.3	-32.2	-37.9	-38.1
MP2	-16.9	-28.1	-23.8	-34.9	-35.3
MP3	-23.7	-35.6	-30.3	-42.0	-42.4
MP4SDQ	-19.4	-30.6	-25.9	-36.9	-37.3
MP4	-13.8	-26.3	-20.7	-32.9	-33.4
CCSD	-19.5	-30.8	-26.0	-37.0	-37.4
CCSD(T)	-15.3	-27.9	-22.1	-34.4	-34.9
BLYP	-20.5	-27.8	-29.9	-37.4	-36.3
B3LYP	-29.2	-35.9	-37.0	-43.9	-43.3

functions on H atoms. The 6-31 + G* extends 6-31G** by adding a set of diffuse s and p functions to Li and heavier atoms. Finally, the 6-31++G** set has a diffuse s set on H atoms. References for these basis sets are given in (Hehre et al. 1986). The basis sets considered here are not expected to give very high accuracy. However, they are a convenient starting point and the results illustrate several important points. The results with the 6-31G* and 6-31 + G* basis sets can be discounted since these basis sets are not adequate for use with correlated methods for this example: in the calculations on H_2 these basis sets do not include any polarization functions. For the MP and CC methods, one can see a large change in ΔE on going from 6-31G* to 6-31G**. Likewise, there is a significant effect on going from 6-31 + G* to 6-31 + G**. Both show the importance of polarization functions on H.

Several features are shown by the data in Table 1. Looking at basis set effects first, we see that for all methods there is an appreciable effect on going from the 6-31G** to the 6-31 + G** basis set. This is the effect of adding diffuse functions to the N atom: it is $4.6 \text{ kcal mol}^{-1}$ for the HF method and slightly more for the correlated methods. Adding the H diffuse s functions, that is, going from 6-31 + G** to 6-31++G**, has a smaller effect. It decreases ΔE for the HF, MP, and CC methods, but increases it for BLYP and B3LYP. It should be borne in mind that effects of diffuse functions depend on the size of the rest of the basis set, but here they are definitely important. These results illustrate that diffuse functions can be important even when all species are neutral.

Several trends among the different correlated methods are noteworthy. The MP2 results are much closer to the MP4 or MP4SDQ results than to the MP3 results. In fact, the MP3 results seem to be somewhat exceptional. This underscores the widely held view that MP3 calculations rarely offer any benefits: they demand considerably more resources than MP2 calculations but do not usually improve MP2 results. This is certainly the case here. For all basis sets, the MP2 results are between the MP4SDQ and MP4 results. This is not to be expected in general, but is nonetheless noteworthy. There are close parallels between CCSD and MP4SDQ and between CCSD(T) and MP4. CC methods can be viewed as summing certain classes of terms in the perturbation series to infinite order. CCSD includes all of the

contributions in MP4SDQ to infinite order, along with certain other terms. There is a similar relationship between CCSD(T) and MP4. The closeness the CC and MP methods suggests relatively minor importance of infinite-order effects. Again, this is certainly not true in all cases, especially difficult cases, for which the CC methods are expected to perform better than the MP methods. The effect of connected triple excitations, that is, the difference between MP4SDQ and MP4 or between CCSD and CCSD(T), is about 3 kcal mol⁻¹, and so triple excitations are necessary for “chemical accuracy.”

It is premature to make judgments about accuracy based on the data in Table 1 since the basis sets are relatively small. A few noteworthy trends are apparent, however. The MP2, MP4, and CCSD(T) results seem to be getting closer to the experimental value as the basis set size increases. The HF result with the largest basis set is actually quite close to experiment, but this cannot generally be counted on. The two DFT sets of results show contrasting behavior. BLYP performs better for the larger basis sets, but this is not true for B3LYP for this example. It should be remembered that the effects of basis set and correlated methods are not additive. Differences between a pair of methods can depend on the basis set being used.

Study with Extended “Pople-Type” Basis Sets

We now examine the results when a set of larger Pople-type basis sets is used. These are based on a triple-zeta valence description. The effects of adding diffuse functions and using multiple sets of polarization effects are considered. The 6-311G** basis set is a triple-zeta valence basis set, augmented with a set of polarization functions for all atoms. 6-311++G** includes in addition a set of diffuse functions on all atoms. In the 6-311++G(2df,2pd) and 6-311++G(3df,3pd), multiple sets of polarizations are used on all atoms.

Comparing the 6-311G** results with the 6-31G** data in Table 1, one sees a change of about 1 kcal mol⁻¹, but not in the same direction for all methods. At the HF, MP2, CCSD, and CCSD(T) levels, the magnitude of ΔE decreases, while it increases for BLYP and B3LYP on going from 6-31G** to 6-311G**. Adding diffuse functions to 6-311G** again has a significant effect, 4 kcal mol⁻¹ for HF, about 5 kcal mol⁻¹ for MP2, CCSD, and CCSD(T), and about 6 kcal mol⁻¹ for BLYP and B3LYP. As in Table 1, diffuse functions make ΔE more negative. The effect of expanding the polarization space illustrates a significant difference between methods. For the HF method, the effect is comparatively small, and it decreases the magnitude of ΔE . For the conventional correlated methods, that is, MP2, CCSD, and CCSD(T), expanding the polarization space increases the magnitude of ΔE by about 7 kcal mol⁻¹, making it more negative. There is a comparatively small difference between the two largest polarization spaces. For BLYP and B3LYP there is little effect on ΔE on going beyond the 6-311++G** basis set. The comparative insensitivity of the DFT methods to polarization is rightly cited as an advantage of these methods over conventional wave function-based correlation methods. It is well known that the results of the latter converge more slowly with respect to basis set expansion.

Table 2 Calculated ΔE values for $3\text{H}_2(\text{g}) + \text{N}_2(\text{g}) \rightarrow 2\text{NH}_3(\text{g})$ using extended Pople-type basis sets (units are kcal mol^{-1})

	6-311G**	6-311++G**	6-311++G(2df,2pd)	6-311++G(3df,3pd)
HF	-32.4	-36.4	-35.1	-34.5
MP2	-27.3	-32.9	-40.7	-40.3
CCSD	-29.4	-34.4	-41.6	-41.2
CCSD(T)	-27.0	-32.2	-40.2	-39.9
BLYP	-28.6	-35.5	-36.5	-36.3
B3LYP	-36.3	-42.3	-42.9	-42.6

Table 3 Effect of geometries on ΔE

HF	MP2	CCSD	CCSD(T)	BLYP	B3LYP
-34.5	-40.3	-41.2	-39.9	-36.3	-42.6
-34.2	-38.6	-40.7	-39.1	-35.9	-42.4

The basis set is 6-311++G(3df,3pd). The first row data are for the HF/6-31G* geometry. The second row data are for the optimized geometries of the different methods with the 6-311++G(3df,3pd) basis set. Units are kcal mol^{-1}

One difference between Tables 1 and 2 is that the effect of triple excitations is slightly smaller with the larger basis sets used in Table 2. As observed in Table 1, the MP2 method seems to perform very well since its results are between those of CCSD and CCSD(T). Concerning agreement with experiment ($-39.5 \text{ kcal mol}^{-1}$ (Helgaker et al. 2000)), results with the largest basis set show best agreement for CCSD(T). Of course, the effects of several additional factors need to be assessed. These include further basis set extensions and the effect of using the actual minimum energy geometries for the respective methods.

Effect of Geometry

Geometries of all species were calculated using the 6-311++G(3df,3pd) basis set for six methods (HF, MP2, CCSD, CCSD(T), BLYP, and B3LYP), and the ΔE values were calculated. The results are shown in Table 3 below, along with the ΔE values obtained with the same basis set and the HF/6-31G* geometries.

The data in Table 3 show that the effect of geometry in this case is comparatively small, just under 1 kcal mol^{-1} for CCSD(T), for example. This is not to say that using a low-level geometry such as HF/6-31G* is generally recommended. Rather, the choice of geometry should be carefully considered, with factors such as desired accuracy and available resources being considered. Certainly, with extended basis sets, obtaining a CCSD(T) geometry can be very time consuming, and this has to be weighed against benefits. In practice, it is advisable to study the effect of geometry on the final results using both some low-level and intermediate-level geometries before obtaining a very high-level geometry.

Results from Correlation-Consistent Basis sets

Within the family of so-called Pople-type basis sets, 6-311++G(3df,3pd) is the largest standard set. Even though this basis set is quite large, it is not a complete basis set for correlated methods. Better basis sets for the systematic description of electron correlation, the so-called correlation-consistent (cc) basis sets, were developed some years later, the first being published in 1989 (Dunning 1989). The largest of such sets are recognized as providing a pathway to basis set completeness. At the same time, these basis sets are quite large, and basis set completeness for wave function-based correlated methods generally requires the use of an extrapolation. The cc basis sets were developed in such a way as to permit reasonably reliable extrapolation to the complete basis set result. Of course, the larger the basis sets used in the extrapolation, the more reliable the extrapolation is likely to be. Helgaker et al. (2000) discuss the convergence and performance of the cc sets quite extensively.

We have calculated the ΔE using two series of the cc basis sets. The first series used the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets. The cc-pVDZ set is a valence double-zeta set with one set of polarization functions (p for H, d for Li-Ar). In terms of contracted functions, it is equivalent to 6-31G**. The cc-pVTZ and cc-pVQZ sets are valence triple- and quadruple-zeta sets. They have multiple sets of polarization functions (2p1d and 2d1f for cc-pVTZ; 3p2d1f and 3d2f1g for cc-pVQZ). In the second series, we have used the augmented sets, namely, aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ. These contain a set of diffuse functions for each angular momentum, for example, there are diffuse s, p, and d sets in cc-pVDZ. In terms of contracted functions, the cc-pVTZ set is the same size as 6-311G(2df,2pd). The aug-cc-pVTZ set would be expected to be similar in accuracy to 6-311++G(2df,2pd). The next table shows the results from the cc-pVDZ, cc-pVTZ, and cc-pVQZ basis sets (Table 4).

At the HF level, there is a small change as the basis set is extended, as observed with the other basis sets considered in this section. For MP2, CCSD, and CCSD(T), there is a considerable change from cc-pVDZ to cc-pVTZ (over 10 kcal mol⁻¹), followed by a much smaller change on going to cc-pVQZ. The changes for BLYP and B3LYP are smaller. It is natural to wonder how well converged are the cc-pVQZ results. Larger basis sets in the cc series have been defined, and one could employ these in calculation on these systems (Helgaker et al. 2000). An alternative,

Table 4 Calculated ΔE values for $3\text{H}_2(\text{g}) + \text{N}_2(\text{g}) \rightarrow 2\text{NH}_3(\text{g})$ using the cc-pVXZ (X = D, T, Q) basis sets (units are kcal mol⁻¹)

	cc-pVDZ	cc-pVTZ	cc-pVQZ
HF	-32.4	-34.0	-34.9
MP2	-23.6	-34.2	-37.3
CCSD	-26.7	-36.9	-39.7
CCSD(T)	-24.0	-34.8	-37.9
BLYP	-26.1	-30.9	-33.3
B3LYP	-34.1	-38.5	-40.6

mentioned above, is to use an extrapolation formula to estimate the complete basis set result. For example, to use the L^{-3} extrapolation the formula (Halkier et al. 1998; Helgaker et al. 2000):

$$W(\infty) = \frac{M^3 W(M) - N^3 W(N)}{M^3 - N^3}$$

could be applied to a pair of ΔE values for a method. In this formula, M and N are the highest angular momentum quantum numbers of two basis sets, while $W(M)$ and $W(N)$ are the values of the quantity being extrapolated with the two corresponding basis sets. For example, suppose we extrapolate the CCSD(T) results with the cc-pVTZ and cc-pVQZ basis sets: $M = 3$ (for the f functions in cc-pVTZ), $W(3) = -34.8$, $N = 4$ (for the g functions in cc-pVQZ), and $W(4) = -37.9$. Applying the extrapolation formula, the result is -40.2 kcal mol $^{-1}$. Applying the extrapolation formula to the cc-pVDZ and cc-pVTZ data for CCSD(T) yields -39.3 kcal mol $^{-1}$. In this case, both extrapolations give quite similar results. This will not always be the case. The larger the basis sets used in the extrapolation, the more reliable the result should be. The next set of results to consider are those with the augmented cc basis sets. These are shown in the Table 5.

Going from cc-pVTZ to aug-cc-pVTZ the ΔE decreases by about 4 kcal mol $^{-1}$ for the correlated methods and just over 2 kcal mol $^{-1}$ for HF. The effect of adding diffuse functions to cc-pVQZ is smaller, namely, about 1.5 kcal mol $^{-1}$ for the correlated methods. The diffuse functions in aug-cc-pVTZ effectively make up for some the valence basis set incompleteness in cc-pVTZ, while cc-pVQZ is more complete, so the effect of the diffuse functions is smaller. The difference between the aug-cc-pVTZ and aug-cc-pVQZ results is very small in this case.

Finally, based on the data in Table 5 and an experimentally derived value of ΔE of -39.5 kcal mol $^{-1}$ (Helgaker et al. 2000), one can see the CCSD(T) results approaching chemical accuracy for this problem. MP2 also performs very well in this case, but this is not always the case. Helgaker et al. (2000) report an extrapolated CCSD(T) result from cc-pcV5Z and cc-pcV6Z values of -39.6 kcal mol $^{-1}$. The cc-pVXZ basis sets include additional functions to account for core correlation effects that are not included in the cc-pVXZ sets.

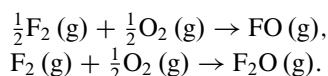
Table 5 Calculated ΔE values for $3\text{H}_2(\text{g}) + \text{N}_2(\text{g}) \rightarrow 2\text{NH}_3(\text{g})$ using the aug-cc-pVXZ (X = T, Q) basis sets (units are kcal mol $^{-1}$)

	cc-pVTZ	cc-pVQZ
HF	-36.2	-35.7
MP2	-38.6	-38.7
CCSD	-40.8	-41.0
CCSD(T)	-39.1	-39.3
BLYP	-35.4	-35.2
B3LYP	-42.3	-42.1

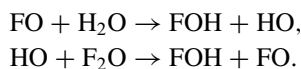
Thermochemistry of Species with F–O Bonds

Species containing F–O bonds have provided a significant challenge for theoretical methods over the years. The most celebrated case is fluorine peroxide, the theoretical structure of which is arguably an unsolved problem, as is its heat of formation. As far as thermochemistry is concerned, getting reliable, chemically accurate heats of formation for several fundamental species (e.g., FO, F₂O, HOF, FOO, and F₂O₂) is very difficult. For some of these species (FOO and F₂O₂) additional difficulties are possible uncertainties in experimental data, which make it more difficult to assess the quality of theoretical data. One theoretical strategy is to use the most advanced theoretical methods (e.g., CCSD(T)) and large basis sets, perhaps including attempted extrapolation to the (estimated) basis set limit. Such calculations place severe demands on resources and are applicable only to relatively small species, but they should be capable of chemical accuracy when carried far enough. Another strategy is to use DFT methods. These are much less demanding computationally and thus can be applied to much larger species. The key issue is their accuracy: can they provide sufficient accuracy on small species to be applied with confidence on larger species? A study by Ventura et al. (1999) involves a comparative study of two functionals and the CCSD(T) method for species with F–O bonds, which we now discuss.

Ventura et al. undertook this study since there were conflicting reports on the suitability of DFT methods for systems containing F–O bonds. Their first goal was to provide a systematic, side-by-side comparison of two functionals, namely, B3LYP and B3PW91, and the CCSD(T) method using the same large atomic natural orbital basis set. The basis set chosen was an uncontracted 14s9p4d3f set (for N, O, and F) combined with an uncontracted 8s4p3d set for H. Two strategies were used to calculate heats of formation: (a) calculating the enthalpy changes for the defining equations of ΔH_f^0 (referred to hereafter as the direct method) and (b) calculating the enthalpy changes for appropriate isodesmic equations and then using experimental data on all but one species to calculate the enthalpy of formation of the unused species. For example, approach (a) would use the following equations for FO and F₂O:



The isodesmic approaches for FO and F₂O used the equations



To calculate the heat of formation of FO, for example, one would calculate the ΔH^0 of the first equation and then combine this result with experimental standard enthalpies of formation of H₂O, HOF, and OH:

$$\Delta H_f^0(\text{FO}) = \Delta H_f^0(\text{HO}) + \Delta H_f^0(\text{HOF}) - \Delta H_f^0(\text{H}_2\text{O}) - \Delta H^0.$$

DFT and CCSD(T) approaches were used to obtain a series of standard enthalpies of formation using both the direct (referred to by the authors as the diatomic approach) and isodesmic approaches. The species with F–O bonds studied were FO, HOF, and F₂O. In addition, the authors calculated heats of formation of the H, N, O, and F atoms, the OH and NO radicals, and H₂O.

The general findings of Ventura et al. can be summarized as follows. The DFT methods performed better using the isodesmic method than they did for the direct method. For the CCSD(T) method, however, the reverse was true. For the isodesmic method, the DFT methods perform better than the CCSD(T) method, but CCSD(T) performs better than the DFT methods for the direct method. Overall the performance of all methods is generally very good. Representative data are shown in the table below (Table 6).

Two values are given for the isodesmic method for FO and HOF, one for each of the two equations given above. Whether the isodesmic or direct method is used for the DFT approaches, the results are impressive, especially considering that the time required for these calculations is a fraction of that needed for CCSD(T). The CCSD(T) results would likely be improved slightly with larger basis sets, while the effect is likely to be much smaller for the DFT methods.

Another instructive aspect of the work of Ventura et al. (1999) comes from the use of the 6-31G* basis set. Such a basis set is obviously not expected to be accurate enough for a direct calculation of the heat of formation of a molecule. However, as the following data show, when the isodesmic approach is used, very promising results are obtained (Table 7).

For HOF in particular, one can see that the isodesmic approach is a major improvement over the direct method for all methods. This illustrates the potential value of isodesmic approaches. The good result obtained for CCSD(T) for FO with

Table 6 Calculated and experimental ΔH_f^0 of FO, HOF, and F₂O (units are kcal mol⁻¹; the data are from Ventura et al. 1999)

	Experiment	B3PW91		B3LYP		CCSD(T)	
		Iso	Direct	Iso	Direct	Iso	Direct
FO	26.06	25.75	24.93	26.50	24.31	28.01	26.42
		26.42		26.29		27.20	
HOF	-20.60	-20.30	-18.84	-21.05	-19.20	-22.56	-20.83
		-20.41		-20.36		-19.45	
F ₂ O	5.86	5.67	5.78	5.62	5.48	4.71	6.52

Table 7 Calculated and experimental ΔH_f^0 of FO and HOF using the 6-31G* basis set (units are kcal mol⁻¹; the data are from Ventura et al. 1999)

	Experiment	B3PW91		B3LYP		CCSD(T)	
		Iso	Direct	Iso	Direct	Iso	Direct
FO	26.06	26.44	23.08	27.04	24.53	30.28	26.24
HOF	-20.60	-20.99	-13.60	-21.56	-13.47	-24.83	-13.16

the 6-31G* is no doubt fortuitous. Ventura et al. (1999) do not report results for F₂O using the 6-31G* basis set.

One cautionary comment worth making here concerns the use of experimental data in schemes with isodesmic reactions. The success of such schemes obviously depends on the reliability of the experimental data, certainly requiring chemical accuracy in the experimental data. For FO, HOF, and F₂O, this appears to be the case. As mentioned above, this is probably not true for all F–O containing species. Some theoretical studies (Feller and Dixon 2003; Karton et al. 2009) have found a disagreement between high-quality calculated data and experimental data for FOO and F₂O₂ and recommend revision and/or a new determination of experimental data for these species.

Calculating the Heat of Formation of H₂SO₂

In a recent study (Napolion et al. 2008), isomers of H₂SO₂ were investigated theoretically using the CCSD(T) and other methods. Part of that study was calculating the standard heat of formation of the lowest energy isomer, namely, the C₂ conformer of sulfoxylic acid, S(OH)₂. This species has been detected experimentally, but no structural, spectroscopic (e.g., infrared), or thermochemical data have been measured. There had been some prior theoretical estimates, but the variations were significant. In 1978, using bond energy estimates, Benson (Benson 1978) estimated ΔH_f^0 to be -69.3 ± 4.1 kcal mol⁻¹. Wang and Zhang (2002) calculated ΔH_f^0 in two ways using the G3B3 and G3//MP2 procedures. First, they calculated the enthalpy change for dissociation to atoms. Then they used the experimental ΔH_f^0 values of atoms to obtain a ΔH_f^0 for S(OH)₂. With this approach they obtained values of -65.2 (G3B3) and -65.9 kcal mol⁻¹ (G3//MP2), which are at the lower end of Benson's estimate. However, Wang and Zhang were concerned about possible inaccuracies in the atomization procedure, so they also used a procedure that involved isodesmic reactions. This gave estimates of -69.5 kcal mol⁻¹ (G3B3) and -68.3 to -70.0 kcal mol⁻¹ (G3//MP2). Wang and Zhang's preferred value was -69.3 kcal mol⁻¹. Our strategy was to use the CCSD(T) method with extended basis sets, including correlation-consistent basis sets and L⁻³ extrapolation.

A direct approach to calculating the ΔH_f^0 was not possible since the standard state of sulfur is solid sulfur. Instead, two schemes were devised, each of which involved calculating the ΔH^0 of a reaction that could be combined with experimental data to yield an estimate of the required ΔH_f^0 . In the limit of an exact calculation, both schemes will give the same final result. Hence, the closeness of the two estimates is to some extent a measure of the accuracy. The two schemes are as follows.

Scheme 1 The standard enthalpy change for the reaction S(OH)₂ (g) → SO₂ (g) + H₂ (g) was calculated. The ΔH_f^0 was then calculated from

$$\Delta H_f^0 = \Delta H_f^0(\text{SO}_2) + \Delta H_f^0(\text{H}_2) - \Delta H_{rxn}^0 = -70.9 \text{ kcal mol}^{-1} - \Delta H_{rxn}^0,$$

using the calculated ΔH_{rxn}^0 and experimental data for SO₂ and H₂.

Scheme 2 The standard enthalpy change for the reaction $\text{S}(\text{OH})_2(\text{g}) \rightarrow \text{H}_2\text{S}(\text{g}) + \text{O}_2(\text{g})$ was calculated. The ΔH_f^0 was then calculated from

$$\Delta H_f^0 = \Delta H_f^0(\text{H}_2\text{S}) + \Delta H_f^0(\text{O}_2) - \Delta H_{rxn}^0 = -4.8 \text{ kcal mol}^{-1} - \Delta H_{rxn}^0,$$

using the calculated ΔH_{rxn}^0 and experimental data for SO_2 and H_2 . The difference

$$\Delta H_f^0(\text{Scheme 1}) - \Delta H_f^0(\text{Scheme 2})$$

is equal to $\Delta H^0(\text{Expt.}) - \Delta H^0(\text{Calc.})$ for the reaction $\text{H}_2\text{S}(\text{g}) + \text{O}_2(\text{g}) \rightarrow \text{SO}_2(\text{g}) + \text{H}_2(\text{g})$. For our most complete calculations (see below), the difference is only $-0.6 \text{ kcal mol}^{-1}$, which is within chemical accuracy. This is a significant achievement since it is well known to be very difficult to obtain accurate thermochemical data on SO_2 . This suggests that our strategy is suitable for calculating the ΔH_f^0 of $\text{S}(\text{OH})_2$.

The results of our calculations are shown in Tables 8 and 9 (see below). Each table shows three quantities. ΔE is the energy change of the reaction for stationary nuclei at 0 K. ΔH^0 is the calculated standard enthalpy change of the reaction at 298.15 K, obtained from ΔE and the unscaled CCSD(T)/6-311++G(2d,2p) thermal correction (-6.4 and $-5.4 \text{ kcal mol}^{-1}$ for Schemes 1 and 2, respectively). ΔH_f^0 is the calculated standard enthalpy of formation of $\text{S}(\text{OH})_2$ at 298.15 K, obtained as indicated above.

Before discussing the data, a note on the cc basis sets used is needed. When the cc basis sets for third-row elements such as sulfur were first developed, the numbers of polarization functions used were the same as in the basis sets for the second-row elements (e.g., 1d for cc-pVDZ, 2d1f for cc-pVTZ, and so on). It was subsequently found that the d set for third-row elements was inadequate and needed to be expanded by adding a set of “tight” d functions. The resulting basis sets are referred to as cc-pV(X+d)Z and aug-cc-pV(X+d)Z. For H and second-row

Table 8 CCSD(T) values of ΔE , ΔH^0 , and ΔH_f^0 from Scheme 1 (see above) in kcal mol^{-1}

	ΔE	ΔH^0	ΔH_f^0
6-311++G(2d,2p)	8.2	1.8	-72.7
6-311++G(3df,3pd)	5.1	-1.3	-69.6
ANO-TZP	8.1	1.7	-72.6
cc-pV(D+d)Z	8.6	2.2	-73.1
cc-pV(T+d)Z	3.9	-2.4	-68.5
cc-pV(Q+d)Z	4.0	-2.4	-68.6
aug-cc-pV(D+d)Z	13.3	6.9	-77.8
aug-cc-pV(T+d)Z	5.5	-0.9	-70.1
aug-cc-pV(Q+d)Z	4.5	-1.9	-69.0
Extrapolation ^a	4.1	-2.3	-68.6
Extrapolation ^b	3.8	-2.6	-68.3

^aExtrapolation based on the cc-pVTZ and cc-pVQZ results

^bExtrapolation based on the aug-cc-pVTZ and aug-cc-pVQZ results

Table 9 CCSD(T) values of ΔE , ΔH^0 , and ΔH_f^0 from Scheme 2 (see above) in kcal mol⁻¹

	ΔE	ΔH^0	ΔH_f^0
6-311++G(2d,2p)	56.5	51.2	-56.0
6-311++G(3df,3pd)	62.6	57.2	-62.1
ANO-TZP	63.9	58.6	-63.4
cc-pV(D + d)Z	43.4	38.0	-42.8
cc-pV(T + d)Z	61.4	56.1	-60.9
cc-pV(Q + d)Z	65.3	60.0	-64.8
aug-cc-pV(D + d)Z	57.5	52.1	-56.9
aug-cc-pV(T + d)Z	64.6	59.2	-64.1
aug-cc-pV(Q + d)Z	66.6	61.3	-66.1
Extrapolation ^a	68.2	62.9	-67.7
Extrapolation ^b	68.1	62.7	-67.6

^aExtrapolation based on the cc-pVTZ and cc-pVQZ results

^bExtrapolation based on the aug-cc-pVTZ and aug-cc-pVQZ results

elements, these sets are the same as the original cc-pVXZ and aug-cc-pVXZ sets, but for third-row elements they have an extra set of d functions.

Looking at Table 8, one sees that there are quite small differences between the triple- and quadruple-zeta results, the largest being 1.1 kcal mol⁻¹ between the aug-cc-pV(T + d)Z and aug-cc-pV(Q + d)Z results. Consequently, the extrapolated results differ little from the cc-pV(Q + d)Z and aug-cc-pV(Q + d)Z values. Also, the two extrapolated results differ by only 0.3 kcal mol⁻¹. The 6-311++G(3df,3pd) basis set performs fairly well overall. The cc-pV(D + d)Z and aug-cc-pV(D + d)Z results are not as good, especially the latter.

The data in Table 9 show greater sensitivity to basis set than do those in Table 8, no doubt a reflection of the demands on the basis set in calculations on SO₂. In Table 9, one can see much larger differences between the triple- and quadruple-zeta results. Not surprisingly, the extrapolation has a larger effect than in extrapolation has a larger effect than in Table 8. However, the two extrapolations give ΔH_f^0 values that differ by only 0.1 kcal mol⁻¹. The 6-311++G(3df,3pd) basis set is seen not to perform as well in Table 9 as in Table 8. The deficiencies of the smaller basis sets are more apparent in Table 9 than in Table 8.

A measure of the quality of the results can be obtained by comparing the ΔH_f^0 values in Tables 8 and 9. Even for the aug-cc-pV(Q + d)Z basis set, there is a 2.9 kcal mol⁻¹ difference, which is not within chemical accuracy, of course. However, the corresponding extrapolated values in the two tables differ by less than 1 kcal mol⁻¹, which is within chemical accuracy. Our best results for ΔH_f^0 , then, are 1–2 kcal mol⁻¹ above Benson's mean value and in the range of values obtained by Wang and Zhang by isodesmic reactions, rather than their estimates based on dissociation reactions. Our reaction enthalpies predict that decomposition of S(OH)₂ to SO₂ + H₂ is slightly exothermic, while decomposition to H₂S + O₂ is strongly endothermic.

Acknowledgments This work was supported in part by the National Science Foundation. Calculations performed in this work on the formation of NH_3 were run on a QS4-2800C computer from Parallel Quantum Solutions, LLC using the Gaussian 2003 software.

Bibliography

- Benson, S. W. (1978). Thermochemistry and kinetics of sulfur-containing molecules and radicals. *Chemical Reviews*, 78, 23–35.
- Dunning, T. H., Jr. (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *Journal of Chemical Physics*, 90, 1007–1023.
- Feller, D., & Dixon, D. A. (2003). Coupled cluster theory and multireference configuration interaction study of FO, F₂O, FO₂, and FOOF. *Journal of Physical Chemistry A*, 107, 9641–9651.
- George, P., Trachtman, M., Bock, C. W., & Brett, A. M. (1975). An alternative approach to the problem of assessing stabilization energies in cyclic conjugated hydrocarbons. *Theoretica Chimica Acta*, 38, 121–129.
- Halkier, A., Helgaker, T., Jørgensen, P., Klopper, W., Koch, H., Olsen, J., & Wilson, A. K. (1998). Basis-set convergence in correlated calculations on Ne, N₂, and H₂O. *Chemical Physics Letters*, 286, 243–252.
- Hehre, W. J. (2003). *A guide to molecular mechanics and quantum chemical calculations*. Irvine: Wavefunction.
- Hehre, W. J., Ditchfield, R., Radom, L., & Pople, J. A. (1970). Molecular orbital theory of the electronic structure of organic compounds. V. Molecular theory of bond separation. *Journal of the American Chemical Society*, 92, 4796–4801.
- Hehre, W. J., Radom, L., Schleyer, P. V. R., & Pople, J. A. (1986). *Ab initio molecular orbital theory*. New York: Wiley.
- Helgaker, T., Jørgensen, P., & Olsen, J. (2000). *Molecular electronic-structure theory*. Chichester: Wiley.
- Irikura, K. K., & Frirup, D. J. (Eds.). (1998). *Computational thermochemistry*. Washington: American Chemical Society.
- Karton, A., Parthiban, S., & Martin, J. M. (2009). Post-CCSD(T) ab initio thermochemistry of halogen oxides and related hydrides XOX, XOOX, HOX, XO_n, and HXO_n (X = F, Cl), and evaluation of DFT methods for these systems. *Journal of Physical Chemistry A*, 113, 4802–4816.
- Lewars, E. (2003). *Computational chemistry*. Boston: Kluwer.
- McQuarrie, D. A., & Simon, J. D. (1999). *Molecular thermodynamics*. Sausalito: Science Books.
- Minkin, V. I. (1999). Glossary of terms used in theoretical organic chemistry. *Pure and Applied Chemistry*, 71, 1919–1981.
- Napoleon, B., Huang, M.-J., & Watts, J. D. (2008). Coupled-cluster study of isomers of H₂SO₂. *Journal of Physical Chemistry A*, 112, 4158–4164.
- Ochterski, J. W. (2000). Thermochemistry in Gaussian. http://Gaussian.com/g_whitepap/thermo.htm
- Pople, J. A., Frisch, M. J., Luke, B. T., & Binkley, J. S. (1983). A Møller-Plesset study of the energies of AH_n molecules (A = Li to F). *International Journal of Quantum Chemistry Symposium*, 17, 307–320.
- Ventura, O. N., Kieninger, M., & Cachau, R. E. (1999). Density functional theory is more accurate than coupled-cluster theory in the study of the thermochemistry of species containing the F–O bond. *Journal of Physical Chemistry A*, 103, 147–151.
- Wang, L., & Zhang, J. (2002). Ab initio calculation on thermochemistry of CH₃SO_xH (x = 1–3) and H₂SO_y (y = 2,3). *Journal of Molecular Structure: THEOCHEM*, 581, 129–138.

Calculation of Excited States: Molecular Photophysics and Photochemistry on Display

15

Luis Serrano-Andrés and Juan José Serrano-Pérez

Contents

Introduction	640
Spectroscopy Overview	642
Electronic Structure Calculations	649
General Overview	649
Methods, Advantages, and Drawbacks	653
Basis Sets	660
Methods for Excited States	662
How to Compute Excited States	669
How to Start: Selection of Goals, Methods, Geometries	669
Molecular Photophysics: Computing Absorption and Emission Spectra	671
Computing Rydberg States	680
Electronic States of Anionic Systems	686
Photochemistry: On the Trail of the Energy	693
Photoinduced Reactions in Bimolecular Systems	707
Conclusions	720
Bibliography	722

Dedicated to the memory of Luis Serrano-Andrés, great scientist, teacher and friend, who passed away recently.

L. Serrano-Andrés
Instituto de Ciencia Molecular, Universitat de València, Valencia, Spain

J.J. Serrano-Pérez (✉)
Department of Chemistry, Imperial College London Computational and Structural Research Group, London, UK
e-mail: Juan.Serrano@uv.es

Abstract

Excited states participate in photoinduced events as well as in thermally activated reactions, even in many cases in which only the ground state is believed to be involved. Life on Earth also depends, both directly and indirectly, on the influence that light has on chemistry. The energy of the Sun's visible and ultraviolet radiation promotes processes that not only permit the continued existence of life on the planet, but which are keys for evolution by means of mutations. To study a system in an excited state, far away from its optimum situation, is a challenge for chemists, both experimentalists and theoreticians. This chapter is focused on the practical aspects related to the calculation of excited states in molecular systems by using quantum-chemical methods, a type of study that escapes in many cases from the well-established computational strategies used for the molecular ground states, both because of the complexity of the problem itself and for the methodological requirements. A short review of the spectroscopic and photochemical panorama will be provided first in order to explain which are the main parameters and processes to be determined, followed by a compact description of the most relevant and employed quantum-chemical methods and computational strategies for excited states. A number of applied examples of actual calculations on paradigmatic excited state problems will be provided in the different subchapters, followed in each case by comments on practical issues occurring in the calculations. With these cases we will try to demonstrate that in the last years the quantum-chemical studies on excited states have reached the required maturity to interpret and predict, at a molecular level, different types of chemical situations.

Introduction

Computing electronic excited states with quantum chemical methods is much more complex than doing it for ground states, because it implies not only coping with higher solutions of the electronic Hamiltonian with diverse character and therefore requiring more complex methods, but also to solve a plethora of new situations, such as hypersurface crossings and coupling between the states, that usually requires to abandon convenient approximations like Born–Oppenheimer (see Szabo and Ostlund 1996), for instance.

Molecules, consisting of electrically charged nuclei and electrons, may interact with the oscillating electric and magnetic fields of light. Spectroscopic experiments demonstrate that energy can be absorbed or emitted by molecules (and atoms) in discrete amounts, corresponding to precise changes in energy of the molecule (or atom) concerned. As matter, light is a form of energy that exhibits both wave- and particle-like properties. Absorption of the relevant frequencies from incident radiation raises molecules from lower to higher levels. Electrons in molecules occupy molecular orbitals (MOs) with precise energy levels. Transitions from lower, filled orbitals, to upper (higher energy), empty orbitals usually involve absorption of radiation in the UV and visible parts of the spectrum. Much smaller quantities of

energy are linked to changes in the vibrational and rotational energy of the molecule. The understanding of the spectroscopic phenomena in the light of molecular orbital theory has opened new avenues in the comprehension of the photoinduced events.

The molecule in excited state is often prone to react in an easier way than in the ground state. The excess energy of an excited species can alter its reactivity and is particularly significant in the case of electronic excitation because of the energies involved are of similar order of magnitude as bond energies. Electronic excitations can then have a considerable effect on the structure of a species. Accordingly, the energies correspond roughly with typical activation energies for many reactions, which are too high to be reached from the ground but not from the excited state. The new electronic rearrangement may be also the key of the reactivity since the molecule in an excited state may exhibit nucleophilic or electrophilic properties different than those of its ground state.

Three modern developments have been produced in the last years that are the key for the comprehension of the photophysics and photochemistry of many chemical and biochemical phenomena: (1) rapid advances in quantum-chemical methods allow to study the excited states with high accuracy; (2) improved molecular beams techniques permit studies of isolated molecules, despite their sometimes low vapor pressure and propensity for thermal decomposition, and (3) the revolutionary impact that femtosecond laser and multiphoton techniques have had on the study of the electronic energy relaxation processes. Indeed, now it is possible to get information about reaction intermediates at very short times from femtochemical techniques, and, more than ever, the participation of quantum chemistry to interpret such findings has become crucial. A constructive interplay between theory and experiment can provide an insight into the chemistry of the electronic state that cannot be easily derived from the observed spectra alone.

From the theoretical viewpoint, the calculation of excited states is still a very complex task. Considering the many different electronic structure situations occurring in the potential energy hypersurfaces (PEHs) of the excited molecular systems, the only methods generally applicable to all of them are the multiconfigurational approaches. The application of these procedures requires a lot of skill and experience, and the limitations on the size of the problem are noticeable. Single-reference (black-box) methods only work in certain regions of the PEHs. In general, the excited state problem can be considered heavily multiconfigurational. New tools and strategies are required for excited states at the highest levels of calculation: optimization of minima, transition states, hypersurface crossings (conical intersections), and reaction paths, whereas states couplings (nonadiabatic, electronic, spin-orbit) need to be computed. This solves only the first part of the problem, that is, the solution of the time-independent Schrödinger equation. Once the potential represented by the PEHs is obtained, time-dependent equations have to be solved to finally determine reaction rates, states lifetimes, or populations. Coupling at the proper level those two types of calculations, static and dynamic approaches representing the electronic structure and reaction dynamics problems, respectively, is still a task under development.

Spectroscopy Overview

The concept of PEHs comes from the Born–Oppenheimer approximation, based on the separation of electronic and nuclear motion due to the large difference in mass between these particles, and assumes that the electrons follow the nuclei instantaneously during the motion of the latter. Therefore, an electronic and a nuclear Hamiltonian can be defined. Solving the electronic Schrödinger equation provides us a description of the movement of electrons, whereas we consider the rotation, vibration, and translation of the molecule solving the nuclear counterpart. The solution of the electronic Schrödinger equation is the energy of a particular nuclear configuration. The total energy for fixed nuclei must also include the constant (within this approximation) nuclear repulsion potential. The value of this total potential energy for every possible nuclear configuration is specifically the potential energy hypersurface. Photophysical and photochemical processes take place through interactions between PEHs. For the specific case of close degeneracies between the surfaces, where ultrafast energy transfers occur, the Born–Oppenheimer approximation breaks down and special methods in order to localize, optimize, and study the crossing structures are required.

The absorption of photons for a molecule is hardly a static problem. After the absorption (ABS) of one photon, a state of the same multiplicity as the ground state is mainly populated. Direct absorption to states of different multiplicity are only possible if the states heavily interact, for instance, by spin-coupling effects. Actually, in the general case, the energy goes to a vibrational excited state of an electronic excited state of the molecule. Straight afterward, a non-radiative decay occurs, with emission of heat (IVR, intramolecular vibrational relaxation), toward more stable structures of the state PEH, in many cases the state minimum. It might frequently happen that along the decay other states cross and, if appropriately coupling, the system can evolve toward other electronic states of the same multiplicity via a non-radiative internal conversion (IC). Finally, the molecule arrives to the lowest-lying singlet excited state, S_1 , from which the molecule may emit (F, fluorescence) and return to the ground state. Alternatively, a non-radiative transition between two states of different multiplicity is also possible (ISC, intersystem-crossing). By successive internal conversions the system reaches the lowest-lying triplet excited state, T_1 , from which the molecule may emit (P, phosphorescence).

Figure 1 contains a simplified Jablonski diagram summarizing the main photophysical and photochemical effects undergone by a molecular system. It is common to reserve the word photophysics to processes not involving the generation of new photospecies, that is, just to decays leading to emission or returns to the ground state. However, nonradiative internal conversions or intersystem crossing are also considered photochemical processes, therefore the use of both terms is somewhat loose. A proper nomenclature for excited states is not easy to establish. The less ambiguous (and less informative too) form is purely enumerative: S_0 , S_1 , T_1 , T_2 , where S represents a singlet state and T stands for the triplet states, and the states are ordered by energy. In symmetric systems, it is convenient to use the labels

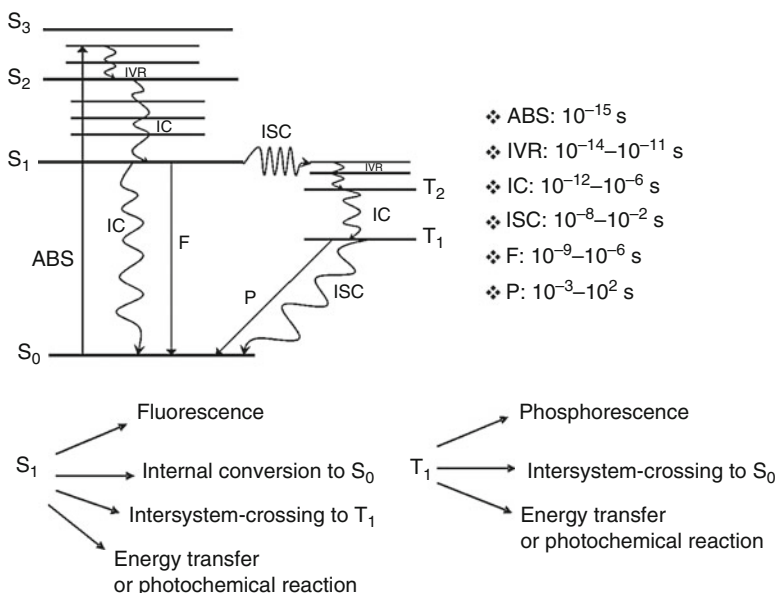


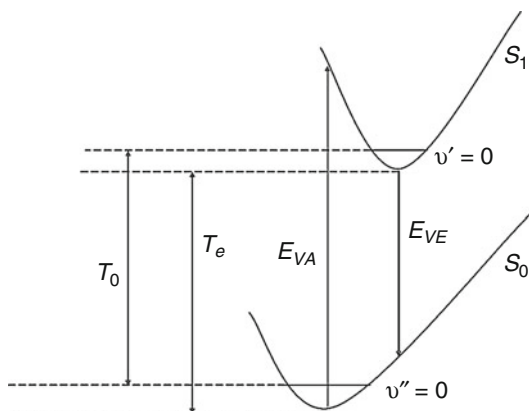
Fig. 1 Jablonski's diagram, lifetimes of the basic photophysical processes and deexcitation pathways from the lowest-lying excited states of a molecule

derived from group theory. A state is then described in terms of the behavior of the electronic wave function under the symmetry operations of the point group which the molecule belongs to $1^1A_g, 2^3B_{3u} \dots$, including the energy ordering, multiplicity, and symmetry label. Classical nomenclatures, such as those developed by Mulliken (N, V, T), Kasha ($^1\pi\pi^*$, $^3\sigma\pi^*$), and Platt (1L_a , 3B_b), highlighting the type of orbitals implied in the excitation, can be useful in different cases.

The basic information required to rationalize a photoinduced phenomenon is the energy levels of the excited solutions and the probability of energy (population) transfer from one state to the other. In the semiclassical treatment of the interaction radiation-matter, whereas we treat the molecule quantum-mechanically, the radiation field is seen as a classical wave obeying Maxwell's equations.

The electric and magnetic fields of the radiation will interact with the atomic or molecular electrons giving a time-dependent perturbation. Solving the time-dependent Schrödinger equation provides us the comprehension of absorption and stimulated emission, whereas to explain spontaneous emission we need the machinery of quantum electrodynamics. The resonance condition provided by the energy differences between the different PEHs of the corresponding states relates to the absorbed or emitted energy quanta. Regarding the transfer probability, it is related with the strength of the interaction between the time-dependent field and the multipolar (the dipole, d , approach is usually enough) charge distribution of the

Fig. 2 Vertical energies and band origins. Vibrational states are labeled with the greek character ν



molecular system. Such strength is proportional to the transition dipole moment (TDM) (Fig. 2).

As an initial feature, the electronic states produce a superposition of bands which characterize the absorption spectrum. The range of absorbed energies fluctuates between the vertical absorption energy, E_{VA} (difference between the minimum of the ground state and the excited state at the same geometry, that is, the Franck–Condon transition), and the adiabatic transition or band origin, T_e (difference between the excited state and the ground state at their respective optimized equilibrium geometries): it is the minimal energy difference allowed in absorption if the assumption that all excitations begin from the relaxed ground state is considered, as well as the largest energy emitted from the relaxed excited state. In many cases, the determination of T_e provides enough information to assign band origins; however, the zero-point vibrational energy (ZPE) has to be included in both initial and final states to get the vibrational band origin, T_0 , which can be directly compared to the experimental value, at least that obtained in the gas phase or in molecular beams. In addition, there is another magnitude named vertical emission energy, E_{VE} , which is the difference between the excited state and the ground state at the relaxed geometry of the former. The Franck–Condon principle stipulates that the vertical absorption can be related with the experimental band maximum. In fact, this is hardly the case, except when the ground and excited states have very similar geometries, and in this case the T_0 transition ($S_0\nu'' = 0 \rightarrow S_1\nu' = 0$) yields the most intense peak. The vertical excitation has, however, no experimental counterpart, whereas to get a true band maximum, the band vibrational profile must be computed. The only direct comparison relates the theoretical and experimental band origins, T_0 actually. Trying to assess the quality of a theoretical approach by comparing theoretical vertical excitations and experimental band maxima is one of the most frequent mistakes seen in the literature.

Regarding the transfer probability, far from the conical intersection regions the Fermi's Golden Rule is employed, in which the one-photon (optical) transition probability between two states $|n\rangle \rightarrow |m\rangle$ is proportional to the square of the

TDM between such states: $\text{TDM} = \langle m | d | n \rangle$. Two- and higher-order multiphoton probabilities can be also obtained. Based on symmetry considerations, selection rules for electronic transitions have been developed because only the totally symmetric matrix elements yield allowed nonzero probabilities (at first order). Using the so-computed electronic TDM, it is useful to estimate the electronic oscillator strength as $f = (2/3) E_{VA} \text{TDM}^2$, with E_{VA} being the vertical energy difference. The oscillator strength is a classically derived magnitude that represents the relative area of the electronic transition band and that it can be compared with the experimental estimation based in shapes and bandwidths. On the other hand, the vibrational contributions to the band intensity (or, in general, the strength of the transfer) can be obtained by computing the TDM between vibrational states. If belonging to the same electronic states, infrared or Raman intensities can be produced, otherwise electronic band vibrational profiles can be obtained. The vibrational TDM is proportional to the vibrational overlap term between the electronic states, $\langle v_m | v_n \rangle$, which are called the Franck–Condon factors (the probability of transition is proportional to its square). The vibrational profiles are basically related to the differences in geometry existing from the initial to the final electronic state and, therefore, the most intense progressions proceed through the normal modes which trigger the aforesaid changes. Within the harmonic approach, the complete TDM with respect the nuclear coordinates Q_i is defined as the Herzberg–Teller expansion:

$$\begin{aligned} \text{TDM} = & \text{TDM}(Q_0) \langle v_m(Q) | v_n(Q) \rangle \\ & + \sum_K \left(\frac{\partial \text{TDM}(Q)}{\partial Q_K} \right)_{Q_0} \langle v_m(Q) | Q_K | v_n(Q) \rangle + \dots \end{aligned}$$

Each one of the terms has an electronic and a vibrational component. The neglect of all terms except the first one is known as the Condon approximation, a usual way to proceed but only applicable for the one-photon dipole-allowed transitions. Otherwise, $\text{TDM}(Q_0)$ is zero by symmetry and the first term vanishes. Nonetheless, this approximation is valid only when the Born–Oppenheimer approximation is also valid. Otherwise, the phenomenon of vibronic coupling arises, which gives rise to other approximations (the term vibronic should be reserved to solutions obtained in non-Born–Oppenheimer cases, although is frequently used improperly). From the calculation of transition dipole moments, radiative lifetimes can also be obtained, both in fluorescence and phosphorescence by using the Einstein coefficients (A_{21}) and the Strickler–Berg relationships (Strickler and Berg 1962):

$$A_{21} = \frac{1}{\tau_{\text{rad}}} = 2.142005 \cdot 10^{10} E_{VE}^3 \text{TDM}_{1 \rightarrow 2}^2 \quad (1)$$

where τ_{rad} is the radiative lifetime measured in s^{-1} . The use of T_e instead of E_{VE} is more representative of the energy of the emission.

To study photophysical and photochemical processes on theoretical grounds, we need to determine the topography of the potential surfaces of the implied states (see Fig. 3). According to the different reaction paths through what a system

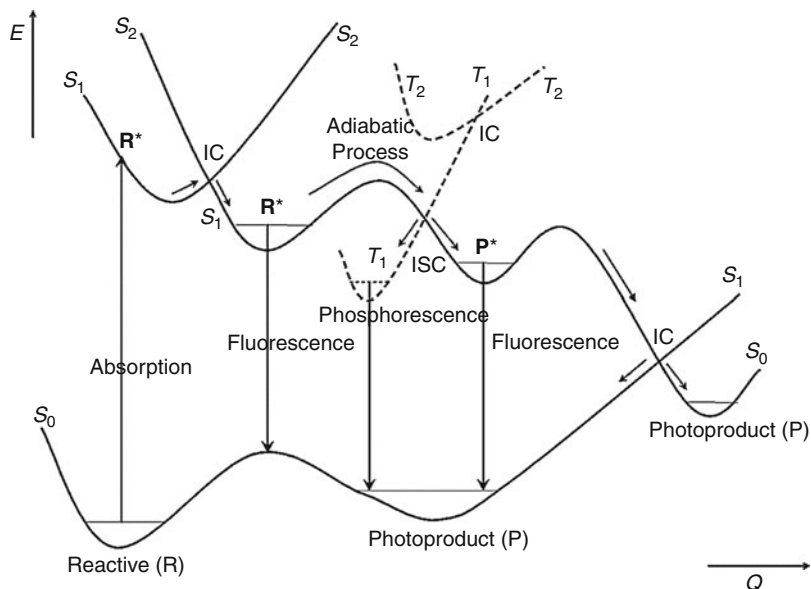


Fig. 3 Scheme of the main photophysical and photochemical molecular events. Notice that the order of the states changes along the nuclear coordinate, Q

might evolve, one can normally make the following classification that defines the photochemical panorama: adiabatic and non-adiabatic photochemistry. In an adiabatic reaction path, once vertical absorption takes place, the system proceeds along the hypersurface of the excited state to reach a local (or absolute) minimum leading in some cases to an emitting feature. On the contrary, in a non-adiabatic photochemical reaction, one part of the reaction takes place on the higher-state hypersurface and after a nonradiative jump at the surface crossing (or funnel) continues on the lower-state hypersurface. In a typical closed-shell ground-state molecule, the reaction usually begins on the potential energy surface of the excited state (S_1 or higher) at the Franck–Condon geometry (i.e., at the ground-state equilibrium geometry) and evolves either to the S_1 state minimum, from which it might emit, or to a crossing region with the ground state. Depending on the properties of such crossing, the process will end up on the reactant minimum or a new photoproduct minimum on the ground-state surface (S_0). The crossings between the excited state relevant from the photochemical view and the ground state are frequent, and they represent the basis of the spectroscopic phenomena. Therefore, a molecule evolving through the PEH of an excited state may well enter in a crossing region between two hypersurfaces during the lifetime of such an excited state. Hence, the lifetime of excited states is determined by the barriers that separate the excited states at the Franck–Condon geometry from the low-lying crossings. There are two types of crossings: conical intersections (CI), when the two interacting hypersurfaces have the same multiplicity, and (within the nonrelativistic approximation) singlet-triplet crossings (STC) or any other crossing between states

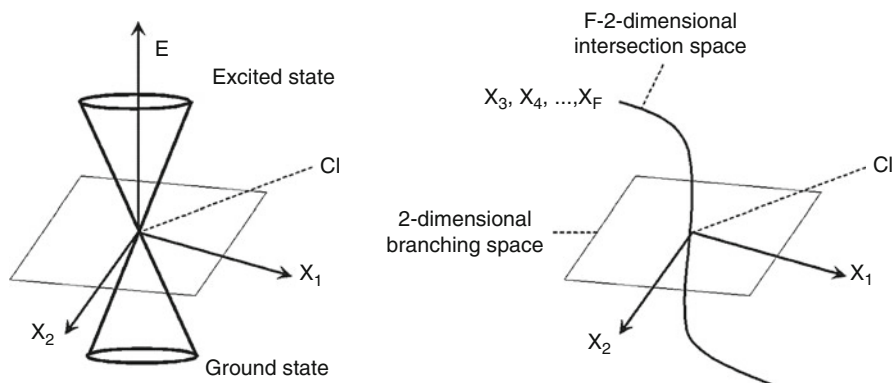


Fig. 4 Description of a conical intersection. The vectors x_1 and x_2 span the branching space. They are called the gradient difference (GD) vector and the derivative coupling (DC) vector. The DC vector measures the distortion of the system providing the maximum coupling between the electronic states involved in the crossing. The GD vector measures the distortion of the system leading to the largest variation of the energy difference between the two electronic states involved in the crossing

that have different multiplicity (Bearpark and Robb 2007; Domcke et al. 2004; Herzberg and Longuet-Higgins 1963; Robb et al. 1996; Teller 1937). Therefore, internal conversions take place through CIs, and intersystem crossings through STCs. The name of conical intersection (see Fig. 4) reflects the fact that a cone-shaped crossing is obtained when the energy of the states is plotted against the two privileged coordinates, the gradient differential vector, x_1 , and the nonadiabatic coupling vector, x_2 . Thus the total coordinate space F -dimensional is divided in two: the intersecting space (of dimension $F - 2$), in which both states are degenerated, and the branching space (of two dimensions). The most pronounced difference between the slopes of both hypersurfaces occurs along x_1 , whereas along x_2 the optimum nuclear displacement which mixes the two adiabatic wave functions in the CI point takes place. Actually, the intersection space is an hyperline that consists of an infinite number of CI points, i.e., a $(F - 2)$ -dimensional intersection space. To locate a CI point is equivalent to minimize the energy in the intersection space. In the case of STCs (if the nonrelativistic Hamiltonian is considered), we have only one privileged coordinates since the nonadiabatic coupling vector vanishes. Therefore we should refer this feature as an hyperplane, since we are moving along a $(F - 1)$ -dimension space.

As the Born–Oppenheimer approximation is not valid in regions where the electronic states become too close, it is logical that the nonadiabatic transfer is faster than the radiative relaxation. In essence, the former is based on the structure of the vibronic states. Certainly, the smaller is the gap between the states, the larger is the transfer probability.

Conical intersections may be peaked or sloped (see Atchity et al. 1991; Ben-Nun and Martínez 2000, and Fig. 5).

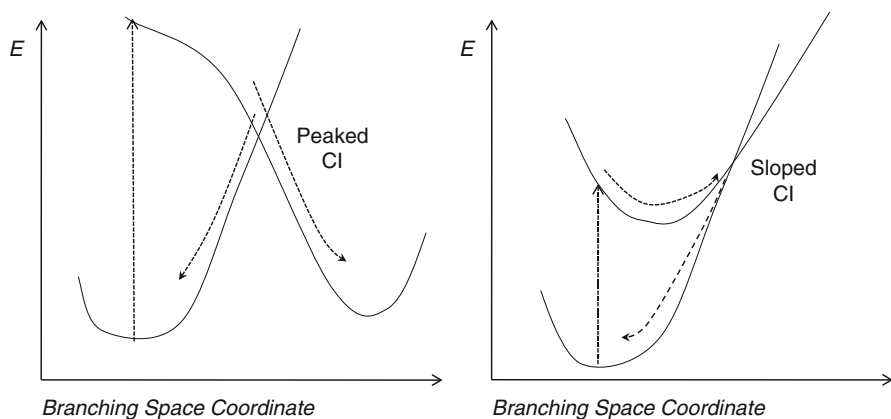


Fig. 5 Scheme of the limit characters displayed by the conical intersections. For a peaked CI, the two minima lie on the two different sides of the CI point and the gradients of the two intersecting PEHs are directed toward different directions. Conversely, for a sloped CI, the excited-state and ground-state minima lie on the same side (the gradients of the two PEHs point in the same direction), then the probability of re-crossing is higher and the efficiency of funneling is lower

If the connection between the two surfaces is sloped, the funnel may well be more efficient. In addition, the degeneracy at a crossing point can also be lifted at second order. As a consequence, we can choose a coordinate system in which to mix the branching and intersection space coordinates to remove this splitting and preserve the degeneracy to second order. These new coordinates give the curvature of the conical intersection hyperline and determine whether one has a minimum or a saddle point on it. These studies may also provide the vibrational modes that must be stimulated in order to enhance nonradiative decay because they decrease the energy gap and can lead to a CI (see Paterson et al. 2004; Sicilia et al. 2007).

A computational strategy can be designed, namely the Photochemical Reaction Path approach, in which the mechanism of the photoinduced process is accounted for by determining the fate of the energy on the populated state by computing the reaction profile. The whole process can be described by computing Minimum Energy Paths (MEPs), describing the lowest-energy, and therefore most favorable, although not unique, path for energy decay. The MEP is often built as steepest descent paths, guaranteeing the absence of barriers along the path. Each step requires the minimization of the PEH on a hyperspherical cross-section of the hypersurface centered on the initial geometry and characterized by a predefined radius. The optimized structure is taken as the center of a new hypersphere of the same radius, and the procedure is iterated until the bottom of the energy surface is reached. Mass-weighted coordinates are used, therefore the MEP coordinate corresponds to the so-called Intrinsic Reaction Coordinate (IRC), measured in a.u., that is, $\text{bohr} \cdot (\text{amu})^{1/2}$. The end of the path and the states crossed along the computed profile will inform about the fate of the energy, and, in particular, of the location of possible radiative minima and surface crossings, CIs and STCs. More crucial than

the presence of a crossing is its accessibility. The path of available energy should reach the crossing region to take place. Otherwise, if a too high energy barrier hinders the access to the crossing, the feature could be totally ineffective.

Electronic Structure Calculations

The development of Quantum Mechanics was spread over by Erwin Schrödinger, Werner Heisenberg, and Paul Dirac in the 1920s. The wave and particle aspects of matter are reconciled by the Schrödinger equation for stationary states, $\mathbf{H}\Psi = E\Psi$. The Hamiltonian operator, \mathbf{H} , is associated to the total energy of a physical system and is the sum of the kinetic energy and the potential energy operators associated with electrons and nuclei ($\mathbf{H} = \mathbf{T}_e + \mathbf{T}_N + \mathbf{V}_{NN} + \mathbf{V}_{Ne} + \mathbf{V}_{ee}$). This is an eigenvalue problem, in which wave functions Ψ are the eigenfunctions of \mathbf{H} and E stands for the corresponding eigenvalues (energies). The main challenge in Quantum Chemistry is that we cannot solve exactly the Schrödinger equation, except for one-electron systems, due to the electron repulsion term present in the Hamiltonian. Quantum-chemical methods look for approximate solutions of the equation, employing computational numerical methods typically based on the variational principle and perturbation theory. A point worth bearing in mind is that none of these models is applicable under all circumstances. Actually, we should get the best method in order to find what it has been wisely defined as “the right answer for the right reason.”

The physics of electron correlation is hidden in the Hamiltonian itself. The Coulomb repulsion given by the term r^{-1} present in the \mathbf{V}_{ee} energy, the inverse distance between two electrons, increases enormously in the regions close to $r_{ij} = 0$, preventing that two electrons may occupy the same space. Therefore, the motion of any two electrons is not independent but it is correlated. The phenomenon is known as electron correlation. Moreover, the statement that two electrons are correlated is equivalent to express that the probability of finding two electrons at the same point in space is zero. The instantaneous position of electron i forms the center of a region that electron j will avoid. For this reason, it is stated that each electron, as described by the exact wave function Φ , is surrounded by a Coulomb hole. However, electron correlation is not taken into account properly by many approximate methods. The effect of neglecting electron correlation partly in approximate quantum-chemical approaches has great impact in the computed molecular spectroscopic properties of interest.

General Overview

We can group computational–chemical methods in three basic categories (see for instance Atkins and Friedman 1997; Helgaker et al. 2004):

- Ab initio methods, in which the complete Hamiltonian is used, all the integrals are solved numerically, and no essential parametrization is employed

- Semiempirical methods, in which a simpler Hamiltonian is used or integrals are adjusted to experimental values or *ab initio* results
- Molecular mechanics, in which we solve Newton's equation of motion, only valid for situations where no bonds are broken or formed, i.e., conformational changes

Obviously, the larger is the system under study, the less accurate is the available method. Despite their inherent drawbacks, classical semiempirical methods are still employed in large systems, whereas modern semiempirical methods, based on the Density Functional Theory, have a widespread use. A combined approach, QM/MM (Quantum Mechanics/Molecular Mechanics) treats an internal part of the problem with QM methods, whereas the surroundings or a large part of a macromolecule (for instance, a protein) is introduced using classical mechanics.

According to the number of configurations used to build the reference wave function, the *ab initio* methods can be classified in the following two categories (see Fig. 6):

- *Single-configuration methods.* They are typically based in the single Hartree–Fock (HF) reference, which determines the optimal ground-state energy and MOs (molecular orbitals). Post-HF methods introduce the electron correlation usually at the Configuration Interaction (CI), Coupled-Cluster (CC) or perturbative (PT) Møller–Plesset (MP, or PT in general) levels. The coupled-cluster methods with singly and doubly excited configurations, including the effect of triple excitations by perturbation theory CCSD(T), as well as related approaches, yield accurate results in well-defined ground-state situations and are considered as benchmark

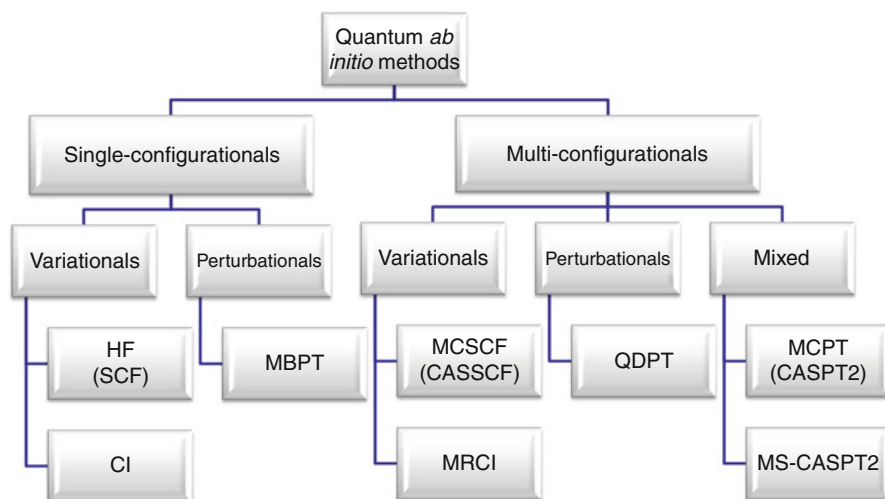


Fig. 6 *Ab initio* methods

results for small to medium molecules. In general, the applicability of the methods in this group is restricted to situations where a single reference wave function is adequate for the description of a chemical process, something not generally true for bond-breaking cases, degeneracies, and excited states

- *Multiconfigurational methods.* Part of the electronic correlation is already included in the reference wave function, normally by using a Multiconfigurational Self-Consistent-Field (MCSCF) wave function, which determines a set of MOs. The remaining electron correlation effects are accounted for by MRCI, MRCC, or MRPT techniques, where MR stands for multireference. They have a more ample range of applicability (ground state, excited states, transition states) than single-reference methods

The variation principle states that given a normalized wave function that satisfies the appropriate boundary conditions, then the expectation value of the Hamiltonian is an upper bound to the exact ground-state energy. In the linear variational problem, the trial function is a linear combination of basis functions, in general using the Linear Combination of Atomic Orbitals (LCAO) approach. On the other hand, in perturbation theory, the total Hamiltonian of the system is divided into two pieces: a zeroth-order part, which has known eigenfunctions and eigenvalues, and a perturbation part. The exact energy and wave function are then expressed as an infinite sum of contributions of increasing complexity. If we have chosen the zeroth-order Hamiltonian wisely, then the perturbation is small and the expansion (i.e., the sum of the 1st-, 2nd-, . . . , n th-order energies) converges quickly.

The simplest wave function to describe a many-electron system is a Slater determinant built by orthogonal one-electron wave functions. Electrons are fermions and accordingly they have to be described by an antisymmetric wave function. For an N -electron system, the Slater determinant has the form:

$$\Psi(x_1, x_2, \dots, x_N) = (N!)^{-1/2} \begin{vmatrix} \chi_i(x_1) & \chi_j(x_1) & \dots & \chi_k(x_1) \\ \chi_i(x_2) & \chi_j(x_2) & \dots & \chi_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(x_N) & \chi_j(x_N) & \dots & \chi_k(x_N) \end{vmatrix} = |\chi_i \chi_j \dots \chi_k| \quad (2)$$

without specifying which electron is in which orbital. To simplify the notation, a normalized Slater determinant is represented by only showing the diagonal elements of the determinant. The constant $(N!)^{-1/2}$ is a normalization factor. The wave function for an electron that describes both the spatial distribution and its spin is called spin orbital, $\chi_i(x_i)$. Since the Hamiltonian employed does not depend on the electronic spin, each spin orbital can be expressed by multiplying the spatial orbital, $\psi_j(r_i)$, by the spin function, ω (α – spin up and β – spin down): $\chi_i(x_i) = \psi_j(r_i) \cdot \omega$. A complete set for describing the spin of an electron consists of two orthogonal functions $\alpha(\omega_i)$ and $\beta(\omega_i)$.

A single-determinant wave function has several interesting properties. Firstly, it is worth noting that spin orbitals must be linearly independent, otherwise the value

of the determinant is zero. It is obvious that interchanging two rows of the Slater determinant, which is equivalent to interchanging the coordinates of two electrons, changes the sign of the determinant. The requirement of the antisymmetry principle is automatically fulfilled. Having two columns of the determinant identical, that is, two electrons occupying the same spin orbital, makes the determinant zero. Thus, no more than one electron can occupy a spin orbital (Pauli exclusion principle). When a linear transformation of the set $\{\chi_i\}$ is carried out, $\chi'_i = \sum_j \chi_j \cdot A_{ji}$, where A_{ji} is an element of the matrix \mathbf{A} of dimension $N \times N$, with a value for its determinant, $\det(\mathbf{A})$, different from zero, then $\Psi' = \det(\mathbf{A}) \Psi$. The wave functions Ψ' and Ψ differ just in a constant and, therefore, represent the same physical situation. Since the set of spin orbitals is linearly independent, we can always choose a transformation matrix \mathbf{A} so that the resulting spin orbitals χ'_i become orthonormal. Therefore, no restriction at all is imposed when we choose from the beginning an orthonormal set of spin orbitals. It just makes the computation of the Hamiltonian matrix elements involving Slater determinants easier. A Slater determinant is completely specified by the spin orbitals used to build it, and any unitary transformation of them is equally valid. Two sets of spin orbitals related by a unitary transformation ($\mathbf{A}^\dagger = \mathbf{A}^{-1}$), which keeps the orthonormality of the spin orbitals, yield the same Slater determinant. Slater determinants formed from orthonormal spin orbitals are normalized and N -electron Slater determinants that have different spin orbitals are orthogonal. In other words, a Slater determinant is completely specified by the spin orbitals used to build it and any unitary transformation of them is equally valid (each result of such a linear transformation represents the same physical situation).

There are two types of spin orbitals: restricted spin orbitals, which are constrained to have the same spatial function for α and β spin functions; and unrestricted spin orbitals, which have different spatial functions for α and β spins. A restricted set of spin orbitals has the form: $\chi_i(x) = \psi_j(r) \cdot \alpha(\omega) // \psi_j(r) \cdot \beta(\omega)$, whereas an unrestricted set has the form: $\chi_i(x) = \psi_j^\alpha(r) \cdot \alpha(\omega) // \psi_j^\beta(r) \cdot \beta(\omega)$.

Essentially, all practical calculations for generating solutions to the electronic Schrödinger equation are performed with molecular orbital methods: the zeroth-order wave function is constructed as one or two Slater determinants and the MOs are expanded in a set of atomic orbitals, the basis set. In a subsequent step, the wave function may be improved by adding electron correlation by either CI, MP, or CC methods. Nevertheless, there is another equivalent theory to get approximate solutions of this eigenvalue equation: the valence bond (VB) theory. The main drawback is that this theory leads to awkward calculations. On the positive side, conceptually it is much closer to the experimentalist's language since it may be regarded as a quantitative version of the resonance structure of electronic structure, with ionic structures and covalent structures of molecules. The VB method assumes that the wavefunction of a molecule may be written as a linear superposition of mathematical functions which represent canonical electronic structures: structures in which electrons are assigned to specific atoms and then paired (spinwise) leading to covalent and ionic structures, the combination of which represents the wave

function. The VB description of a bond as the result of two overlapping and localized orbitals is in contrast to the MO approach where a bond between two atoms arises as a sum over (small) contributions from many delocalized orbitals.

Classical valence bond theory is very successful in providing a qualitative explanation for many aspects. One of the great merits of VB theory is its pictorially institutive wave function that is expressed as a linear combination of chemically meaningful structures. It is this feature that has made VB theory so popular in the 1930–1950s. However, VB theory was “defeated” by MO theory for two main reasons: (1) mathematical simplicity and (2) several “failures” of the VB theory (actually due to misuse of very simplified versions of the theory). Currently, VB theory is coming of age in all the branches of theoretical chemistry (excited states, dynamics, environmental effects, and so on), with the development of faster, and more accurate *ab initio* VB methods, and with generation of new post-Pauling concepts. The Renaissance of VB theory is marked by surge in the following two-pronged activity: (1) creation of general qualitative models based on VB theory and (2) development of new methods and softwares that enable applications to moderate sized molecules.

Methods, Advantages, and Drawbacks

The Hartree–Fock method is the simplest *ab initio* approach. We can equate closed-shell HF theory to single determinant theory, and we are thus interested in finding a set of spin orbitals $\{\chi_a\}$ such that the single determinant formed from them, $|\Psi_0\rangle = |\chi_1\chi_2\chi_3\dots\chi_a\chi_b\dots\chi_N\rangle$ is the best possible approximation to the ground state of the N -electron system described by an electronic Hamiltonian \mathbf{H} . The expectation value of the energy, $E_0 = \langle\Psi_0|H|\Psi_0\rangle$, is a linear combination of one-electron integrals, $\langle i|h|j\rangle$ and two-electron integrals, $\langle ij|kl\rangle$. According to the properties of a Slater determinant, electrons are not uncorrelated in a HF wave function. At least the probability of finding two electrons with parallel spins at the same point in the space is zero, the so-called exchange correlation, which is incorporated by the antisymmetric condition of the wave function for fermions. The phenomenon is known as the Fermi hole. We are, therefore, facing a model of independent particles where the behavior of certain electrons is not fully independent because the Fermi hole simulates somehow the Coulomb hole. Consequently, only Fermi correlation is accounted for by the HF wave function.

By minimizing E_0 with respect to the choice of spin orbitals $\{\chi_a\}$, varying them with the only restriction that they remain orthonormal, $\langle\chi_a|\chi_b\rangle = \delta_{ab}$, one can arrive to the Hartree–Fock conditions. In doing so, one obtains an equation that defines the best spin orbitals, the ones that minimize E_0 . This equation for the best (Hartree–Fock) spin orbitals is the Hartree–Fock integro-differential equation $f|\chi_a\rangle = \varepsilon_a|\chi_a\rangle$, for the N -occupied spin orbitals. Each of the solutions $\{\chi_j\}$ has a spin orbital energy ε_j . The N spin orbitals with the lowest orbital energies are just the spin orbitals occupied in $|\Psi_0\rangle$ for which we use the indices a, b , etc. The remaining

infinite number of spin orbitals with higher energies are the virtual spin orbitals, which we label with the indices r, s , etc.

The Fock operator f is the sum of a core-Hamiltonian operator $h(1)$, which is the kinetic energy and potential energy for attraction to the nuclei of a single electron, and an effective one-electron potential called the Hartree–Fock potential $v^{\text{HF}}(1)$, which contains the Coulomb term (the total averaged potential acting on the electron in χ_a , arising from the $N - 1$ electrons in the other spin orbitals) and the exchange term (which arises because of the antisymmetric nature of the determinantal wave function). This is the essence of the HF approximation: to replace the complicated many-electron problem by a one-electron problem in which electron–electron repulsion is treated in an average way. What is more, since the Fock operator has a functional dependence, through the coulomb and exchange operators, on the solutions $\{\chi_a\}$ of the pseudo-eigenvalue equation, Hartree–Fock equations are actually nonlinear equations and will need to be solved by iterative procedures. The matrix representation of the Fock operator in the basis of spin orbitals eigenfunctions is diagonal with diagonal elements equal to the orbital energies.

The exact solution to this integro-differential equation corresponds to the “exact” HF spin orbitals. In practice, it is only possible to solve this equation exactly (i.e., as an integro-differential equation) for atoms. In practice, for molecules, the spin orbitals are expanded as a combination of one-electron atomic base functions and the set of matrix equations are solved. Only as the basis set approaches completeness, i.e., as one approaches the HF limit, will the spin orbitals be the exact HF orbitals.

It can be shown that any single determinant wave function $|\Psi_0\rangle$ formed from a set of spin orbitals $\{\chi_a\}$ retains a certain degree of flexibility in the spin orbitals; the spin orbitals can be mixed among themselves without changing the expectation value $E_0 = \langle \Psi_0 | H | \Psi_0 \rangle$. Furthermore, for a single determinant wave function, any expectation value is invariant to an arbitrary unitary transformation of the spin orbitals, since a transformed single determinant $|\Psi'_0\rangle$ can at most differ from the original determinant $|\Psi_0\rangle$ by a phase factor and, obviously, any observable property depends on $|\Psi|^2$, as we stated before. Thus the spin orbitals that make the total energy stationary are not unique, and no particular physical significance can be given to a particular set of spin orbitals. Localized spin orbitals, for instance, are not more “physical” than delocalized spin orbitals. Indeed, there exist a set of spin orbitals for which the eigenvalue matrix is diagonal, and we obtain just the equation $f | \chi_a \rangle = \varepsilon_a | \chi_a \rangle$, and this set is called “canonical spin orbitals.” The canonical spin orbitals, which are a solution to this equation, will generally be delocalized and form a basis for an irreproducible representation of the point group of the molecule.

It was in 1951 that Roothaan published his equations (see Roothaan 1951), considering molecular orbitals that were restricted to be linear combinations of a set of three-dimensional one-electron functions, ϕ_μ . Thus, ψ_i is a linear combination of these functions ϕ_μ (basis set): $\psi_i = \sum C_{\mu i} \cdot \phi_\mu$. If the set $\{\phi_\mu\}$ was complete, this would be an exact expansion, and any complete set $\{\phi_\mu\}$ could be used. Unfortunately, this one is always restricted, for practical computational reasons,

to a finite set of N basis functions. As the basis set becomes more and more complete, this expansion leads to more and more accurate representations of the “exact” molecular orbitals, which are eigenfunctions of the Fock operator. For any finite basis set, we will obtain molecular orbitals which are exact only in the space spanned by the aforesaid finite basis. In addition, since the electronic energy (the output of the quantum mechanical calculation) is variational, the better the basis set, the lower the total energy. As the basis set becomes more and more complete, the total energy approaches the Hartree–Fock limit. Of course, by the variational principle, the Hartree–Fock-limit energy is still above the “exact” energy, which here can be taken as the energy obtained from an exact solution to the nonrelativistic Schrödinger equation in the Born–Oppenheimer approximation.

Variation of the total energy was then carried out respect to the coefficients $C_{\mu i}$ of such a linear combination. This leads to a set of algebraic equations which can be written in matrix form, $\mathbf{FC} = \mathbf{SCE}$, where \mathbf{F} is the matrix representation of the Fock operator (which can be divided into the core-Hamiltonian matrix, that is, integrals involving the one-electron operator $h(i)$ describing the kinetic energy and nuclear attraction of the electron i , plus the two-electron terms), \mathbf{C} is the matrix of coefficients, \mathbf{S} is the overlap matrix, and \mathbf{E} is the diagonal energy matrix. The columns of \mathbf{C} describe the molecular orbitals, i.e., the coefficients describing ψ_1 are in the first column, with the corresponding energy ε_1 , those describing ψ_2 are in the second column, and so on. If \mathbf{S} is the unit matrix (i.e., if we have an orthonormal basis set), then we would have $\mathbf{FC} = \mathbf{CE}$, and Roothaan’s equations would just have the form of the usual matrix eigenvalue problem. Increasing the flexibility of the one-electron basis set $\{\phi_{\mu}\}$, the HF energy E_0 will progressively reach the Hartree–Fock limit.

The approximation has its limitations, in particular the lack of electronic correlation energy. For instance, Restricted Hartree–Fock (RHF) solution predicts incorrectly the dissociation of molecules into open-shell fragments (like $\text{H}_2 \rightarrow 2\text{H}$) given that both electrons are forced to occupy the same spatial molecular orbital (MO), when they should belong to infinitely separated atoms in the dissociated solution. This means that some electron–electron repulsion remains even at infinity, which is the source of a spurious term in the total energy (different from twice the energy of two hydrogen atoms). Alternatively, the products of dissociation are not just $2\text{H}\cdot$, but also include, incorrectly, H^- and H^+ . However, the poor behavior of closed-shell RHF calculations upon dissociation to open-shell products does not detract from their utility as reference function in the region of equilibrium: the potential surface obtained from a closed-shell RHF calculation will not be parallel in regions of the surface characterizing dissociation limit, but it will be reasonably parallel near the region of the equilibrium geometry. It is worth mentioning that the VB wave function for H_2 , $\Psi = C_{\text{cov}}\Psi_{\text{cov}} + C_{\text{ion}}\Psi_{\text{ion}}$, with the covalent and ionic counterparts, contains “left-right” correlation which is necessary for a correct description of dissociation of the H_2 molecule.

Within the HF approach, a slight improvement can be achieved in dealing with open-shell systems with the restricted open-shell (ROHF) and the unrestricted Hartree–Fock (UHF) procedures. In the former, all electrons, except those that

are explicitly required to occupy open-shell orbitals, occupy closed-shell MOs. In the latter, electrons of different spin, in general, are described by different sets of spatial orbitals, although the obtained wave function is typically contaminated with higher spin state solutions. The approximation leads to two integro-differential eigenvalue equations which are coupled and cannot be solved independently. The introduction of a basis leads to the Pople–Nesbet equations, $\mathbf{F}^\alpha \mathbf{C}^\alpha = \mathbf{S}^\alpha \mathbf{C}^\alpha \mathbf{E}^\alpha$ and $\mathbf{F}^\beta \mathbf{C}^\beta = \mathbf{S}^\beta \mathbf{C}^\beta \mathbf{E}^\beta$. $\mathbf{N}^\alpha = \mathbf{N}^\beta$, a restricted solution to the Roothaan equations is a solution to the unrestricted Pople–Nesbet equations. In general those approaches, which can serve as reference for post-HF methods in ground-state cases, are not accurate enough to compute excited states.

The lack of correlation is the actual source of all errors. In particular, a Slater determinant incorporates exchange correlation, i.e., the motion of two electrons with parallel spins is correlated (the so-called Fermi correlation). Unfortunately, the motion of electrons with opposite spins remains uncorrelated. It is common to define correlation energy, E_{corr} , as the difference between the exact nonrelativistic energy of the system, ε_0 , and the Hartree–Fock energy, E_0 , obtained in the limit that the basis set approaches completeness: $E_{\text{corr}} = \varepsilon_0 - E_0$. The simplest manner to understand the inclusion of the correlation effects is through the method of configuration interaction (CI). The basic idea is to diagonalize the N-electron Hamiltonian in a basis of N-electron functions: we represent the exact wave function as a linear combination of N-electron trial functions and use the linear variational method.

$$|\Phi\rangle = C_0 |\Psi_0\rangle + \sum_{ra} C_a^r |\Psi_a^r\rangle + \sum_{a<b;r<s} C_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \dots \quad (3)$$

If the basis were complete, we would obtain the exact energies of all the electronic states of the system. In spite of providing the exact solution of a many-electron problem, we can handle only a finite set of N-electron trial functions. As a result, the CI method provides only upper bounds to the exact energies. Specifically, the lowest eigenvalue, ε_0 , will be an upper bound to the ground-state energy of the system. When all the N-electron wave functions are taken into account, the calculation is named full configuration interaction (FCI) and the corresponding eigenvalues and eigenvectors computed are exact within the space spanned by the finite basis set. Despite the great advances in FCI technology in the last few years, the size of the eigenvalue problem becomes rapidly too large to be handled by modern computers. As a result, FCI solutions are only available for very small molecular systems. In contrast to HF, the FCI energy of H_2 properly describes the dissociation. However, within a minimal basis set approach, the FCI potential curve does not agree very well with the exact one provided by Kolos and Wolniewicz due to the lack of flexibility of such a basis set. The truncation of both N-electron basis and one-electron basis is the main source of inaccuracies in quantum-chemical calculations (Fig. 7).

As mentioned earlier, for computational reasons, we have to truncate the full CI matrix or equivalently the CI expansion for the wave function, considering only those configurations which differ from the HF ground-state wave function

$$\begin{aligned}
 & \text{N-electron basis set} \\
 \langle \Phi \rangle &= C_0 |\Psi_0\rangle + \sum_{ra} C_a^r |\Psi_a^r\rangle + \sum_{a<b; r<s} C_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \dots \\
 & \text{1-electron basis set} \\
 \langle \Phi_i \rangle &= |\chi_i \chi_j \dots \chi_N\rangle \quad \chi_i = \phi_i \vec{\omega} / \vec{\omega} = \alpha, \beta \\
 \phi_i(\vec{r}) &= \sum_{\mu=1}^K C_{\mu i} \phi_{\mu}(\vec{r}) \quad \phi_{\mu}(\vec{r}) = \sum_K d_{K\mu} g_K
 \end{aligned}$$

Fig. 7 Many-electron expansion (CI) and one-electron expansion (basis set). The total wave function, Φ , is a linear combination of N-electron wave functions Ψ_0, Ψ_a^r , etc. Each one of these functions is an antisymmetrized and normalized product of spin orbitals, χ_l . Each of them is constituted by a one-electron wave function, ψ_l , and a spin function, ω . Each one-electron wave function is defined as a linear combination of a set of basis functions, ϕ_{μ} , which are used to be contracted gaussian functions, CGTFs (linear combinations of a set of primitive functions, g_{κ})

by no more than a predetermined number of spin orbitals. The simplest version of this scheme is to deal with single and double excitation out of $|\Psi_0\rangle$: a SDCI calculation. SDCI and, in fact, all forms of truncated CI deteriorate as the number of electrons increases. The truncation of the CI expansion leads to problems of size-extensivity. For instance, in a DCI calculation, by definition the wave function of each monomer contains double excitations within the monomer. If we restrict the supermolecule trial function to double excitations, we exclude the possibility that both monomers are simultaneously doubly excited, since this represents a quadruple excitation in the supermolecule. It is now common to differentiate size-extensivity and size-consistency. The latter was originally employed by Pople as a criterion for a well-constructed quantum-chemical method, indicating that the energy computed for two noninteracting molecules should be exactly twice that calculated for only one isolated molecule, that is, a property that described the additive separability of the wave function. Recently, the concept has been extended to include not only the fragmentation limit but the entire process, that is, qualitatively all regions of the potential energy hypersurface must be qualitatively correct. For instance, both RHF and UHF wave functions are size consistent in the first sense, considering that they properly describe a separated H_2 dimer system, but for a closed-shell molecule dissociating into open-shell fragments the RHF descriptions fails, and therefore does not conform to the broader size-consistency concept. On the other hand, size extensivity, analogously to the extensivity concept in thermodynamics, refers to the correct, linear, scaling of a method with the number of electrons, and its fulfillment leads to methods in which calculations with differing number of electrons can be compared, like those related to ionization processes. Otherwise, the error of the method increases with the size of the system. At the noninteracting limit size extensivity implies size consistency, but this latter property has the requirement of correct fragmentation, which does not depend on the mathematical scaling. Therefore, RHF and UHF approaches are always size extensive, but RHF does not provide

proper dissociation in open-shell cases, and it is therefore non-size-consistent, as any other method such as single-reference coupled-cluster (CC) or perturbation approaches (PT) which use RHF as reference wave function. On the other hand any truncated configuration interaction approximation is not size extensive, whereas CC and PT are. In the case of the multiconfigurational methods, the size-consistency depends on the selected reference – generally they are size consistent – and the size extensivity varies with the approach. MRCI is formally a truncated CI approach and therefore is not size extensive, as CASPT2 – multiconfigurational complete active space perturbation theory up to second order – and similar methods, which are formally not size extensive either, but in practice the effects are irrelevant, in particular for the most common implementation (see Taylor 2007).

It is worth mentioning that using a better-variational method only ensures to get a better variational energy, but a priori we cannot say nothing regarding properties other than the energy. For instance, the dipole moment of CO is ca. zero, and to obtain the correct sign of the vector (actually, greater than zero: +0.044 a.u. experimentally) is a tricky problem. A DCI calculation with the 200 most important double excitations has a lower energy than a SDCI procedure with 138 doubles + 62 singles. Still, the DCI calculation predicts the wrong sign of the dipole moment (−0.072 a.u.), just the contrary that the SDCI calculation (+0.030 a.u.). To improve the result it is necessary to employ CCSD(T) (+0.047 a.u.) (Paschoal et al. 2009).

There is another clever way to introduce in the wave function enough correlation effects to get a good account of the nature of the chemical problem, for instance, in a dissociation path of a diatomic molecule. Such procedure implies to improve the reference wave function by including more than one determinant or electronic configuration in its description. Adding more than one configuration means to get their weights in a linear combination. Therefore, the central idea of a multiconfigurational self-consistent field (MCSCF) calculation is to build the wave function as a truncated CI expansion

$$|\Psi_{\text{MCSCF}}\rangle = \sum_I c_I |\Psi_I\rangle, \quad (4)$$

in which both the expansion coefficients and the orthonormal orbitals contained in $|\Psi_I\rangle$ are optimized simultaneously. For a closed-shell system, if only one determinant is included in such an expansion, the MCSCF and HF methods become identical. In the minimal basis set description of H_2 , Ψ_{MCSCF} is identical to the FCI wave function. However, if an extended basis set is used, the MCSFC energy will be above and nearly parallel to the FCI energy but below the energy obtained from any two-configuration CI expansion based on canonical HF MOs.

The correlation energy introduced in a MCSCF wave function is usually named nondynamic, static, or large-range correlation energy, and the corresponding wave function must be build in such way that includes most of near-degenerate electronic configurations, that is, all those which basically contribute to define the reference function along the chemical process under study. An example is a bond breaking. A proper MCSCF description should include both the bonding and antibonding dis-

tribution of the electrons, and the corresponding MOs. The optimization procedure at each region of the PEH will determine the relative weight of each configuration at that point and the proper MOs. Improving the reference in such manner helps to the post-HF approaches to account for the remaining correlation effects, named dynamic or short-range. It is like starting the race from a more advantageous position.

One of the most successful and systematic procedures to account for the correlation energy, which is not variational but is size extensive at each order, is the perturbation theory. Using different zeroth-order Hamiltonians, Møller–Plesset perturbation theory (MPPT), also named many-body perturbation theory (MBPT), yields the energy and the wave function corrected up to n order following a Taylor series expansion. Although, in general, higher perturbation orders improve the result, it has been proved that it leads to divergences at infinite order. Second- and fourth-order perturbation (MP2, MP4) results are much better than the SCF and SDCI ones. The most recent family of methods based on a single reference is that based on the coupled-cluster (CC) approach, in which the correlation energy is introduced size-extensively by increasing the excitation level. CC methods for the ground state, especially when including up to triple excitations, have become, in practice, the most accurate quantum-chemical approaches for many systems.

No matter how good is the quality of the post-HF method for the inclusion of the correlation energy, when based on a single HF reference, the accuracy of the results largely decreases in many situations, like those related with dissociations, degeneracies, and excited states. For instance, regarding the dissociation problem, the potential curve calculated at any level based on a RHF reference is not satisfactory at long range. The problem is the starting point: the exact wave function at large R cannot be described by a single determinant. The results largely improve its accuracy when a multiconfigurational or multireference (MR) wave function is used as a reference, for instance, in the MRCI, MRCC, or MRMP approaches. Recovering correlation energy does not require so large effort in this case, for instance, lower perturbation (MP) or excitation (CC) levels. As a drawback, the complexity of the methods increases.

On the other hand, the semiempirical DFT approach has become the most widespread quantum-chemical method in the last years. It allows us to treat large systems in a relatively short time taking into account electron correlation. The basic idea is that the energy of an electronic system can be written in terms of the electron density, ρ . For a system of N electrons, $\rho(\mathbf{r})$ denotes the total electron density at a particular point in space \mathbf{r} . The electronic energy E is said to be a functional of the electron density, denoted $E[\rho]$, in the sense that for a given function $\rho(\mathbf{r})$, there is a single corresponding energy. It is not necessary to compute a wave function, but, unfortunately, the Hohenberg–Kohn theorem does not inform about the form of the functional dependence of energy on the density: it confirms only that such a functional exists. So the procedures should resort to approximate derivations with adjustable parameters. This is the reason why functionals proliferate to hundreds in the literature. Some of them, like the Becke3–Lee–Yang–Parr (B3LYP) functional, is the most popular in the literature, not because its overall accuracy but of its balance

and its controlled flaws. There is no systematic way to improve a functional, just to derive another one with different parameters. It is typical to use one functional per each situation, something that makes the predictability of the results somewhat questionable. On the contrary, *ab initio* calculations, if allowed by the size of the problem, can be hierarchically improved in a known direction enlarging the N -electron and/or the one-electron basis.

Basis Sets

The choice of the one-particle space is a most important decision when setting up any calculation, and there is no point in trying to improve the result if the selection of the one-electron basis set is not adequate. This is especially true for the calculation of excited states, in which states of very different nature (for instance, compact and diffuse) have particular requirements that must be fulfilled simultaneously when selecting the basis set. In a strict mathematical sense, many different types of basis set functions $\{\phi_\mu\}$ can be used. Gaussian-type functions (GTFs, $g \propto e^{-\alpha r^2}$) are the more widespread in spite of being ill-behaved in both close and far away of the nuclei. Nevertheless, four-center integrals (that is, two-electron repulsion integrals) are very easy to evaluate with GTFs, and, in addition, the basis set can be improved by employing contracted Gaussian functions, CGTFs, which are linear combinations of primitive GTFs. To build a CGTFs, the primitive functions have to be optimized, both exponents and contraction coefficients. According to the number of basis functions, it is common to distinguish among minimal basis set (a single basis function per occupied atomic orbital), double- ζ (two basis functions per atomic orbital), triple- ζ , split-valence (duplication only affects to valence layer whereas minimal basis set is used for core orbitals), etc.

Ground-state calculations have made extensive use of the so-called Pople basis sets: STO-G, 3-21G, 6-31G. These basis sets are built using segmented contraction, which means that each primitive functions contributes to one or a limited number of contracted function only. Exponents and coefficients were optimized using the HF approach. Valence basis sets are too limited to get accurate results, and therefore additional functions have to be added, named polarization functions (6-31G* or 6-31G(d), which means that polarization functions are added to second-period atoms; and 6-31G** or 6-31G(d,p), which means that, additionally, polarization functions are added to hydrogen atoms). These are functions of higher angular momentum than those of valence type and are necessary in order to describe the changes of the electronic density of an atom in the molecule. Diffuse functions (6-31+G, 6-31++G) can be also added which are more extended in the space, that is, with smaller exponents, to describe better situations such as anions or Rydberg states. A 6-31G basis set for the C atom ($1s^2 2s^2 2p^2$) is described as (10s4p)/[3s2p], i.e., 10 primitives s and 4 primitives p (22 primitives: $1s \times 6 + 2s \times 3 + 2s \times 1 + 2p \times 3 + 2p \times 1$) are contracted to three combined functions s and two combined functions p (nine contracted functions: $1s, 2s, 2s'$,

2p, 2p'). The basic flaw of Pople's basis sets is that they have been built with a method lacking correlation energy. The consequence is that they are very poor in recovering correlation effects in post-HF methods. In order to get better values, more polarization of diffuse functions have to be added to compensate the lack of flexibility of these basis set.

If the ultimate goal is to perform correlated calculations, it would seem preferable to include correlation in the construction of the basis set. For correlated calculations, the basis set requirements are different and more demanding since we must also provide a virtual orbital space capable of recovering a large part of the correlation energy. It is often sufficient to correlate only the valence electrons. This is the case for correlation consistent (cc) or Atomic Natural Orbitals (ANO) basis sets. Correlation consistent basis sets are designed so that functions which contribute in similar amounts of correlation energy are included at the same stage, independently of the function type. We can add polarization functions (cc-pVDZ) or diffuse functions (augmented, aug). They need fewer primitives than ANOs, and each contracted basis comes from a different primitive set and the exponents of polarization functions are optimized by correlated calculations within a segmented contraction scheme.

On the other hand, ANO-type basis sets arise as eigenfunctions of the first-order reduced density matrix of the atom. They give the most rapidly convergent CI expansion, and to obtain a given accuracy one requires fewer configurations formed from NOs than configurations formed from any other orthonormal basis. The occupation number is a reference of its significance in the wave function and the truncation of the orbital space by elimination of the ANOs whose occupation numbers are small produces the least possible error on the wave function. The coefficients of ANOs obtained in a correlated calculation of the ground state of the atom represent the coefficients of the basis functions within a general contraction scheme (all primitives on a given atom and of a given angular momentum enter all the contracted functions having that angular momentum, but with different contraction coefficients, which improves the flexibility). This yields basis of any size with a unique calculation with the primitive set, and since the same primitive set for all contractions is used, a smaller basis is always a subset of a larger basis. In other words, we can enlarge our basis set without changing of vector space (Almlöf and Taylor 1986, 1991). An improvement of the original scheme takes into account an average density matrix to build the basis set (including ground state, excited state, cation, anion, atom into an electric field, see Widmark et al. (1990)). In general, ANO basis set will require less contracted functions than the other basis to get the same results. Recently, all electron ANO basis sets, including correlation effects for all periodic system elements, have been generated (Roos et al. 2005). It has also been observed that for extremely accurate calculations, only ANO basis sets can provide the best answer (see, e.g., Martin et al. 1997).

A final type of approach regarding the one-electron basis sets is the combination of inner-shell pseudopotentials (including typically relativistic effects) with valence basis sets and the use of embedded potentials to represent atomic environments.

Methods for Excited States

Based on the methods described in the previous section, specific algorithms to deal with electronic excited states have been developed. In this section, the most commonly employed approaches will be briefly summarized (see also Merchán and Serrano-Andrés 2005; Serrano-Andrés and Merchán 2005).

Starting by the single-configurational methods that use the HF solution as reference wave function, an approach which has lost popularity because of its poor performance is the method named Configuration Interaction-Singles (CIS). The essence of the method is to consider that an excited state can be described by a singly excited determinant formed by replacing, with respect to the reference wave function, an occupied spin orbital with a virtual spin orbital. The drawbacks of such a description may be partially compensated if a linear combination of all possible single excited determinants is used to build the excited state wave function. In general, the CIS excitation energies are largely overestimated due to the absence of correlation energy effects. It is more common to find in the literature cases in which CIS yields the wrong order and state nature than the proper ones, simply because the differential correlation energy affects the excited states unevenly and because the intrinsic character of the states is multiconfigurational (Foresman et al. 1992).

In the light of the previous discussion, the logical method to deal with excited states should be multireference CI (MRCI) (Buenker et al. 2000). Consider a wave function of CI type expanded in a many-electron basis set of determinants. As in the H_2 molecule, it is possible to select a number of determinants to describe the correct dissociation limit. When the energy is minimized with respect to the coefficients of the expansion, the configuration interaction (CI) method is employed. It should be kept in mind that actual calculations are performed using either spin-adapted Configuration State Functions (CSFs: appropriate linear combination of Slater determinants) or determinants as a N -electronic basis set (Slater determinants lead to more efficient CI algorithms), whereas CSFs lead to shorter CI expansions and, obviously, spin eigenfunctions, i.e., singlet, doublet, triplet states . . .). In case that the expansion contains more than one configuration, the process is denoted as MRCI. The wave function $|\Psi\rangle$ is a multireference function and, at least, the singly and doubly excited determinants generated from each reference determinant $|m\rangle$ are taken into account. Unfortunately, MRCI-based approaches are restricted to molecular systems of small molecular size (containing up to 6–7 atoms), where dynamic correlation can be fully retrieved. Otherwise, truncation of the many-electron functions involved in the MRCI expansion might easily lead to large errors in the computed excitation energies.

Another type of methods applied to the calculation of excited states are the propagator approaches (Oddershede 1987). The underlying technique, also called Green's function approach, equation-of-motion or linear response theory in its different forms, can be applied to various types of methodologies: single- or multi-reference configuration interaction, coupled-cluster, or density functional.

The essentials of the technique consider that once a molecule is subjected to a linear time-dependent electric field fluctuating with frequency ω , a second-order property as the frequency-dependent ground-state polarizability of the system is well approximated by

$$\alpha_{\omega} = \sum_{i \neq 0}^{\text{states}} \frac{|\langle \Psi_0 | r | \Psi_i \rangle|^2}{\omega - \Delta E_i}, \quad (5)$$

where the denominator of the expression involves the frequency of the field and the excitation energies (ΔE_i) characterizing the excited states (Eq. 1), while the numerator of each term is the square of the transition dipole moment between the ground and the corresponding excited state. Using complex function analysis, it is possible to obtain the poles of the expression, that is, the values for which the frequency corresponds to the excitation energies and the denominator goes to zero, while the residues provide the numerators, in this case, the one-photon absorption matrix elements. The peculiarity of the propagator approaches is that the wave functions of the individual states are not necessarily computed to obtain excitation energies and transition probabilities, while its quality relies on the type of reference wave function. Except for very elaborated implementations, usual propagator approaches have also the same problems as all single-reference procedures.

The family of methods CCS, CC2, and CC3 is based on response theory (Koch and Jørgensen 1990). The CCS approach is equivalent to the single excited configuration interaction or Tamm–Dancoff approach. The iterative hybrid CC2 and CC3 procedures introduce approximations to account for the level of excitation. They have been defined for systems with closed-shell ground-states, although some Equation of Motion EOM-CC procedures also deal with open-shell ground state cases. In order to get accurate excitation energies and properties, the single-configuration coupled-cluster methods should include high excitation levels to compensate both the poor reference wave function and the multiconfigurational character of the excited states. CC-based methods are, up-to-date and in practice, the most accurate methods to compute excited states in small to medium size molecules with closed-shell ground states, but only for those states which are well described by singly excited configurations, in systems where the ground state has a clear single-configuration character, and close to the equilibrium geometry. Triple excitations have to be anyway included in the cluster expansion if accuracy is intended. The precision of the single-reference CC methods decreases in systems with open-shell ground states and vanishes (up to several eVs of error) for multiconfigurational cases, like the $2^1 A_1$ state of ozone. The inclusion of quadruple excitations, unpractical so far, would improve some of those results, although the only solution in prospect to beat in accuracy the lower level and less expensive multireference perturbation approaches such as CASPT2, is to use multireference coupled-cluster (MRCC) methods, in which the required excitation level will be expected lower. However, the development of such methods is still in their infancy, and the initial results are not very promising.

Before continuing with the *ab initio* procedures, in particular the MRMP theory most employed in this chapter, the use of DFT approaches for excited states will be discussed. The implementation, known with the unfortunate name of Time-Dependent DFT (TD-DFT) approach (no time-dependency is accounted for) should be expected to be able to deal with large systems were *ab initio* methods become too expensive. Unfortunately, these methods fail dramatically in too many situations: charge transfer states, multiconfigurational states, doubly or highly-excited states, and even introduce large and systematic errors in valence states of large π extended systems. In some cases, the deviations can be as large as 5–6 eV, especially in cases in which the ground state is poorly defined by the HF configuration. Even when many different parameterizations have been tried and different functionals developed in order to correct the flaws of the method, so far no single functional is able to solve most of them simultaneously. Much worse than that (and this is common for all single-configurational approaches), they cannot describe at all degenerate situations like conical intersections, which are the core of the quantum-chemical description of the excited states.

An hybrid derivation, known as DFT/MRCI is unarguably the best DFT-based procedure for excited states. It consists of an MRCI expansion with multiconfigurational wave functions replacing the HF orbitals with Kohn–Sham orbitals in the building of CSFs. The Hamiltonian, however, is heavily modified with empirical parameters, and only a few situations can be handled. The objective is to recover dynamic correlation by means of DFT and static correlation by means of MRCI. In this way, severe size-extensivity problems can be avoided even for systems with many valence electrons. Currently, optimized parameter sets for the effective DFT/MRCI Hamiltonian are available in combination with the B3LYP functional. The accuracy in energy is about 0.2–0.3 eV, and other properties are well balanced as well. Additionally, it is very appropriate to compute SOC, including a spin-dependent Hamiltonian (Grimme and Waletzke 1999; Kleinschmidt et al. 2001).

As already mentioned, a multiconfigurational description of the reference wave function helps enormously to recover the overall correlation effects. Starting from a good reference, the method used for such a recovery does not need to be extremely elaborated. That is why multiconfigurational second-order Møller–Plesset theory, the CASPT2 method, based on a Complete Active Space SCF (CASSCF) zeroth-order wave function, has become the most successful method for excited states. As this is the procedure in which most of the examples of the chapter will be based on, a more detailed description of the procedure will be given. Nowadays, the CASPT2//CASSCF methodology has proved the best ratio quality of the results/computational cost. At the CASSCF level (a particular case of MCSCF), both the many-electron-function coefficients of the MCSCF expansion and the coefficients included in the expansion of each molecular orbital are optimized simultaneously (see Eq. 3). Their variations are considered as rotations in an orthonormalized vector space. In the CASSCF method, the orbitals are classified in three categories, depending on the role they play in building the N -electron wave function: inactive, active, and secondary orbitals. Inactive and active orbitals are occupied in the wave functions, whereas the remaining of the orbital space, given by

the size of the one-electron basis set employed, is constituted by secondary orbitals, also called external or virtual. Inactive orbitals are doubly occupied in all the CASSCF configurations. The rest of the electrons (called active electrons) occupy active orbitals. The CASSCF wave function is formed by a linear combination of all the possible configurations that can be built by distributing the active electrons among the active orbitals and are consistent with a given spatial and spin symmetry. That is, in the configuration space spanned by the active orbitals, the CASSCF function is complete (CAS-CI space, equivalent to FCI). Inactive orbitals are also optimized in the variational process, but they are treated as in the restricted HF function. The CASSCF energy is invariant to rotations among the active orbitals (Roos et al. 1979).

Essentially, we construct for a given state a multiconfigurational wave function which includes all configurations generated by a set of active orbitals and active electrons that fulfils spin and spatial symmetry requirements. This provides the nondynamic correlation effects due to configurations which are very close in energy. Several states that belong to a same symmetry are usually computed by means of a State-Average (SA) CASSCF calculation, where a functional of energy is defined as average of a number of states, that, if required, although it is not recommended, can be weighted. From a SA-CASSCF calculation, a single set of average orbitals and a number of orthogonal wave functions equal to the number of roots used in the average process are obtained. In this manner, it is sometimes possible to overcome the problem of “root flipping,” that is, the interchange of roots along the CASSCF optimization procedure. For a given spatial and spin symmetry, the treatment of excited states is preferably performed by using SA-CASSCF calculations. In principle, it is also possible to make a single CASSCF calculation for higher roots, optimizing just one state. Nevertheless, experience shows that in most cases it can only be achieved for the second root of a given irreducible representation. Wave functions obtained in a SA-CASSCF calculation are orthogonal among them, whereas those obtained from different CASSCF calculations are not. In the latter cases, the wave function is rather poor and cannot be used for a further correlated calculations, although sometimes it may be useful to perform CASSCF optimizations.

On the other hand, the Restricted Active Space Self-Consistent Field (RASSCF) method is a more general extension of the CASSCF method. Now, there are three subspaces within the active orbitals: RAS1 (orbitals that are doubly occupied except for a maximum number of holes allowed in this orbital subspace), RAS2 (in these orbitals, all possible occupations are allowed), and RAS3 (orbitals that are unoccupied except for a maximum number of electrons allowed in this subspace). CASSCF calculations can be performed by allowing orbitals only in the RAS2 space. A single reference SDCI wave function is obtained by allowing a maximum of two holes in RAS1 and a maximum of two electrons in RAS3, while RAS2 is empty (Malmqvist et al. 1990; Olsen et al. 1988; Fig. 8).

Either by using CASSCF or RASSCF, the active space provided by the user of a CASSCF calculation represents a key point to obtain accurate theoretical predictions once dynamic correlation has subsequently been taken into account,

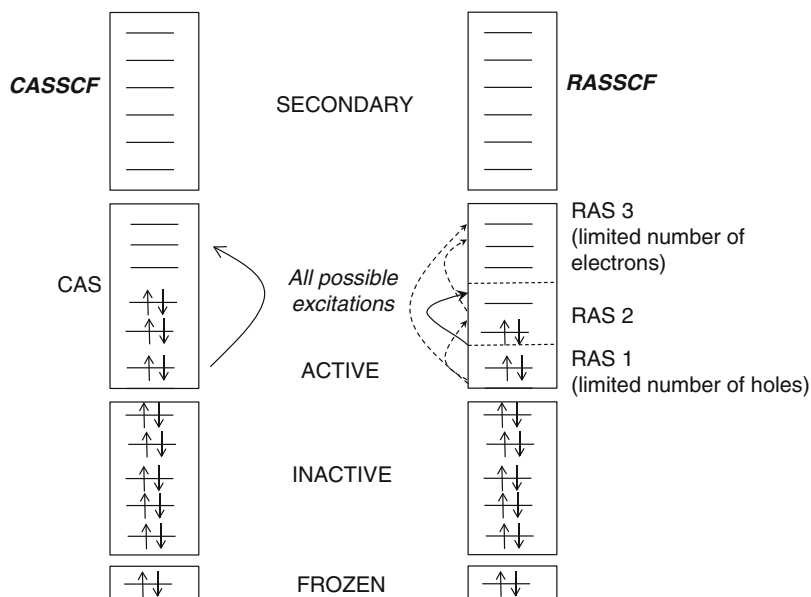


Fig. 8 CASSCF and RASSCF methods

for instance, at the CASPT2 or RASPT2 levels. The properties of a CASSCF wave function depend on the active space. Thus, a valence CASSCF is size-extensive (the computed energy within a given methods scales linearly with the number of particles) and the corresponding CASPT2 results become also nearly size-extensive (formally the CASPT2 method is not size-extensive, but it is in practice, in particular the MOLCAS implementation). As in any quantum-chemical approach, one has to make sure that the method has enough flexibility, i.e., the active space is the appropriate one, to describe the chemical process under consideration. It is important to mention here the power of the CAS State Interaction (CASSI) approach, which provides orthogonal wave functions and transition densities from CASSCF or RASSCF wave functions optimized independently for a number of excited states of whatever symmetry (Malmqvist and Roos 1989).

Valence-bond methods have increased its applicability recently. One example is the CASVB (complete active space valence bond) method. A CASVB wave function can be obtained simply by transforming a canonical CASSCF function and readily interpreted in terms of the well-known classical VB resonance structures. The total CASVB wave function is identical to the canonical CASSCF wave function. In other words, the MO description and the VB description are equivalent, at least at the level of CASSCF. The CASVB method provides an alternative tool for describing the correlated wave functions.

The CASPT2 (complete active space perturbation theory to second order) method (Andersson et al. 1992) includes the remaining dynamic correlation due to short-range electronic interactions. This method can be seen as a conventional non-

degenerate perturbation theory, that is, a single state is independently considered, with the particularity that this zeroth-order wavefunction is multiconfigurational (CASSCF). The wave function is corrected up to first order and, consequently, the energy is corrected up to second order. The set of functions required to compute the first-order correction of the wave function is formed by those that interact with the zeroth-order wave function through the Hamiltonian in the Rayleigh–Schrödinger perturbation theory, and it is known as the first-order interacting space. Taking into account the one and two particle nature of the Hamiltonian, the first-order interaction space, called hereafter V_{SD} , comprises the functions generated by singly and doubly excited configurations from the zeroth-order (CASSCF) wave function. As a matter of fact, the configurational space is divided in four subspaces: V_0 (one-dimension space expanded by the reference function $|0\rangle$ of the studied states), V_K (orthogonal to $|0\rangle$ in the restricted FCI subspace used to generate the CAS function), the aforesaid V_{SD} (space expanded by the singles and doubles replacements from $|0\rangle$), and $V_{TQ} \dots$ (space that contains the excitations of higher order). Regarding orbitals, they are classified into frozen (doubly occupied, not correlated), inactives (doubly occupied in the reference function), actives (with any occupation, between 0 and 2, in the reference function) and secondary (empty in the reference function). Recently, a new shifted zeroth-order Hamiltonian named IPEA has been designed and set up with a value 0.25 a.u. into the default for the MOLCAS implementation (Ghigo et al. 2004). The new formulation solves previous effects for open-shell cases in which the correlation effects were overshoot. The net effect of this correction is to slightly increase the excitation energies. In any case care has to be taken when comparing present with previous CASPT2 results.

The normalized wave function is corrected up to first order. The weight of the reference function, C_0^2 , can be used as a simple and rapid criterion of quality for the perturbation treatment carried out. Ideally, in order to get a fast convergence in the perturbation series, the weight should be close to unity. Nevertheless, its value depends on the number of correlated electrons. Thus, upon enlarging the molecular system the reference weight decreases. The electronic excited states considered should have a similar magnitude for the weight as compared to the ground state, employing the same active space. Sometimes intruder states appear in the second-order calculation, which are normally related to the occurrence of large coefficients in the first-order expansion, leading to a low value for the reference weight. Analysis of the states with large coefficients (intruder states) may give a hint about the type of reformulation in the perturbation partition necessary to overcome the problem. Thus, a new CASSCF calculation might be designed comprising in the active space the MOs implied in the description of the previous intruder states. It is the proper action to be taken when intruder states are strongly interacting with the CASSCF reference wave function, because it points out to obvious deficiencies in the choice of the active space. Intruder states are often present in the treatment of excited states of small organic compounds when the active space does not include the full π valence system. Thus, the low weight for the zeroth-order wave function in such a case just tells us that the active space has to be enlarged in a way that previous intruder states would be treated variationally, that is, they should be moved to the

active space. It is also frequent to find calculations where the reference weight of the excited state is “somewhat low” compared to that of the ground state, but a particular state cannot be identified as intruder in the first-order wave function, which is instead characterized by a large number of low-energy minor contributions. It occurs often in the simultaneous computation of valence and Rydberg states, where the one-electron valence basis set has been augmented with Rydberg-type functions. This is a typical case when using large and diffuse basis sets like ANOs. We have to face then accidental near-degeneracy effects, implying weakly interacting intruder states, and the level-shift (LS) technique is especially useful in order to check the validity of the perturbation treatment performed. The level-shift CASPT2 (LS-CASPT2) method removes efficiently weak intruder states by the addition of a shift parameter ε , to the zeroth-order Hamiltonian and a subsequent back correction of its effect to the second-order energy.

Many times, one has to apply both strategies: enlargement of the active space to overcome the problem of severe intruder states, and, with the enlarged active space, the LS technique is applied in order to minimize the effect of weakly interacting intruder states. Since a constant added to a linear and Hermitian operator (like \mathbf{H}) does not affect its eigenfunctions but it is added to its eigenvalues (let k be a constant; as long as $\mathbf{H}\Psi = E\Psi$, then $(\mathbf{H} + k)\Psi = \mathbf{H}\Psi + k\Psi = E\Psi + k\Psi = (E + k)\Psi$), the energies of the states may be altered, but this is a price that must be paid. However, it is desirable that the shift parameter is as small as possible. For instance, results at $\varepsilon = 0.0$ (standard CASPT2), 0.1, 0.2, 0.3, and 0.4 a.u. are sufficient to establish the proper behavior of the LS-CASPT2 results. It is extremely dangerous to rely on just one result because the appearance of an accidental near degeneracy might lead to large errors in the excitation energies. In order to demonstrate the proper performance of the LS-CASPT2 technique, calibration calculations of that type always have to be carried out. The best choice for ε is the lowest possible value capable of removing intruder states. In order to avoid singularities, currently the technique so-called imaginary level-shift should be employed. The dependence of the energy on the imaginary level-shift parameter is of minor relevance (see Forsberg and Malmqvist 1997; Roos et al. 1996).

Finally, the multistate CASPT2 (MS-CASPT2) procedure has to be mentioned. It represents an extension of the CASPT2 method for the perturbation treatment of chemical situations that require two or more reference states. The procedure implies the use of an effective Hamiltonian in which the diagonal terms are the single-root, and non-orthogonal, CASPT2 solutions, whereas the interacting terms form the off-diagonal components. The diagonalization of the Hamiltonian, after symmetrization, provides a set of MS-CASPT2 energies and orthogonal states ready to describe situations such as avoided crossings and near-degeneracy of valence and Rydberg states, which cannot be fully accounted for by just using a single-reference perturbation treatment. A new wave function, named Perturbatively Modified CAS-CI (PMCAS-CI), is obtained built as a combination of the previous SA-CASSCF states, which has been in many cases improved from the previous set. This approach is not free of problems, as it will be discussed along this chapter (Finley et al. 1999; Serrano-Andrés et al. 2005).

The CASPT2 approach, especially when combined with ANO-type basis set functions, has proved to give a balanced and accurate description of all types of excited states and electronic structure cases, independently of their nature. The energy and the wave function are treated in general in an unbiased way. As all other quantum-chemical methods, the results heavily rely on the proper determination of the structural parameters, that is, the geometry optimizations. Because of their computational cost, geometries are obtained at lower levels of calculations than energies. In many cases, analytical gradients, more convenient for optimization processes, are not available at the highest level of calculation (they are, for instance, at the MRCI level, see Shepard et al. (1992) or for CASPT2 applicable to small molecules, see Celani and Werner (2003)). This is why most of the calculations use the mixed strategy, for instance, CASPT2 energies and CASSCF geometries (CASPT2//CASSCF), but also MRCI//CASSCF or CC3//CC2. Some problems related with this inconsistency will be discussed during the chapter.

It can be finally commented that, recently, the RASPT2 approach, in which a RASSCF wave function is used as reference for multiconfigurational perturbation theory has been made available, using all tools related with CASPT2, for instance allowing MS-RASPT2 calculations (Malmqvist et al. 2008). Benchmark studies are currently under way in order to determine the best partition schemes for the RASSCF active spaces. Prospectively, the approach will largely extend the applicability of the MRMP calculations.

How to Compute Excited States

How to Start: Selection of Goals, Methods, Geometries

Quantum-chemical methods provide information for excited states directly applicable to explain and predict the spectroscopy, photophysics, and photochemistry of molecular systems. A balanced description of the different electronic states is required in order to obtain the initial, basic data, that is, energy differences and transition probabilities, in an accurate way. This goal is a much more difficult task for excited states as compared to the ground state. First, one has to deal with many classes of excited states, each one showing different sensitivity to the amount of electronic correlation energy and, also, flexible one-electron basis functions able to describe all effects simultaneously are required, in general larger than that used in ground-state quantum chemistry. Then, it is necessary to compute extremely complicated potential energy hypersurfaces where the number of minima, transition states, and surface crossings like conical intersections, is multiplied. Because of the inherent complexity of the problems, the methods and algorithms to compute excited states are not as widespread as for ground states or are still under development.

In this section, we are going to illustrate different examples for the calculation of excited states using presently available quantum-chemical techniques. In each subsection, we will proceed in the same manner, first, by describing an actual

example taken from the literature, and, after, practical aspects which explain the choices made. The results obtained will be then discussed, including new cases if necessary.

Before introducing the first example, let's make some comments on the initial way to tackle the problem. Our goal, in principle for a typical ground-state closed-shell case, will be focused in the lowest-lying singlet and triplet excited states of the molecule. Most of the interesting spectroscopy and photochemistry will take place in the low-lying states, although higher states are relevant in other contexts. Properties of interest provided by static electronic structure quantum-chemical calculations that help to rationalize the photophysical and photochemical processes in a molecule are: molecular structures, charge and spin distributions, electronic and vibrational energies, oscillator strengths, dipole and transition moments and their directions, radiative lifetimes, nonadiabatic, vibronic, spin-orbit, and electronic couplings. In a second step, in reaction dynamics calculations, reaction rates, lifetimes, and population distributions can be provided.

Which method is the most appropriate for the purpose of that research? The initial questions to answer is what is the goal of the study, how far can it go, how many questions can be solved, and, especially, which is the accuracy required in the study. All this depends in most cases on the complexity and the size of the system. Of course, the best ratio quality/CPU time should be looked for, but it is useless to carry out calculations, no matter how cheap they are, with methods that cannot provide accuracy to make conclusive predictions. Apart from checking carefully the literature, calibrations will be typically necessary. Both methods and basis sets should be tested, and several steps to approach to the problem are maybe necessary. We can start by using less expensive semiempirical classical or TD-DFT methods in order to understand the problem and its requirements, and then move to more sophisticated coupled-cluster or multiconfigurational approaches, depending on the goals of the research. The same with basis sets. 6-31G-type basis sets are practical because of their reduced time requirements, but they are less accurate and in many cases using them may degrade too much our results. Calibrating the obtained values with fully correlated methods or higher-quality basis sets like Dunning's correlation consistent or ANO-type basis sets is always a requirement. Obviously the program package we should use has to balance all aspects we have mentioned. Normally a combination of them will be necessary: versatile programs, such as GAUSSIAN, GAMESS, or QCHEM; packages specialized in DFT implementations, like ADF, TURBOMOLE, or NWCHEM; suites focused on CC algorithms such as DALTON or ACES II; or programs especially designed for multiconfigurational methods like MOLCAS or MOLPRO, these more generally applicable to the calculation of excited states.

A point to bear in mind is that there are different types of excited states, and each of them has different theoretical requirements. A valence state can be viewed as a promotion from one occupied molecular orbital to a virtual one. On the other hand, a Rydberg state, in a simple MO theory, is the result of exciting one electron from one occupied molecular orbital to an atomic-like orbital of higher quantum number. Valence excited states are more compact than the diffuse Rydberg states. A valence

state can be denoted as covalent or zwitterionic according to the type of valence bond (VB) structures. Covalent and zwitterionic states are described by hole-hole and hole-pair VB structures, respectively. What is more, the label ionic is reserved to states with an actual charge separation. On the other hand, if a pair of nearly degenerated nonbonding orbitals is occupied with a total of two electrons in the ground state, the molecule is called biradical. In addition, there are charge transfer states, in which an electron is transferred from one occupied orbital of a molecule to a virtual orbital of another molecule (or the same molecule, if it is intramolecular).

Excited states can be also classified according to the main types of configurations involved in their description, in relation to the ground state: singly excited states, doubly excited states, and so on. Rydberg states are usually well described by only one singly excited configuration. Multiconfigurational singlet excited states with a large contribution of doubly excited configurations are normally covalent. Singlet excited states of zwitterionic character are described by one or several singly excited configurations. In anionic systems, new types of excited states are present, such as the resonance states or the multipole-bound (dipole, basically) states, where the bonding energy is the result of the interaction between the additional electron and the multipole moments of the molecule.

Molecular Photophysics: Computing Absorption and Emission Spectra

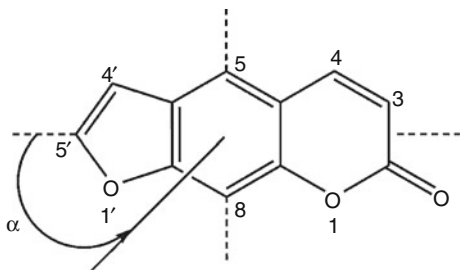
An Application Example: Psoralen

The simplest approach to the electronic state problem is the determination of the photophysics of an isolated molecule, that is, to reproduce its absorption and emission spectra. As mentioned, let's first describe in its full extent a case, the psoralen molecule, and later the choices made can be analyzed (see Serrano-Pérez et al. 2006, 2008a, b, c, d).

Furocoumarins (also named psoralens, see Fig. 9) are a class of heterocyclic compounds with a known phototherapeutic activity. These systems have been found to possess mutagenic properties when applied in conjunction with near UV-A light (320–400 nm) exposure. The technique so-called PUVA therapy (psoralen + UV-A) has been specifically designed to treat different skin disorders such as psoriasis and vitiligo. It is generally assumed that there is an oxygen-independent mechanism, which implies a [2 + 2]-photocycloaddition of psoralen and a pyrimidine DNA base monomer, and there is also an oxygen-dependent mechanism, in which energy transfer between the furocoumarin and molecular oxygen present in the cellular environment is produced generating cytotoxic singlet molecular dioxygen. It is believed that the state protagonist of the photosensitizing action, in both cases, is the lowest-lying triplet excited state of psoralen. Initially, our main goal is to describe the absorption spectrum of the molecule, and we do it by computing the lowest-lying singlet and triplet states at the optimized structure of the ground state (i.e., those distances, angles, and dihedral angles which make the energy minimum), what is

Fig. 9 Psoralen labeling.

The *arrow* defines the positive angles of both the dipole moment and electronic transition moment directions with respect to the pseudosymmetry long axis



typically known as the Franck–Condon geometry. Later, the mechanisms for triplet state population can be determined.

Using CASSCF multiconfigurational wave functions as reference, the second-order perturbation theory through the CASPT2 method was employed to include dynamic correlation energy in the calculation of the electronic excited states. The imaginary level-shift technique was employed in order to prevent the effect of intruder states. A shift parameter of 0.3 a.u. was selected by checking the stability in the excitation energies. The molecular symmetry was constrained to the C_s point group. An atomic natural orbital (ANO-L)-type basis set, contracted to C,O [4s3p1d]/H [2s1p] was used throughout. The carbon and oxygen 1 s core electrons were kept frozen in the second-order perturbation step. Geometries were obtained by computing analytical gradients at the RASSCF level of calculation for the ground and the lowest singlet and triplet excited states. In the optimization of the A' states, an active space of 14 a'' active orbitals and 16 electrons, i.e., the full π system of the molecule, has been employed, and up to quadruple excitations were considered (eight orbitals in RAS1 space and six orbitals in RAS3 space). Within the irreducible representations (a' , a'') of the C_s group, this active space can be labeled as (0, 14). An additional oxygen lone-pair orbital was included in the active space (1, 14) in order to optimize the lowest A'' excited state. In all the remaining calculations, CASSCF wave functions were generated as state-average (SA) CASSCF roots of a given symmetry. Based on preliminary RASSCF calculations and using the criterion of the largest natural orbitals occupations for the states of interest, the CASSCF active space was reduced to include 12 active electrons and 12 active orbitals (0, 12) for A' roots and 14 active electrons and 13 active orbitals (1, 12) for A'' roots. The CAS state interaction method (CASSI) was used to compute transition properties, including the spin-orbit coupling (SOC) elements between selected states.

At the Franck–Condon (FC) geometry, the lowest singlet excited states $2^1A'(\pi\pi^*)$, $1^1A''(n\pi^*)$, and $3^1A'(\pi\pi^*)$ lie at 3.98, 5.01, and 5.03 eV, respectively (see Table 1). Whereas the transition to the $n\pi^*$ state is predicted with negligible intensity, the $\pi\pi^*$ states have related oscillator strengths of 0.027 and 0.107. Unlike other states, the $3^1A'$ state has a high dipole moment, 8.70 D, differing by more than 2.5 D from that of the ground state, and therefore the associated transition is expected to undergo a red shift (bathochromic effect) in polar environments. The recorded absorption spectra in different solvents, from cyclohexane to water,

Table 1 Computed excitation energies ΔE (eV) at CASPT2 level, oscillator strengths f , dipole moments $\mu(D)$, dipole moment directions μ_{dir} (deg), and transition dipole moment directions TDM_{dir} (deg) for the low-lying electronic transitions of psoralen

State	ΔE	f	μ	μ_{dir}	TDM_{dir}
$1^1A'$	—	—	6.25	128	—
$2^1A'(\pi\pi^*)$	3.98	0.027	6.50	132	4
$1^1A''(n_o\pi^*)$	5.01	0.000	2.07	130	—
$3^1A'(\pi\pi^*)$	5.03	0.107	8.72	132	5
$4^1A'(\pi\pi^*)$	5.22	0.064	7.07	−23	−10
$5^1A'(\pi\pi^*)$	5.30	0.331	7.07	−37	102
$6^1A'(\pi\pi^*)$	5.70	0.091	7.37	−37	30
$1^3A'(\pi\pi^*)$	3.27	—	5.40	124	—
$2^3A'(\pi\pi^*)$	3.55	—	5.35	125	—
$3^3A'(\pi\pi^*)$	4.08	—	5.48	127	—
$1^3A''(n\pi^*)$	4.85	—	2.15	111	—

display a weak and structured band ranging from 360 to 270 nm (3.44–4.77 eV). Depending on the band resolution and the environment, one or two maxima near 330 and 280 nm (3.76 and 4.43 eV, respectively) have been described. The present computed results suggest that this set of features can be better assigned just to the $2^1A'(\pi\pi^*)$ transition, with the weak $n\pi^*$ band lying beneath. In that case, the observed band profile should be attributed to vibrational structure. This explanation is not unlikely, considering that a noticeable rearrangement of the molecular bond distances occurs at the $2^1A'$ minimum.

Although less relevant regarding the phototherapeutic properties, we will describe the higher-energy region of the absorption spectrum. Transitions to the $4^1A'$ (5.22 eV) and $5^1A'$ (5.30 eV) ($\pi\pi^*$) excited states have oscillator strengths of 0.064 and 0.331, respectively. The recorded spectra show a single and sharp band peaking near 248 nm (5.00 eV) in cyclohexane and 240 nm (5.16 eV) in ethanol and water. In principle, the observed feature can be assigned to transition to the $3^1A'$ state at 5.03 eV. Moreover, taking into account our computed results, an additional and more intense band can be expected at higher energies (5.30 eV). The measured band is probably a combination of both transitions.

Three other $^3A'\pi\pi^*$ states are next in energy at 3.55, 4.08, and 4.66 eV. The nature of the low-lying transitions of each symmetry, which are those basically responsible for the photophysical properties of psoralen, can be graphically described by computing the differential electron density plots as displayed in Fig. 10. Transition to the $S_1\pi\pi^*$ state is mainly benzene-like, with the charge migration concentrated in the central benzenoid ring. On the contrary, that related to the $T_1\pi\pi^*$ state has its major contributions in the pyrone ring, with high participation of the carbonyl oxygen and a shift in the density away from the pyrone ring C_3 – C_4 bond, which will be later discussed as an essential feature of the photophysics of the system. Also in Fig. 10, we find the expected differential density plots of the $n\pi^*$ states centered on the carbonyl group. Transition dipole moment directions (TDM_{dir}) indicate that the three lowest $\pi\pi^*$ features have nearly parallel polarizations, that is, they are aligned with the long axis of the molecule, while

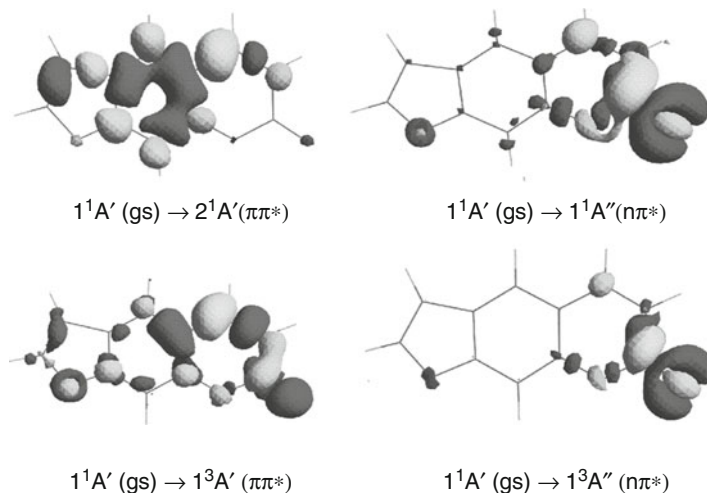


Fig. 10 Differential electron density for the main valence transitions in psoralen computed at the ground-state optimized geometry. The electron density is shifted upon light-induced excitation from *darker* to *lighter* regions

transition to the $5^1A'$ state has essentially perpendicular polarization. Regarding the vertical excitations to the triplet states, the $1^3A'\pi\pi^*(T_1)$ state lies at 3.27 eV, near 0.7 eV below the $2^1A'(S_1)$ state.

Fluorescence has been reported for psoralen in polar solvents starting (T_0) at 350 nm (3.54 eV) with a maximum at 409 nm (3.03 eV). Phosphorescence has also been recorded in solution with band origin at 456 nm (2.72 eV) and a maximum between 460 and 490 nm (2.7–2.5 eV).

The fluorescence and phosphorescence quantum yields were measured in ethanol as $\Phi_F = 0.019 - 0.02$ and $\Phi_P = 0.13$, respectively. The obtained ratio Φ_F/Φ_P is approximately 7.1. The total phosphorescence decay time (τ_P) has been reported 1.1 s in glycerol-water and 0.66 s in ethanol. With those data, the phosphorescence radiative lifetime $\tau_{\text{rad}} (= \tau_P \cdot \Phi_P)$ can be therefore expected between 8 and 5 s.

The low-lying singlet excited state $2^1A'(\pi\pi^*)$ is responsible for the lowest-energy absorption and emission fluorescence bands (see Table 2). Vertically, at the ground-state geometry, the transition energy is computed to be 3.98 eV and, upon relaxation of the geometry, the band origin (T_e) decreases to 3.59 eV. This means that the range of absorption goes from 3.59 to 3.98 eV, well within the PUVA action. A similar relaxation is observed experimentally between the lowest-energy absorption band maximum and the band origin. The structural changes of the computed equilibrium geometries for the ground (S_0) and the $2^1A'(\pi\pi^*)$ states affect the bond alternation of the system, mainly in the central ring (cf. Figs. 10 and 11), as expected from the differential charge density plots. By using the Strickler–Berg relationship, a fluorescence radiative lifetime of 74 ns is calculated for the S_1 state. The low-lying $1^1A''(n\pi^*)$ state (vertically S_2) becomes relaxed by more than

Table 2 Computed and experimental excitation energies (eV) and emission radiative lifetimes (τ_{rad}) relevant for the photophysics of psoralen

Theoretical (CASPT2)				
State	E_{VA}	T_e	E_{VE}	τ_{rad}
$2^1A'(\pi\pi^*)$	3.98	3.59	3.45	74 ns
$1^1A''(n_o\pi^*)$	5.01	3.91	2.78	3 μs
$1^3A'(\pi\pi^*)$	3.27	2.76	2.29	28 s
$1^3A''(n_o\pi^*)$	4.85	3.84	2.79	9 ms
Experimental (data in ethanol)				
State	Abs _{max}	T_0	E_{max}	τ_{rad}
$2^1A'(\pi\pi^*)$	3.7-4.3	3.54	3.03	–
$1^1A''(n_o\pi^*)$	–	–	–	–
$1^3A'(\pi\pi^*)$	–	2.7	2.7	5–8 s
$1^3A''(n_o\pi^*)$	–	–	–	–

E_{VA} vertical absorption, T_e adiabatic electronic bandorigin, E_{VE} vertical emission, Abs_{max} experimental absorption maximum, T_0 experimental band origin, and E_{max} emission maximum

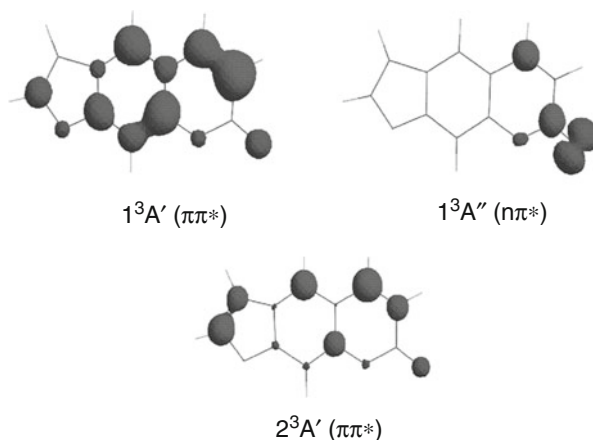


Fig. 11 Spin density for the low-lying triplet states in psoralen computed at the ground-state optimized geometry

1 eV upon geometry optimization. Although the $1^1A''(n\pi^*)$ minimum belongs to the S_1 hypersurface, the final T_e value is about 0.3 eV higher in energy than the computed and measured band origin for $2^1A'(\pi\pi^*)$. Therefore the $n\pi^*$ state is not a plausible candidate for the fluorescence, which is better attributed to the $\pi\pi^*$ state.

The $1^3A'(\pi\pi^*)$ state is the clear protagonist of the phosphorescence. The computed band origin at 2.76 eV perfectly relates to the observed value in solution at 2.72 eV. The relaxation energy is about 0.5 eV. The largest structural change is produced in the C_3 – C_4 bond of the pyrone ring, which enlarges by near 0.13 Å from the ground-state value. The computed spin population, displayed in Fig. 11,

is mainly placed on each of the carbon atoms forming the bond. In that way, psoralen becomes highly reactive in its lowest triplet state through its pyrone C₃–C₄ bond. This finding is the cornerstone of the photophysics of psoralen, which has been repeatedly proposed to take place through a reactive triplet state. The computed phosphorescence radiative lifetime is 28 s, somewhat higher than those estimated experimentally from the quantum yield and the total relaxation time, 5 and 8 s. For the $2^3A'(\pi\pi^*)$ state, the spin population is placed mainly on the carbon atoms forming both C=C bonds, that is, C₃–C₄ (pyrone) and C₄, –C₅, (furan).

Practical Aspects

Selection of Geometries

Any quantum-chemical calculation starts with the definition of the molecular geometry. As mentioned in section “[Methods, Advantages, and Drawbacks](#),” we can safely assume that the absorption spectrum is initiated at the ground state equilibrated geometry. An optimization is therefore required, in principle at the highest reasonable level of theory. DFT (with a proper functional) works well for most closed-shell cases, and it is a fully correlated method. Other ab initio methods that also include all correlation energy will be more accurate, but also more computationally expensive: MP2, CCSD, CASPT2 (with numerical gradients for small molecules). Although the best choice is not to use any symmetry restriction and let the system to find the lowest-energy solution, it is possible to make approximations and, for instance, let the system to be planar (C_s point group), like in the case of psoralen and save computational time, especially if experiment give some indication along these lines. In order to confirm that the obtained point is a true minimum, it would be desirable to perform a frequency analysis, although computing Hessians (second derivatives) is an expensive task. If too demanding, maybe is feasible to compute the Hessian at lower levels of theory and check the outcome. Of course, if possible, one should compare the result with experimental (X-ray or electron diffraction, for instance) data. To compute emission, we have to optimize excited states too, a more difficult task than for ground states only at reach of few methods: CASSCF, RASSCF, CASPT2, and CCSD, to mention the most reliable ones (the two latter only for small systems). If the ground state is optimized with a better method than the ground state, some unbalance in the results may occur. It can be a good idea to get the different states optimized at the same level, like CASSCF or RASSCF. The choice of RASSCF in the psoralen example is motivated by the impossibility of including all π space into a CASSCF calculation. Although safe in many cases, the exclusion of some MOs, like, for instance, the lowest-energy nodeless π orbital, may lead to problems if the excluded orbital, instead of becoming delocalized and having then no biased effect, tends to localize unbalancing the calculations. It is worth noticing that CASSCF (RASSCF) gives very good geometries in π -conjugated systems, probably because of some compensation between the obtained single and double bonds lengths.

Accuracy of the Excitation Energies and How to Compare with Experiment

It is time to decide if the method and strategy employed is accurate enough to solve the problem. In principle, the accuracy should be established by previous experience if the electronic structure problem is under control. We cannot expect a TD-DFT calculation to be more accurate than 0.5–0.4 eV, much less if the problem has no closed-shell ground state or a charge transfer character (except for specific functionals like CAMB3LYP, see Peach et al. (2008)). Choosing one functional for each molecule makes theoretical chemistry useless as a predictive tool. It is much better to stick into one procedure and know their limitations. Coupled-cluster results, EOM-CC(T) or CC3, are much more accurate, sometimes near 0.2–0.1 eV if pushed to the limit, but this also relies on being far from degenerate situations or multiconfigurational ground states (see Grein (2009), 1^1B_2 state of ozone: CCSD(T) 1.81 eV vs. best estimation 4.11 eV, or Kowalski (2005), 2^1A_1 state of ozone: EOMCCSD 9.82 eV vs. best estimation 4.33 eV). Multiconfigurational calculations, CASPT2 in particular, can reach also 0.2–0.1 eV in most cases, although the result may degrade if the reference is not well described. More important that all we have said: accuracy with respect to what? Many studies discussing the precision of a method for excited states compare their vertical excitation energies with what experiment initially offers, absorption maxima. This is wrong and misleading. A vertical excitation energy computed at the electronic ground-state equilibrium geometry is a purely theoretical concept without experimental counterpart. The Franck–Condon principle relates such value with the absorption (not emission) maximum, but this approximation will be true only in the case that the ground and excited minima take place at the same geometry. Apart from that, an experimentally determined absorption maximum depends on the concentration of the sample, the apparatus, the environment, the temperature, the pressure, etc. In the most favorable case, vertical absorption and emission maximum will differ by near 0.1 eV from an experimental gas-phase or molecular beam maximum. Typically, those cases display as maximum feature the band origin or T_0 transition. Indeed, the true experimental data to compare with this band origin, that theory computes as the energy difference between the states minima, also named the adiabatic energy difference, T_e if it is just electronic energies or T_0 – the true comparable value – if the zero-point energy correction has been included. The situation is even worse for the emission, where the FC principle is hardly fulfilled. In the psoralen example, a good agreement has been obtained for the band origins: computed T_e 3.57 and 2.76 eV and measured T_0 3.54 and 2.7 eV for fluorescence and phosphorescence, respectively. In most cases, T_e and T_0 differ by close to 0.1 eV. On the other hand, the vertical emission is typically much lower than the emission maxima.

How to Deal with Symmetry

Unfortunately, the topic is too extensive to be treated here. For practical cases, the reader is referred to the advances examples in the MOLCAS program manual (see <http://www.teokem.lu.se/molcas>). Group theory may be very useful in a calculation on excited states, especially because it allows dividing the states in each one

of the irreducible representations (since a symmetry operation leaves a molecule unchanged) and consequently reduce the computational effort. Unfortunately, the molecular symmetry is typically only found for ground states, and most of the photochemistry takes place at distorted asymmetric structures, especially for organic systems. Linear molecules or many inorganic complexes can however be studied making an extensive use of the symmetry, as well as other systems if absorption, and sometimes emission, spectra are studied. Symmetry is particularly useful in multiconfigurational calculations to reduce the active space requirements. Also, symmetry permits prediction of intensities when following selection rules or analysis of the polarization or dipole moment directions. Normally, one should have at hand the character tables and also the Wigner–Witmer rules, as those found in Herzberg books. Finally, it is necessary to warn about symmetry breaking problems, in which a bad selection of the method or the active space may lead to find spurious lower-energy solutions.

How to Obtain Intensities and Band Shapes: Vibrational Contributions

Apart from the energies, computing electronic oscillator strengths, f , provides information about the relative intensity of the different transitions, initially for those states allowed in one-photon (optical) spectroscopy. Group theory indicates that one-photon allowed transitions of the molecule are those in which the direct product of the symmetries of the initial state, the corresponding dipole moment component (x , y or z), and the final state belong to the totally symmetric irreducible representation of the point group of the molecule:

$$\int \psi_m^* \hat{d}_{el} \psi_n d\tau_{el} \Rightarrow \Gamma(\psi_m) \otimes \Gamma(\hat{d}_{el}) \otimes \Gamma(\psi_n). \quad (6)$$

For instance, in the case of a C_S molecule like psoralen (in the XY plane), the allowed transitions are those in which the aforesaid direct product is A' . Otherwise, the transition is forbidden. In other words, from the ground state A' , the accessible states are A' if the light is x -polarized or y -polarized, and A'' if the light is z -polarized. If the light is x -polarized, for instance, the transition $A' \rightarrow A'$ is allowed, whereas the transition $A' \rightarrow A''$ is forbidden. The magnitude of the adimensional f has to be compared to the area beneath the band representing the transition, not to the height. Frequently, especially in gas-phase spectra, weak bands such as the Rydberg transitions are the highest because they sit on top of the more intense, although broader, valence bands. Notice that the range of energies of a band depends on the length of the vibrational progressions, which, in turn, relies on the difference in geometries with the ground state. The larger is the difference, the broader and extended will be the band. Flexible systems like NH_2 or $C=O$ -based spectra like that of formaldehyde extends their valence bands and spread their intensity by several eVs. In many occasions it may be required to compute the vibrational profile of different electronic bands and plot them together. This is an expensive task because, requires the calculation of Hessians, FC factors, and sometimes the derivatives of such factors and of TDMs in order to obtain vibrational TDMs. This is, however,

the only procedure to elucidate complex cases, for instance, that of the lowest single valence 5.5–6.5 eV band of pyrrole, which contains two valence and four Rydberg bands which have all to be resolved (see Roos et al. 2002). Ideally, and in order to get all actual excitation bands, calculations on vibrational TDMs should be complemented with the inclusion of vibronic couplings, in which more than one electronic state is considered at the same time (breaking the Born–Oppenheimer approach), an example of what can be found in the pioneering studies of Domcke and coworkers on the pyrazine molecule (Domcke et al. 1993).

How to Add Environmental Effects

Adding the effects of the environment for excited states accurately is, if possible, even more complex than for the ground state. Usual procedures use cavity models such as Onsager's or the Polarized Continuum Model (PCM), with the additional consideration of the non-equilibration of the electronic response for the excited states that leads to divide the reaction field in slow, inertial, and fast, optical, parts. Results obtained with cavity models cannot be expected to be as accurate as those for the isolated system when compared with gas-phase results, among other things, because using large basis sets as those required for excited states will force the charge to leave the cavity and provide non-physical results. In many cases, the information yielded by the dipole moment of the states will be informative enough for qualitative purposes. In the psoralen example, it was discussed how states with dipole moments larger than that of the ground state were expected to stabilize in polar solvents (and undergo a spectral red-shift) than those with smaller dipole moments (blue-shift), typical case of the $n\pi^*$ states, which, additionally, tend to directly interact with protic solvents forming hydrogen bonds and pushing the excitation energy up in energy, sometimes even 0.5 eV. These interactions cannot be included by the cavity models and specific molecules have to be applied, even several solvation shells. Careful microhydration experiments allow nowadays comparison with such type of calculations.

Solvation is a very dynamical phenomenon which requires also the inclusion of statistical effects. More sophisticated studies require the employment of dynamical approaches making use of statistical mechanics, such as Monte Carlo type of calculations. Solvent molecules can be then simulated by point charges (like in QM/MM approaches as it will be discussed later) and dynamical time shots with their positions taken for a subsequent quantum chemical calculation. The required property will be obtained as an average of the different conformations, as some studies on 2-aminopurine reported recently have shown.

Active Spaces for Multiconfigurational Methods

As a large part of studies on excited states, and especially in photochemistry, employ multiconfigurational approaches, it is necessary to understand the process of selection of an appropriate active space (AS) for such computations. An active space should contain all orbitals and electrons relevant for the chemical process under study. The size and nature of the AS define the type of states and processes the multiconfigurational method will compute. Obviously, there is no single choice,

and the size is limited to 14–15 or 30–30 orbitals/electrons in CASSCF and RASSCF calculations, respectively, depending on the type of partitions made. It is important to emphasize that the selection depends on the problem. If one is only interested in low-lying excited states, maybe a small AS is enough for that purpose. On the other hand, more states imply larger ASs. In photochemistry the situation is worse, because the AS must be flexible enough to include all MOs participating in the process in all different regions of the PEHs which are chemically relevant. For instance, each bond that breaks means that the AS should contain both the corresponding bonding and antibonding MOs. In the psoralen example above the highest-lying $\pi\pi^*$, MOs plus one lone-pair MO from the carbonyl oxygen were selected because the purpose was to compute the low-lying spectra. The selection was a combination of experience and use of tools like performing a RASSCF calculation on several roots with enlarged active spaces to analyze the MOs participation in the different states and eliminate those orbitals with occupation numbers <0.05 and >1.95 in all roots of interest. If Rydberg states were to be computed, they should be added into the active space too. Making a TD-DFT calculation also helps to select the AS. As a general rule, it is wise to have in the AS one “virtual correlating” MO for each heavily occupied MO, for instance, each π MO with its corresponding “correlating” π^* MO. This is not necessary for lone-pair orbitals. Other well-known requirements relates to the need to include a second “correlating” d shell (4d) when computing the excited states of first-row transition metals, or more specific rules for very heavy elements. One big advantage of using symmetry in the calculations is the possibility of splitting the AS following the symmetry requirements, always when the energy differences are computed between states computed with the same AS size. The reader is referred to the specialized literature and the MOLCAS manual.

The recently developed RASPT2 method has opened the field to the use of larger RASSCF ASs. Special care has to be taken in the proper distribution of the MOs into the three RAS subspaces, and in the inclusion of an excitation level high enough (triples typically) from and to RAS1/RAS3 to get accurate results. The method is still under calibration, but it is a big step forward for the application of the methodology to larger and more complex systems. Typical improvements include the placement of Rydberg MOs or the second d-shell into the RAS3 space.

Computing Rydberg States

An Application Example: Water

How do we understand a Rydberg state? If we extract one electron from each of the MOs of a neutral molecule, the system becomes positively charged. An electrostatic interaction is therefore established between the molecular cation and the negative electron up to the moment that this ends up as a free electron, that is, at the energy named ionization potential (IP). While the electron leaves the molecule many metastable situations, the Rydberg states, take place, one series from each of the MOs converging to the respective IPs. Obviously, the electrons will be located

relatively far from the molecule, therefore the Rydberg states will differ in extension from the compact valence ones and will require basis sets that generate large and diffuse orbitals able to represent them. Typically, the Rydberg states are labeled, and even represented, in the united-atom approach, that is, by using atomic-type orbitals, because they “see” the molecule as a single atom. The Rydberg states are of interest in gas-phase photochemistry; because of their diffuseness, they are strongly affected by external fields and solvation. Solution chemistry does not contain Rydberg transitions.

Treating simultaneously valence and Rydberg states is not that simple. It requires methods able to deal with the mixing of configurations and orbitals in the wave function. In the past, the literature has been plagued with discussions about how physical is the effect of the valence-Rydberg mixing. In most cases, such mixing was just a consequence of the lack of electron correlation in the calculation of the wave function. Then, both valence and Rydberg orbitals and configurations mix, and valence states, become more diffuse and Rydberg states more compact, with specially dangerous consequences for the valence states, which are extremely sensitive to the mixing. Therefore, if a calculation does not provide clear and compact valence states the result is always suspicious.

Not always the mixing is spurious. One intriguing case is the water molecule. A comprehensive *ab initio* study performed in 1974 by Goddard and Hunt characterized all the computed states below 11.7 eV as having Rydberg nature, a result supported at the configuration interaction (CI) level. The electronic spectrum of water in the gas phase is currently interpreted as composed either of different Rydberg series or implying excited states with a significant Rydberg character at the ground-state equilibrium geometry. The lowest-energy band of the gas-phase electronic spectrum of water is broad (6.8–8.2 eV), with poorly defined vibrational progressions, and has its maximum absorption around 7.4 eV. There is a unanimous agreement in assigning the lowest-energy band to the 1^1B_1 state, but the valence, Rydberg or mixed valence-Rydberg, nature of the state is still under debate.

In C_{2v} symmetry, the ground state of the water molecule is mainly described by the electronic configuration $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^2(1b_1)^2$. The latter orbital is nonbonding. The valence orbitals comprise the four highest occupied MOs listed above and two unoccupied orbitals of a_1 and b_2 symmetries ($4a_1$ and $2b_2$ orbitals). The valence excited states can interact with the corresponding Rydberg states of the same symmetry and close in energy. For instance, the valence singly excited $1b_1 \rightarrow 4a_1$ configuration could be mixed with the $1b_1 \rightarrow 3s$, $1b_1 \rightarrow 3p_z$, $1b_1 \rightarrow 3d_{x^2-y^2}$, and $1b_1 \rightarrow 3d_{z^2}$ Rydberg states.

The properties of the lowest-lying electronic states were studied in the light of CASPT2//CASSCF procedure (see Rubio et al. 2008). The coupling of the CASSCF wave functions via dynamic correlation was dealt by using the MS-CASPT2 method. In this way, all the states of a given symmetry are allowed to interact under the influence of dynamic correlation, and the possible erratic valence-Rydberg mixing can be removed. Properties of the states were determined from the PMCAS-CI wave functions.

Table 3 Computed MS-CASPT2 vertical transition energies (ΔE , eV), oscillator strengths (f), and orbital extensions ($\langle r^2 \rangle$, a.u.) from the PMACAS-CI wave functions for the lowest-lying singlet states of the water molecule at the ground-state equilibrium geometry

State	ΔE	f	$\langle r^2 \rangle$
1^1A_1			13
$1^1B_1 (1b_1 \rightarrow 3s/4a_1)$	7.50	0.033	30
$1^1A_2 (1b_1 \rightarrow 3p_y/2b_2)$	9.27	Forbidden	43
$2^1A_1 (3a_1 \rightarrow 3s/4a_1)$	9.86	0.032	45

The next table lists some properties of the low-lying electronic states of the water molecule, employing (after calibration) a one-electron basis set O[5s4p2d1f]/H[3s2p1d] + (2s2p2d), being the diffuse 2s2p2d set, located at the oxygen atom only, necessary to properly describe Rydberg states. Indeed, if no Rydberg orbital is included into the active space, the CASSCF calculation will only yield roots corresponding to valence states. Rydberg and valence orbitals must be treated simultaneously, and this is not possible if there is no Rydberg orbital in the active space. The full valence active space comprises six orbitals with eight electrons and corresponds to (3120), labels that represents each one of the irreducible representations of the point group ($a_1 b_1 b_1 a_2$).

Regarding the first excited state, 1^1B_1 , it has a diffuse nature according to the value of $\langle r^2 \rangle$, more than twice that of the ground state, but appreciably lower than that obtained for a pure Rydberg state as the $2^1B_1 (1b_1 \rightarrow 3pz)$ state ($\langle r^2 \rangle = 64$ a.u.). As the state is mainly described by the singly excited $1b_1 \rightarrow 3s/4a_1$ configuration (75 % in the PMCAS-CI wave function), the valence-Rydberg character of the state arises mainly from the nature of the excited orbital, labeled here as $3s/4a_1$ to highlight its intermediate character valence ($4a_1$)-Rydberg (3s), and not from configurational mixing (Table 3).

We can get further insight into the nature of such a state analyzing its evolution with respect to the O–H internuclear distance. For the sake of simplicity, we have considered the symmetrical stretching of both O–H bonds, preserving thus the C_{2v} symmetry of the system. MS-CASPT2 calculations were performed for the three lowest roots of 1^1B_1 symmetry using the O[5s4p1d]/H[2s1p] + (2s2p2d) basis set (since now we will use the same basis sets for each state, one for vertical transitions, another for potential curves analysis) and the (5120) active space (full valence active space plus two extra Rydberg orbitals of a_1 symmetry). The bond angle was kept fixed at 104.5° .

The second and third 1^1B_1 states have Rydberg character at all distances examined since they show high values of $\langle r^2 \rangle$ in comparison with the ground state. However, the situation is clearly different for the 1^1B_1 state. The evolution of the computed $\langle r^2 \rangle$ with the O–H distance shows that the state is getting more and more valence character when stretching the O–H bond as a result mainly of the changes undergone by the virtual orbital: the energy of the antibonding $4a_1$ orbital decrease when lengthening the O–H distance separating this orbital from the related 3 s-Rydberg orbital. In other words, the nature of the 1^1B_1 state evolves from a valence-Rydberg

mixing type ($1b_1 \rightarrow 3s/4a_1$) at the ground-state equilibrium geometry to a valence character ($1b_1 \rightarrow 4a_1$) at O–H bond lengths around 1.25 Å and longer, which corresponds only to a symmetrical bond stretching of 0.3 Å from the ground-state equilibrium value. Accordingly, the second 1B_1 root becomes the Rydberg state ($1b_1 \rightarrow 3s$) at O–H distances longer than 1.32 Å. This state constitutes an example of MO Rydbergization postulated by Mulliken. The reason for such behavior is probably that both valence and 2 s Rydberg MOs share the same energy and position in space, yielding therefore a common mixed state (Fig. 12).

A similar analysis of the 1A_2 state shows another example of Rydbergization, as its nature evolves from ($1b_1 \rightarrow 3p_y/2b_2$), with a 73 % weight at PMCAS-CI wave function at geometries close to the ground-state equilibrium geometry to

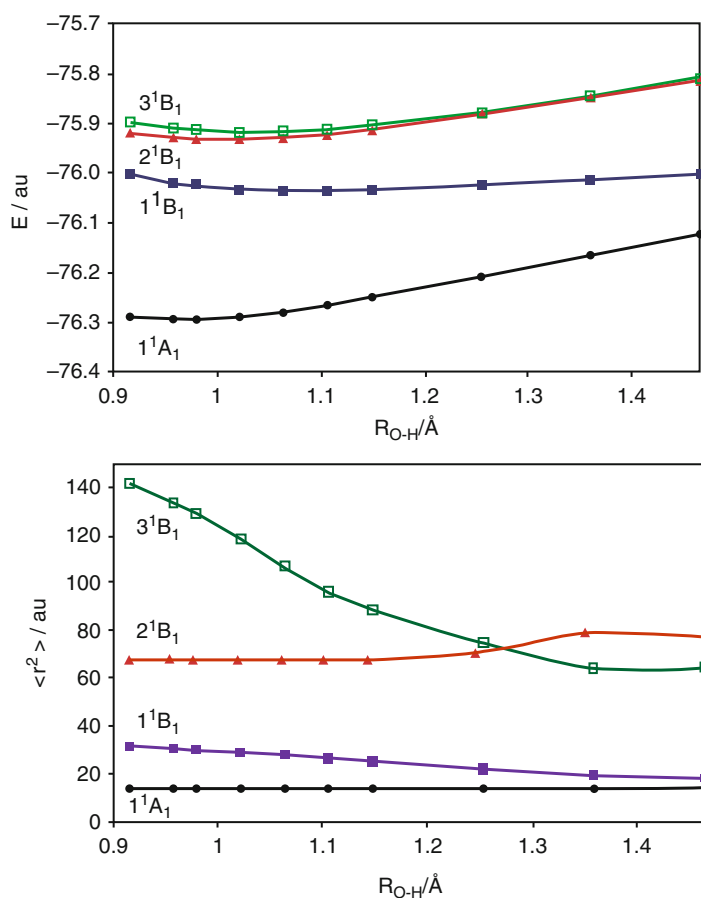


Fig. 12 Symmetrical stretching of the O–H bonds in the ground and low-lying 1B_1 states of the water molecule. Potential energy curves referred to the corresponding ground-state minimum (*left*) and computed $\langle r^2 \rangle$ values

($1b_1 \rightarrow 2b_2$) at O–H distances longer than ~ 1.217 Å, where the second 1A_2 state becomes the $1b_1 \rightarrow 3p_y$ Rydberg state. Indeed, the diffuse nature of the 1A_2 state decreases when increasing the O–H distance. Calculations of the stretching potential energy curves were performed using the active space (3141) with eight active electrons.

On the other hand, the 2A_1 state is characterized mainly by means of the $3a_1 \rightarrow 3s/4a_1$ configuration, with a weight of 70 % at the PMACAS-CI wave function, and by $1b_1 \rightarrow 3p_x$, with a weight of 28 %. This state is closely related to the 3A_1 state, which is characterized by the same configuration but in different proportion (69 % and 28 %, respectively). Since the $1b_1 \rightarrow 3p_x$ one-electron promotion is a pure Rydberg excitation, its contribution increases the configurational mixing of the state. Calculations carried out using the active space (4220) show that the valence character of the state increases with the stretching the O–H bonds due to the increment of the weight of the singly excited $3a_1 \rightarrow 3s/4a_1$ configuration in the PMACAS-CI wave function of the state together with the transformation of the excited orbital toward a valence $4a_1$ orbital. To sum up, the 2A_1 state is a valence-Rydberg state at the ground-state equilibrium geometry and the mixing is not only due to the nature of the excited orbital ($3s/4a_1$), but also to configurational mixing involving the Rydberg $1b_1 \rightarrow 3p_y$ excitation.

In the light of these results, it is concluded that the electronic spectrum of gas-phase water has a predominant Rydberg character at the Franck–Condon geometry, in line with the general view, although a certain degree of valence character has been shown to exist in the three lowest singlet states, a property that increases far from the FC region.

Practical Aspects

How to Solve Valence-Rydberg Mixing

Many methods are unable to solve properly the valence-Rydberg mixing because they do not include enough correlation effects in the calculation of the wave function, for instance, TD-DFT, CASSCF, RASSCF, or even CC2. In many cases, the problem is aggravated by the use of excessively diffuse basis set functions. For instance, aug-cc- or 6-31++G-type basis sets are already too diffuse, becoming even more when atom-centered and combined in the LCAO approach. The excess of uncontracted diffuse functions has been found necessary to compensate the poor recovery of the correlation energy of the underlying valence basis sets. A procedure to reduce the mixing problems and even to make the obtained Rydberg solutions identifiable is to employ molecule-centered uncontracted functions, as explained in the previous example. Typically, orbitals, excitations, and even population analysis are more clearly represented in such cases.

If the mixing persists, and it can be observed by comparing the orbital extension (second Cartesian moment $\langle r^2 \rangle$) of the valence state with that of the ground state, which should be similar and different from that of the Rydberg state, the only solution is to increase the level of correlation used to compute the wave

function. In the CC treatments, it might mean to include the triple excitations. For multiconfigurational studies, it means to make the CASPT2 states to interact and display orthogonal solutions via the MS-CASPT2 method. Ethene is the most paradigmatic case. A CASPT2(2e,9MOs)/ANO-L C[5s4p2d]/H[3s2p] + 1s1p1d calculation yielded an expected $^1B_{1u}$ valence and Rydberg states at 8.45 and 8.93 eV with orbital extensions $\langle x^2 \rangle$ 52 and 50 a.u., and with related oscillator strengths 0.261 and 0.166, respectively (see Finley et al. 1998; Müller et al. 1999). Several errors plague these results. First, the ground state extension was 12 a.u., therefore none of the two states seemed to have proper valence character. Also, Rydberg states cannot have oscillator strength values close to those of valence states. Second, the correct energetic values are known to be ~ 8.0 and 9.33 eV, respectively, leading then to deviations of ± 0.4 eV. A clear mixing of states was obtained at the CASSCF level (in which our wave function was obtained) that CASPT2 could not solve properly for the energies. Using the multistate MS-CASPT2 method makes both states to interact. By diagonalization of the effective interaction Hamiltonian, a new set of states is obtained. The new wave functions, PM-CASCI, are linear combination of the previous SA-CASSCF references built with the eigenvectors of the MS treatment. The new results provide two new solutions for the $^1B_{1u}$ valence and Rydberg states at 7.98 and 9.33 eV with orbital extensions $\langle x^2 \rangle$ of 20 and 82 a.u., and with related oscillator strengths 0.360 and 0.051, respectively, leading to a proper interpretation of the spectrum.

A proof that the valence-Rydberg mixing problem is just a problem of lack of correlation in the wave function is obtained when performing RASSCF/RASPT2 calculations on the ethene molecule. When a calculation including all valence σ , π , π^* , σ^* and Rydberg orbitals in the RAS active spaces is carried out, the mixing vanishes even at the RASSCF level and compact valence and diffuse Rydberg states are obtained separated. The RASPT2 step, without the need of a multistate treatment, produces already the correct results.

Focusing on the Valence States

Is it possible to ignore the Rydberg states and focus on the valence spectrum? Logically, this depends on the system and also on our goals. A system with a solvated or proteinic environment perturbs so much the diffuse transitions that it is safe to ignore the Rydberg solutions. For isolated systems, the simultaneous treatment of valence and Rydberg states seems unavoidable, for instance, in pyrrole, where the lowest-energy excited single state is of Rydberg character. Using poor basis sets like 6-31G* type or excluding in general the diffuse functions from the basis set or the Rydberg orbitals from an active space may seem a solution (the Rydberg states cannot be represented in such case), but the obtained “valence” solutions will most probably be of mixed type.

For larger systems the, exclusion of the Rydberg orbitals from the basis or the active space may be less dangerous. Valence states decrease much more in energy with the enlargement of the molecule or the complexity of the chromophore than Rydberg states. It is unusual to find Rydberg series below 4.5–5.0 eV. Therefore, if one is interested in computing low-lying valence states in a large molecule, it

might be safe to first estimate where the lowest-energy Rydberg state is placed and consider that valence states below such energy are treated accurately with compact basis sets. From one to other molecule, this estimation can be done using the respective IPs, which are a good measure of where the Rydberg series are being located energetically. On the other hand, the geometry of the Rydberg states tends to be similar to that of the molecular cation, typically different from that of the valence states. A way to simplify the calculations in many systems may be to use a compact valence basis set to get the geometry of valence states that, far from the FC region, might not suffer of mixing with the Rydberg states. A recent example can be found in the literature for 1,1'-bicyclohexyliden (see Pérez-Hernández et al. 2008).

Electronic States of Anionic Systems

An Application Example: p-benzosemiquinone Radical Anion

When computing excited states in anionic systems, several cautions should be taken into account. In some molecules, certain states in the anion may lie below the ground state of the neutral system, for instance, those in which the actual electron affinity is positive. Most of the electronic states in anions, including in many cases the ground state, are, however, higher in energy than the ground state of the neutral system and represent temporary anion states, which means that they are unstable with respect to electron detachment. They are typically named Temporary Negative Ion (TNI) states, resonances, or anionic valence-bound states. Conventional quantum-chemical techniques cannot be applied in general to the study of these temporary states since they lie in the continuum of the neutral species plus the free electron. It has been shown that it is, however, possible to obtain well-localized solutions with *ab initio* methods such as CASSCF and others. Those eigenvalues can be regarded as discrete representations of the TNI or anion resonance states. The metastable states of anions can be classified as either shape or core-excited resonances. From the electronic structure standpoint, shape resonances originate in the attachment of the electron to a virtual orbital of the neutral ground state. Alternatively, they can be viewed as the result of the promotion of the unpaired electron of the LUMO to higher-lying virtual orbitals. Core-excited resonances can be seen, on the other hand, as arising from the attachment of an extra electron to an excited state of the neutral molecule. They can be divided into Feshbach and core-excited shape resonances, depending on their energetic position with respect to the excited state of the neutral molecule involved. The former lie below the parent state of the neutral, whereas the latter are located above.

The theoretical treatment of temporary negative ions can proceed via the use of scattering theory or the employment of modified bound-state quantum chemical techniques. The latter requires the use of stabilization methods for the obtained solutions, for instance, by decreasing the exponents of the diffuse functions. It is of general knowledge that any quantum chemical calculation in anions requires a basis set including extra diffuse functions. Besides the discrete valence TNI states, also named valence-bound (VB) anion states, diffuse states described by a singly

excited configuration involving a diffuse orbital can be also obtained, and they have been sometimes erroneously interpreted as Rydberg states. Singly excited Rydberg states are not present on monoanions because there is no specific direct electrostatic interaction between the neutral molecule and the electron; doubly excited Rydberg states would be instead present at high energies. Then, if not Rydberg states, what is the nature of the diffuse states that a quantum-chemical calculation finds interleaved between the valence excited states of the anion? In the cases where the neutral system has a dipole moment larger than a critical value established between 1.26 and 2.50 D, metastable solutions known as dipole-bound (DB) anion states should be found consequence of the interaction between the additional electron and the dipole moment of the neutral system. These states will have small binding energies (i.e., they will be more stable than the neutral ground state), but in some cases, they may appear higher than the neutral state because of deficiencies of the quantum-chemical treatment. Upon improvement of the level of the calculation, they should end up with a positive binding energy. In systems with smaller dipole moments, however, spurious diffuse solutions are also present together with the VB and DB states, but they are an artificial consequence of the calculation, in particular of the structure of the one-electron basis set, which forces a confinement effect (known as the basis set cage effects) leading to erroneous results. In the low dipole moment molecules, these diffuse solutions for the excited states of the anion try to simulate the neutral molecule plus a free electron by placing the extra electron into the most diffuse orbital available. The computed energy for these states strongly depends on the diffuseness of the basis set employed. In practice, the only way to check if the obtained result is correct is to analyze the stability of the energy, for instance, with the increase of the diffuseness of the basis set, as it will be discussed below.

We have a hint of how to analyze ionic systems in the monodeterminant HF approach. The Koopman's theorem stated that given an N -electron HF single determinant $|\Psi_0\rangle = |\chi_1\chi_2\cdots\chi_c\cdots\chi_N\rangle$, the ionization potential (IP) to produce and $(N-1)$ -electron single determinant by removing an electron from spin orbital χ_c , and the electron affinity (EA) to produce and $(N+1)$ -electron single determinant by adding an electron to spin orbital χ_r , are just $\text{IP} = E_c - E_0 = -\varepsilon_c$ and $\text{EA} = E_0 - E_r = -\varepsilon_r$, respectively; that is, the corresponding orbital energies of the protagonist spin orbitals. This "frozen orbital" approximation assumes that the spin orbitals in the $(N \pm 1)$ -electron states are identical with those of the N -electron state, neglecting the relaxation of the spin orbitals in the ionized states. As a result, Koopman's theorem calculations tend to produce too positive IPs and too negative EAs. In addition, we should take into account the correlation effects, which one obtains in going beyond the HF approximation, which will produce further corrections. In general, Koopman's ionization potentials are reasonable first approximations to the experimental ones. On the contrary, Koopman's electron affinities are often inaccurate. Obviously, correlated methods are needed to determine quantitatively such properties.

The ground state of the anion can be then lower (positive EA) or higher (negative EA) in energy than the ground state of the neutral system. How about the excited states of the anion? They will lie in general higher than the neutral ground state,

in the continuum region. As mentioned, it is possible to obtain discrete solutions of the Hamiltonian corresponding to the valence-bound states of the anion, but, because of the confinement effect of the basis set functions, other solutions in which the electrons are placed in the diffuse orbitals will be also present. In a series of calculations on the *p*-benzosemiquinone radical anion (pBQ^- , nonpolar in the neutral form), a number of experiments were performed (see Pou-Amérgo et al. 2000). Apart from the clearly localized valence-bound anionic states, a number of supposedly diffuse states were obtained interleaved with the valence states. What are those solutions? They are spurious solutions caused by the cage effect of the basis set. By decreasing the orbital exponents in the C and O basis sets, it was observed that, while the valence states remained in energy, the “diffuse” solutions started to decrease their absolute energy. The instability remained until their energy converged to the energy of the neutral system, while the electron becomes free and detached from the molecule, an effect that is observed because it always ends up on the most diffuse orbital available. Once identified, the stable solutions above the neutral ground state energy can be TNI states or anion resonances, whereas the instable solutions can be neglected as spurious. The apolar character of pBQ^- prevents the molecule to have dipole bound anion states (although it has quadrupole-bound states).

The CASPT2 study on pBQ^- used a basis set of ANO-L type C[4s3p1d]/H[2s1p] plus a set of 1s1p1d diffuse functions centered in the molecule and an active space of 9 MOs of valence $\pi\pi^*$ type, (03010301), distributed into the D_{2h} irreducible representations ($a_g b_{3u} b_{2u} b_{1g} b_{1u} b_{2g} b_{3g} a_u$). The adiabatic electroaffinity (AEA), obtained as the difference between the ground states of the neutral and anion pBQ at their respective geometries was computed as -0.33 eV (CASSCF) and 2.01 eV (CASPT2), as compared to the 1.9 eV measured in experiment. The result clearly highlights the fundamental role of the correlation effects. Two series of calculations on the excited states were performed to analyze the spectrum of the system at the geometries of neutral and anion pBQ . The former results can be compared with those obtained from photodetachment measurements, whereas the latter are better related to the electron attachment or electron transmission spectroscopy (ETS). Table 4 summarizes the results.

Whereas most of the assignments, made to $\pi\pi^*$ valence-bound states, seem to be explained, the observed band near 2.7 – 2.9 eV in the pBQ^- absorption spectrum could not be initially resolved because the degeneracy of the 1^2B_{3u} and 1^2A_u states, both leading to one-photon allowed transitions. This broad band had two main peaks at 2.7 and 3.1 eV, which could be attributed to one or other states. In principle, the transition to the 1^2A_u was computed with larger oscillator strength, but that does not guarantee that the highest peaks correspond to such a state. The vibrational profile of the transitions to both states was therefore computed at the CASPT2//CASSCF level of calculation by obtaining vibrational energies and TDMs, which required geometries optimizations and Hessians for both states and resolution of the vibrational Hamiltonian. As a result, it was confirmed that the two highest bands at 2.7 and 3.1 eV corresponded to the initial quanta in the vibrational progression of the breathing mode in the 1^2B_{3u} state, whereas the 1^2A_u transition,

Table 4 Computed and experimental excitation energies (eV) and oscillator strengths of the p-benzoquinone radical anion

State	Anion absorption				State	Attachment energies ^a			
	Theoretical ^b		Experimental ^c			Theoretical	Experimental ^d		
	E_{VA}	f	A_{max}	f		AE ^e	SF ₆	ETS	EI
1^2B_{2g}	—	—	—	—	1^2B_{2g}	-1.64			
1^2B_{3g}	2.25	Forbidden	2.41	Forbidden	1^2A_u	0.91	0.70	0.72	0.77
1^2B_{3u}	2.80	0.05	2.7/3.1	0.15/0.06	1^2B_{2u}	0.96			
1^2A_u	2.82	0.17	2.7/3.1	0.15/0.06	1^2B_{3u}	1.31	1.35	1.46	1.6
1^2B_{1g}	3.25	Forbidden	—	Forbidden	2^2B_{3u}	1.87	1.90	2.15	2.0
2^2B_{3u}	3.56	0.32	3.8/3.9	0.50/0.35	1^2B_{1g}	1.99			

^aElectron attachment energies at neutral ground state geometry

^b E_{VA} : vertical excitation energy at the anion ground state geometry

^cAbsorption maxima in acetonitrile/water

^dSF₆ scavenger spectra (SF₆), electron transmission spectroscopy (ETS), and vibrational excitation by electron impact (EI)

^eAdiabatic energy difference: energy difference between the anion state and the neutral ground state geometry at such geometry

even when broader, extended to higher energies displaying a larger band area but not so high peaks. Such result proves the danger implicit in the use of vertical excitation energies matching band maxima to assign molecular spectra.

Practical Aspects

Basis Sets and Spurious Solutions in Anions

Closely controlling the quality of our basis set and its effects on our states will confirm the validity of our results. Let us illustrate this aspect with another example. Neutral nitromethane (CH₃NO₂) has a dipole moment of 3.46 D. A standard ANO-L C,N [4s3p2d]/H[3s1p1d] basis set is used supplemented with a set of 1 s diffuse functions (eight primitives) with exponents for the diffuse functions explicitly optimized to deal with Rydberg states. At the CASSCF optimized geometry for the neutral molecule, the lowest state of the anion is 0.530 eV above the neutral ground state at the MS-CASPT2 level. Just by scaling the exponents of the diffuse functions by a factor 0.5 in both the neutral ground state and the lowest anion state, the energy difference decreases to 0.250 eV. Scaling instead the exponents by 0.1 leads to an energy difference of -0.022 eV, that is, the lowest state of the anion has converged below the energy of the neutral system. Increasing further the diffuseness of the basis set does not vary the result, which corresponds to a dipole-bound state of the anion with a small binding energy of 0.022 eV (experimental value 0.012 eV). The lowest valence-bound anion state is computed, adiabatically, 0.23 eV below the neutral ground state of the molecule at the CCSD(T) level. This state remains basically stable with the increase of the diffuseness of the basis set and it is, in general, better represented by using diffuse functions localized in the positive sites of the molecule. To test the stability of the solutions obtained, other techniques have

been developed, such as increasing the effective positive nuclear charge, adding specific counter ions to fix the negative charge, or using penalty functions or electric fields, such the dielectric continuum cavities, which will highly perturb the diffuse solutions.

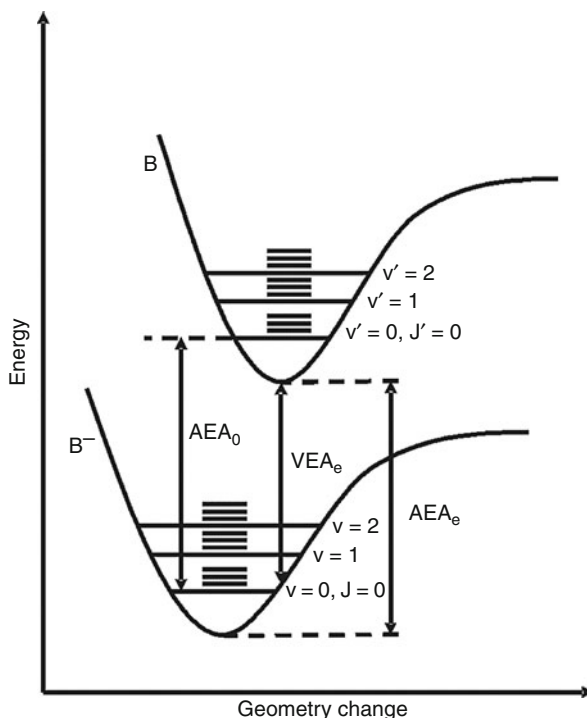
Negative Electron Affinities

Ionization potentials and electron affinities are intrinsic properties of the DNA and RNA nucleic acid bases (NABs) whose determination enables a deep understanding of all phenomena related to the electron donor and acceptor abilities of the NABs, such as those involving charge transfer and transport along the DNA strand. Determination of the EAs of NABs is difficult both experimentally and theoretically, and the uncertainties range up to several eV, including also changes in the sign of the energies, because both VB and DB anions may be located within a small range of energies, and therefore determination of accurate EAs is uncertain, especially because the type of anion formed may vary with the different experimental conditions. Apart from that, new difficulties interfere the experimental determination of EAs of nucleobases as the presence of different tautomers of the nucleobases which are close in energy in the gas phase. In particular, the canonical (keto) form of guanine, which is the biologically relevant tautomer, has a very low concentration in the vapor, and there is no direct experimental value reported for the corresponding EA.

The electron affinity of a neutral molecule is the energy required for detaching an electron from a singly charged negative ion, or equivalently, the energy released when an electron is attached to the neutral system. Thus, the electron affinity of a neutral molecule can be defined as the energy difference between the ground state of the neutral system and that of the anion. A positive EA implies that the anion is more stable than the neutral species. As in the case of the ionization potential, three theoretical magnitudes (see Fig. 13) are used for describing this transition: the vertical electronic energy difference (VEA_e or VEA) between the ground states of the neutral system and the anion at the equilibrium geometry of the neutral molecule, the adiabatic energy gap (AEA_e) between the minima of the neutral and anion molecule ground states, and the corrected adiabatic property (AEA_0 or AEA) with the addition of the zero-point vibrational energy correction (ZPE). Thus, positive $VEAs$ indicate that the molecule acts as a trap for an excess electron, with an attachment energetically favored and, therefore, the anion can be created spontaneously. In this case, positive $AEAs$ follows, and the system becomes stable, that is, it does not undergo autodetachment and can take part in chemical reactions. On the other hand, negative values for $VEAs$ and $AEAs$ represent the TNI states, existing in short periods of time and becoming prone to photodetachment.

Analysis of the experimental literature on nucleobase EAs shows an extremely confuse situation, ranging from clearly negative values (-0.56 eV) up to largely positive energies (1.51 eV), and including EAs close to 0 eV. In general, determination of EAs represents a technical challenge, especially when they have negative values, and in many cases it is based on indirect measurements. Negative electron affinities can be experimentally measured by electron transmission spectroscopy.

Fig. 13 EA diagram ($B + 1e^- \rightarrow B^-$). Definitions of the theoretical magnitudes related to EA are graphically shown through the electronic, vibrational, and rotational potential energy levels. Those magnitudes are VEA_e (vertical electronic electron affinity, from the neutral ground-state minimum), AEA_e (adiabatic electronic electron affinity, from minimum to minimum), and AEA_0 (adiabatic electron affinity including the zero-point vibrational corrections of the minima)



The technique is able to detect negative ion resonance states, which are energetically unstable with respect to electron autodetachment. It is unclear when the experiment is measuring vertical (VEA) or adiabatic (EAs) attachments, or if the indirectly obtained data truly represent the molecular EAs. In general, however, ETS is the only direct experimental technique which is expected to provide actual VB anions in the region of the resonance states. In particular for gas-phase NABs (except G, which cannot be isolated), ETS measurements report EAs values clearly in the negative region (from -0.22 to -0.54 eV).

In order to compute accurate theoretical results for the VEA and AEA of NABs different levels of theory were taken into account, employing the MP2, CCSD, and CCSD(T), and the CASSCF and CASPT2 methods, in conjunction with the 6-31G*, cc-pVDZ, aug-cc-pVDZ, ANO-L C,N,O[4s3p1d]/H[2s1p] (hereafter ANO-L 431/21), and ANO-L C,N,O[4s3p2d1f]/H[3s2p1d] (hereafter ANO-L 4321/321) basis sets. Methods and basis sets were selected to obtain the most accurate values from preliminary calculations on atomic systems, in which the required levels of highly flexible enough basis sets and strongly correlated methods to obtain predictive EAs were determined. Geometry optimizations of both neutral and anionic NABs were carried out at the MP2/6-31G(d), MP2/aug-cc-pVDZ, CASSCF/cc-pVDZ, CASSCF/ANO-L 431/21, and CCSD/aug-cc-pVDZ levels of theory. No symmetry restrictions (C_1 symmetry) were imposed, whereas all minima

were characterized by computing second derivatives at the same level, except in the case of CCSD/aug-cc-pVDZ where the geometries were tested comparing with the optimized parameters at the other levels of theory. At the respective equilibrium structures, additional CASPT2 and CCSD(T) calculations were performed to account for the most accurate energy values. ZPE corrections were included at different levels using the harmonic approach. The active space for the CASSCF calculations in geometry optimizations comprises the full π -valence system, except the molecular orbital localized mainly on the nitrogen atom of the NH_2 group in the case of cytosine, adenine, and guanine, whose occupation number is very close to two. This MO is further included in conjunction with the lone pair electrons and orbitals of the heteroatoms in the final CASSCF and CASPT2 calculations of VEAs and AEAs, except when the large ANO-L 4321/321 basis set was employed (see Roca-Sanjuán et al. 2006, 2008a).

The analysis of the results shows that neither DFT procedures nor the MP2 method have the required accuracy, either by the known problems of DFT to deal with negative centers or by the spin contamination problem that affects MP2. CCSD(T)//CCSD/aug-cc-pVDZ and CASPT2(IPEA)/ANO-L 4321/321//CASSCF/ANO-L 431/21 will be established as the most accurate procedures, both for vertical and adiabatic EAs. It must be emphasized here that the main factor to achieve accurate results for the VEA of NABs by using ab initio methods is the employment of atomic one-electron basis sets flexible enough to describe both the spatial distributions of electrons and their correlation effects and including functions decaying slowly with the radial distance. The CASPT2//CASSCF strategy has the advantage to compute several states of the system, and it is possible to easily distinguish between the different solutions. Single-reference methods are not free of problems, because they can only obtain the lowest solution. For instance, at the corresponding geometry of the neutral species, adenine and guanine, in which the VB anion state lies much higher than other diffuse states, the CCSD and CCSD(T) computations lead initially to a diffuse and low-energy spurious solution in which a delocalized electron is located far from the molecule in a diffuse orbital. Finally, the sequence of stable solutions obtained for the VEAs of NABs using the reference CCSD(T) and CASPT2(IPEA) levels of theory is established as $U \approx T > C > A > G$, ranging from -0.61 eV (U) to -1.14 eV (G). Therefore, for guanine is less favorable to accept an electron at the neutral molecule geometry.

Regarding gas-phase adiabatic EAs, those for uracil and thymine were determined very close to zero, whereas cytosine has a small negative AEA. The sequence of AEAs for isolated NABs can be established as $0 \text{ eV} \sim U \sim T > C > G > A$. Purines are much less favorable than pyrimidines to retain the electron attached to the neutral nucleobase, and after geometry relaxation adenine becomes the poorest electron acceptor of all NABs, in contrast to what occurred for the vertical EAs, in which guanine had the more negative value. To understand the differences among these compounds, it can be also analyzed how the inductive effect (and the number of stable resonance structures) makes more or less stable the new center of negative charge created by the addition of the new electron. It should not be forgotten

that the accepting properties of the systems will largely change in solvated or biological environments. For instance, a recent calculation performed for the EAs of the cytosine molecule within a chain of oligomers, $dC_{18} \times dG_{18}$, employing a CASPT2/MM approach led the AEA of cytosine to change from -0.26 eV in vacuo to 0.69 eV in the biological surrounding, becoming therefore the molecule a strong electron acceptor.

Photochemistry: On the Trail of the Energy

An Application Example: Thymine

The energy absorbed by a molecule can be released radiatively, that is, slowly emitted via fluorescence or phosphorescence, or nonradiatively. In that case, it can give rise to productive photochemistry, yielding photoproducts different from the initial species, or it can become unproductive, meaning that it is dissipated to the environment through the vibrational degrees of freedom, ending in some cases in the initial ground state of the system. Photophysics and photochemistry is typically a combination of all such processes. From the theoretical viewpoint, the best initial strategy to understand the photochemical processes is trying to follow closely the path of the energy from the initially populated states at the FC region toward favorable regions of the PEHs. That means to trace the lowest-energy possible pathway until reaching an energy barrier, that is, a minimum, or a transition state, or a hypersurface crossing, in particular conical intersections (CIs), the protagonists of the ultrafast radiationless nonadiabatic energy transfers between PEHs. The only strategy that guarantees finding PEH points along the lowest-energy path is the computation of a minimum energy path (MEP), that is, a steepest-descendent pathway. The procedure is performed generally in mass-weighted coordinates, and it is equivalent to the Intrinsic Reaction Path approach. Once localized, the crossings and CIs, the estimation of their accessibility, and the calculation of the interstate couplings (like the SOC between singlet and triplet states) and transfer probabilities will help understand and predict favorable IC and ISC processes. The described strategy is named as the Photochemical Reaction Path approach, and it can be applied to solve a large number of photochemical problems.

An interesting example of nonproductive photochemistry can be found in important compounds such as the DNA/RNA natural nucleobases, which were determined a long time ago as basically nonfluorescent. Modern femtochemical techniques have determined fluorescence decay times in the DNA/RNA nucleobase monomers, nucleosides, and nucleotides in different media to be ultrafast. In particular, molecular beam measurements reported two main decay lifetimes near 100 fs and few ps in all five natural NABS: thymine, uracil, cytosine, adenine, and guanine (Crespo-Hernández et al. 2004). Finding ultrafast decays in a molecule suggests the presence of extremely efficient internal conversion channels, and therefore the presence of accessible conical intersections, in particular one connecting the initially populated bright spectroscopic $\pi\pi^*$ singlet and the ground state, yielding therefore nonfluorescent species. The ultrashort lifetime of nucleobases

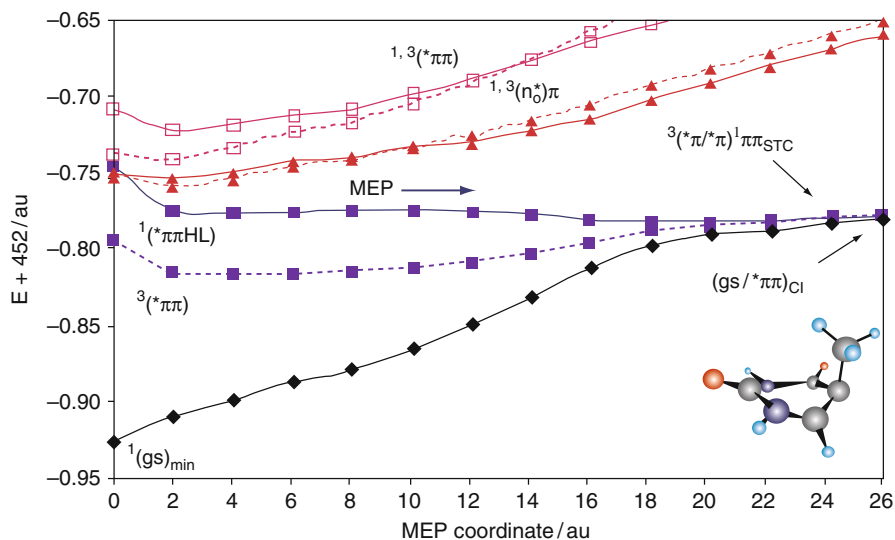


Fig. 14 Evolution of the ground and lowest singlet excited states for thymine from the FC geometry along the $^1(\pi\pi^*HL)$

is an intrinsic molecular property as it has been proved in recent years both from theory and experiment. Indeed, such photoprotective properties may well be very important at the beginning of life in our planet given that there are evidences that point out that life began on Earth millions of years before the development of the ozone layer. The ultrafast decay channels of the nucleobases may have favored photostable natural nucleobases against less stable derivatives (Serrano-Andrés and Merchán 2009). This type of photochemical mechanisms may have been operative as natural selection processes reaching in that way the genomic stability until the development of the ozone layer and complexes mechanisms of reparation of DNA.

Thymine photophysics is here selected as an illustration of the computational strategies mentioned above (see Merchán et al. 2006; Serrano-Pérez et al. 2007b). Figure 14 displays the CASPT2 energies of the lowest-lying singlet and triplet states of thymine along the MEP calculated for thymine on the bright $^1(\pi\pi^*HL, \text{HOMO} \rightarrow \text{LUMO})$ state from the FC geometry. The level of the calculation was CASPT2//CASSCF(14e⁻,10MOs)/ANO-S C,N,O[3s2p1d]/H[2s1p]. Apart from the radius of the hypersphere controlling the distance from the initial geometry, not any other restriction was imposed to the calculations.

From thymine ground state and upon near-UV absorption at the FC region, most of the population reaches initially singlet excited states, and in particular, it is the transition to the $^1(\pi\pi^*HL)$ excited state at 4.89 eV which has the largest oscillator strength up to 6 eV, that is, 0.167. The ultrafast nonradiative decay undergone by thymine in the femtosecond range can be rationalized by the barrierless character of the path leading from the FC region toward a CI seam with

the ground state, $(\text{gs}/^1\pi\pi^*)_{\text{CI}}$. Unlike simple geometry optimizations, the use of the MEP technique guarantees the absence of energy barriers along the lowest-energy path. The structure of the CI at the end of the MEP can be characterized as ethene-like, a diradical species (as many CI are) involving combined stretching and twisting of the ethylenic bond and leading to a screw-boat (S) conformation $^5\text{S}_6$ for the six-membered ring. The presence of an accessible CI explains also the low fluorescence quantum yield ($\sim \phi_F = 10^{-4}$) detected for thymine with band origin at 4.5 eV in water. This weak emission can be related to the presence of a high-lying planar $^1(\pi\pi^*HL)$ minimum computed adiabatically from the ground state at 4.49 eV, whereas a nonfluorescent $^1(n_o\pi^*)$ minimum is found at 4.05 eV with a minor contribution to the emissive properties. As recently estimated by means of reaction dynamics, the near 100 fs decay detected in natural nucleobases can be related to the barrierless path from the FC to the CI region. Although other slower singlet decay pathways could be estimated proceeding through the low-lying $n_o\pi^*$ state, that crosses the MEP at high energies and has its own CI with the ground state, still the main relaxation path for the energy runs along the computed $^1(\pi\pi^*HL)$ barrierless MEP, which will transfer most of the energy toward the ground state and back to the original species. Such behavior means that the systems are largely photostable, as it has been proved for natural and methylated nucleobases, and that the mechanism can be considered an intrinsic property of the systems. Figure 15 illustrates such mechanism for the pyrimidine nucleobases, thymine, uracil, and cytosine, although the basics also hold true for the purine systems, adenine and guanine (Merchán et al. 2006; Serrano-Andrés et al. 2006, 2008; ?). Further insight can be only obtained when reaction dynamic calculations are performed on larger regions of the PEHs (see, for instance, Szymczak et al. (2009)).

The decay along the singlet manifold is not the only procedure for energy relaxation. Efficient population of the triplet manifold can also take place, essentially through intersystem crossing (ISC) processes taking place along the main decay process on $^1(\pi\pi^*HL)$ or via photosensitization from endogenous or exogenous species. Triplet states are frequent intermediates in different types of photoinduced reactions. Both their usual diradical character and long lifetimes make them reactive species prone to interact with other systems. Among the most important reactions involving triplet states, those related to DNA/RNA purine and pyrimidine nucleobases have undoubtedly attracted more attention, in particular the photodimerization of pyrimidine nucleobases, considered to be the most frequent DNA lesion taking place after UV light irradiation. Through their triplet states, DNA/RNA nucleobases may not be as photostable as expected. As seen in Fig. 14, along the $^1(\pi\pi^*HL)$ state, MEP is clear that two singlet-triplet crossings (STC) are accessible, and therefore two ISC processes may take place: at 4.8 eV with the $^3(n_o\pi^*)$ triplet state, $(^3n_o\pi^*/^1\pi\pi^*)_{\text{STC}}$, and at 4.0 eV, further along the relaxation path and near the ethene-like CI with the ground state, directly with the lowest $(^3\pi\pi^*)T1$ triplet state, $(^3\pi\pi^*/^1\pi\pi^*)_{\text{STC}}$, a structure displaying the same type of screw-boat puckered geometry with a stretched and twisted double bond $\text{C}_5=\text{C}_6$ as at the $(\text{gs}/^1\pi\pi^*)_{\text{CI}}$ CI. Efficient ISC requires both small singlet-triplet energy gaps and large spin-orbit coupling elements at the regions of degeneracy.

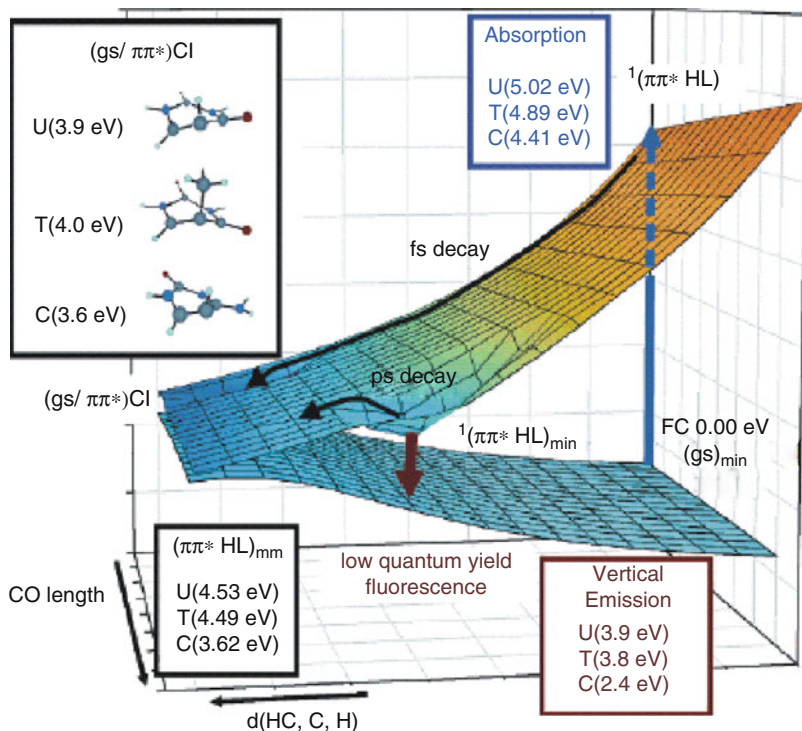
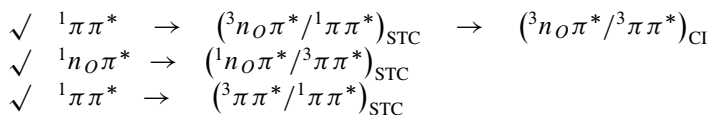


Fig. 15 Global scheme of the photochemistry of uracil (U), thymine (T) and cytosine (C) as suggested by the CASPT2 calculations

As compared to IC, taken place essentially in small zones where the seam of CIs becomes accessible, the regions of the potential energy hypersurfaces for effective ISC are more extensive.

Figure 16 includes a scheme describing the population of T_1 based on our CASPT2 calculations, and three are the suggested competitive mechanisms from the initially populated singlet state by means of STC processes:



Regarding the higher-energy ISC process, the ${}^3(n_o\pi^*)$ triplet state can be populated from ${}^1(\pi\pi^*)$ in the STC crossing region, in which a high SOC of 8 cm^{-1} has been computed favoring the process. From such region, a MEP on the ${}^3(n_o\pi^*)$ state leads directly to its energy minimum, ${}^3(n_o\pi^*)_{min}$, placed at 3.93 eV adiabatically from the ground state. Basically degenerated, we have located a conical intersection connecting the triplet states ${}^3(n_o\pi^*)$ and ${}^3(\pi\pi^*)$, facilitating the population switch

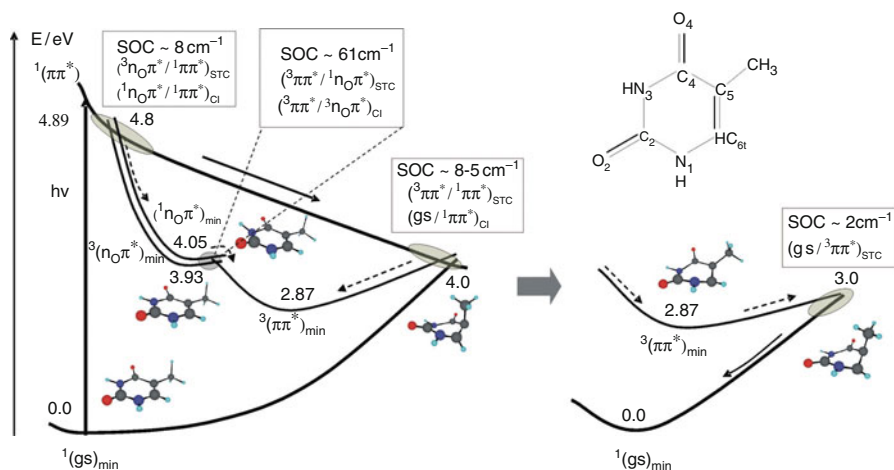


Fig. 16 Scheme, based on CASPT2 results, of the photochemistry of thymine focused on the population of the lowest-energy triplet state

toward the lowest ${}^3(\pi\pi^*)$ state. Another computed MEP on the T_1 hypersurface leads from $({}^3n_O\pi^*/{}^3\pi\pi^*)_{CI}$ to the ${}^3(\pi\pi^*)$ state minimum. A second ISC channel also relates with an $n\pi^*$ state. In our results, as well as in previous theoretical studies, the presence of the low-lying ${}^1(n_O\pi^*)$ state is confirmed, lying almost isoenergetic with the lowest ethene-like $(gs/{}^1\pi\pi^*)_{CI}$ CI near 4.0 eV. The ${}^1(n_O\pi^*)$ state may then be related to the observed dark state in pyrimidine nucleobases in solution. Additionally, we have found that, as both singlet and triplet ${}^{1,3}(n_O\pi^*)$ states have almost parallel hypersurfaces because of the typical small singlet-triplet splitting of $n\pi^*$ -type states, near degeneracy structures with the lowest ${}^3(\pi\pi^*)$ state are equal to both states. Therefore, close to the $({}^3\pi\pi^*/{}^3n_O\pi^*)_{CI}$ CI, we have found a singlet-triplet crossing $({}^3\pi\pi^*/{}^1(n_O\pi^*))_{STC}$. The extremely large computed SOC terms, $\sim -61\text{ cm}^{-1}$, guarantee also an efficient ISC process in the region, confirming this mechanism as relevant for the overall process. In principle, in different environments such as in polar solvents, it is expected that the $n\pi^*$ -type excited state will destabilize with respect to $\pi\pi^*$ -type excited states. Despite those effects, both singlet and triplet $n_O\pi^*$ -type states are estimated to lie in the solvent below the ${}^1(\pi\pi^*HL)$ state at the FC geometry, confirming the existence of the STC crossing upon decay along the ${}^1(\pi\pi^*HL)$ state. Finally, a third ISC channel directly connecting the lowest $\pi\pi^*$ states is found at low energies with SOC values ranging 5–8 cm^{-1} . The profile of the computed MEP in Fig. 16 suggests that the lowest-energy ISC mechanism may enhance its efficiency with respect to the other nucleobases resulting in larger quantum yields because the region for the energy transfer near to the end of the MEP seems to be much more extended. The presence of three basic ISC funnels in thymine (occurring also in uracil and adenine but not in guanine and cytosine where the higher-energy $n\pi^*$ -mediated channels are

absent) successfully explains for the first time the wavelength dependence on the measured ISC quantum yield (ϕ_{ISC}) reported in thymine (uracil and adenine too) on the basis of the location and accessibility of the two STC crossing regions upon the initial excitation conditions. In the case of thymine, the ϕ_{ISC} value increases from 3.9×10^{-3} at 280 nm (4.43 eV) to 5.2×10^{-2} at 240 nm (5.17 eV). At low excitation energies, only the lowest-lying ISC computed to take place close to 4.0 eV will be activated. At higher-energies, however, the channels near 4.8 eV will be additionally activated, increasing the overall triplet formation yield.

From any of the STC or CI regions, the lowest triplet state is populated by any of the previous ISC processes and reaches its minimum, as proved by the computed MEPs. The high reactivity attributed to this triplet state originates from its partial diradical character on C₅ and C₆. The minimum is placed at 2.87 eV, adiabatically from the ground state optimized minimum. As a final aspect of the evolution along the triplet manifold in thymine, we have located the singlet-triplet crossing connecting the $^3(\pi\pi^*)$ and the ground state, and mapped the MEP leading from such STC toward $^3(\pi\pi^*)_{min}$. The crossing is placed at near 3.0 eV from the ground state minimum, what means that there is a relatively small barrier of 0.13 eV ($3.0 \text{ kcal mol}^{-1}$) to reach $(gs/^3\pi\pi^*)_{STC}$ from $^3(\pi\pi^*)_{min}$. The computed electronic SOC is, however, somewhat low, $\sim 2 \text{ cm}^{-1}$, predicting for the triplet state a long lifetime and a slow relaxation, becoming therefore prone to undergone reactivity (Merchán et al. 2005; Serrano-Pérez et al. 2007b).

Practical Aspects

Reaction Paths: MEP Verses Other Approaches

Computing minimum energy paths (MEP) is an expensive procedure in which each point of the MEP is obtained from individual optimizations. Special care is required in the selection of active spaces, which should be flexible enough and appropriate along the whole path, and the radius of the hypersphere, which should be not too small or too large. When the steepest-descendent MEP ends finding an energy barrier, for instance, as frequently in the minimum of the state, other procedures should be employed. It must occur also that the calculation of the MEP is too expensive. It is possible to continue using the MEP algorithm upward, that is, by increasing the hypersphere radius increasing energy paths can be computed. Other procedure is to search for a transition state (TS) along the path that represents the energy barrier. More general strategies are normally required to map the full profile of the reaction linking two predefined regions of the PEHs. One possibility is performing a relaxed scan, that is, varying by specific amounts one internal coordinate which is expected to be the leading structural modification along the path and, at each selected value, to perform a restricted geometry optimization in which only such coordinate is frozen. Relaxed scans are extremely dangerous and should be used with caution, because they often produce unconnected and useless paths. There is no guarantee that the lowest-energy value for each of the selected frozen-coordinate structures belongs to a connected path. In many cases, the results will

jump from one region of the PEH to another along a profile that should represent a reaction path but does not at all. In general, the only relatively safe case in which a relaxed path can be used is that leading to a dissociation of an atom, because in this case the path has a descendent character while the bond length is increased.

Instead of a relaxed scan is preferable to make a linear interpolation in internal coordinates (LIIC). This is a predefined path in which each structure is generated from the previous one by adding linearly to each internal coordinate the geometrical change of the initial and final structures divided by the number of steps. The advantage of this profile is that the points belong to a connected path. The disadvantage is that it will most probably give rise to too high energy barriers, although they will always be upper bounds of the actual barrier. For instance, in a study of the photophysics of psoralen, it was determined that the MEP from the FC region led the system to the minimum of the initially populated $S_1^1(\pi\pi^*)$ state. Psoralen is a strongly emitting system, especially by phosphorescence (in a ratio 7/1 with respect to fluorescence). Emission from the $S_1^1(\pi\pi^*)$ minimum explains the fluorescence of the system, but, what about phosphorescence? The population of the triplet manifold seems not to take place close to the FC region, and no relevant STC crossing takes place soon along the MEP. Therefore, an effective triplet formation has to take place by depopulating the singlet state minimum by an accessible STC. The $T_1(T_\pi)$ state lies much lower in energy than the $S_1(S_\pi)$ state, both vertically (~ 0.6 eV) and adiabatically (~ 1.2 eV). Direct interaction between the singlet and triplet vertical excited states is unlikely, and therefore another mechanism has to be found, involving most probably population of higher-energy triplet states and subsequently internal conversion toward T_1 . The qualitative El-Sayed rules indicate that the spin-orbit coupling is large between states of $\pi\pi^*$ and $n\pi^*$ types and small between states of the same character. In addition, it is known that the ISC process is enhanced in molecules with heavy atoms (the heavy-atom effect) because the SOC terms increase. There are two necessary conditions to be fulfilled for an effective ISC to take place: low energy gap and high spin-orbit coupling between a singlet and a triplet state. Thus, the nonradiative decay to a triplet state should occur along the relaxation pathway of S_π , starting from the Franck–Condon region, and in close vicinities of a singlet-triplet crossing. A LIIC profile was computed (see Fig. 17) from $(S_\pi)_{\text{MIN}}$ to the minimum of the $1^1A''$, $(S_n)_{\text{MIN}}$, a path considered to be appropriate to find the singlet-triplet crossing, since both S_n and T_n states share the same basic structure and energetics (see Serrano-Pérez et al. 2007a). At half of the path the S_π and $T_n^3(n\pi^*)$ states cross, and the SOC value is high enough (9 cm^{-1}) to guarantee an efficient ISC. A barrier of 0.4 eV, surely an upper bound, has to be overcome to reach the T_n state.

The geometry closer to the crossing point was used as initial geometry to initiate a search at the CASSCF level for a singlet-triplet crossing (STC) between S_π and T_n , that is, $(S_\pi/T_n)_X$. As expected, the main change in geometry from the minimum of S_π to the optimized structures of S_n or T_n is related to the length of the C=O bond, much larger (~ 0.15 Å) than the ground state value for the $n\pi^*$ states. It is not surprising that the STC structure displays an intermediate value of 1.220 Å. The

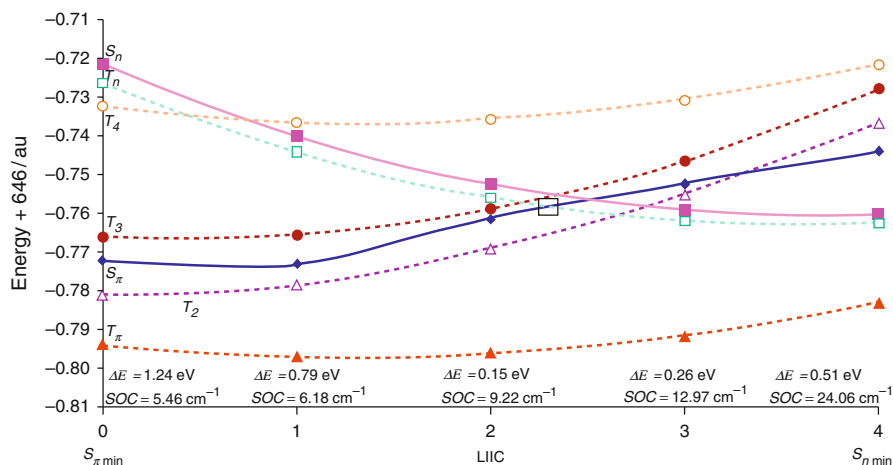


Fig. 17 LIIC path for psoralen. The crossing region between S_1 and T_n is highlighted with a square

SOC term in STC is computed as 6.4 cm^{-1} . As displayed in Fig. 18, the barrier from $(S_\pi)_{\text{MIN}}$ to $(S_\pi/T_n)_X$ was computed 0.36 eV at the CASPT2 level. This is still a large energy barrier; however, the crossing is now below the energy of the S_π state at the FC geometry, and therefore the system has enough excess energy to access the STC region and make the ISC process to take place.

Once the T_n state is efficiently populated through a favorable ISC mechanism, it will quickly evolve toward the nearby energy minimum, $(T_n)_{\text{MIN}}$, placed 0.11 eV below $(S_\pi/T_n)_X$. At those geometries, T_n is the second-energy excited triplet state and can be expected that the energy follows a pathway for favorable internal conversion (IC) toward the low-lying T_π state. Actually, a conical intersection has been found, $(T_n/T_\pi)_{\text{CI}}$, placed isoenergetic with $(T_n)_{\text{MIN}}$. An efficient IC will therefore take place transferring the population to T_π , which will subsequently evolve to its own minimum, $(T_\pi)_{\text{MIN}}$, from where phosphorescence (P) will take place. Additionally, the molecule may react with thymine or transfer its energy to the molecular oxygen exerting the photosensitizing action. The family of furocoumarins (8-MOP, 5-MOP, khellin, TMP, and 3-CPS) was studied at the same grounds showing that all the compounds share the same mechanism of population of T_π . The differences in the values of the barriers (to reach the crossing regions on the hypersurfaces), besides the values of SOC, are actually the key points in considering the photosensitizing ability of each compound.

A final remark should be introduced. Considering that $n\pi^*$ states tend to destabilize in polar solvents, the (T_n) -mediated mechanism here proposed will surely decrease its importance in polar media, remaining a plausible candidate for the efficient population of the low-lying T_π triplet state in the gas phase. Indeed, the relevance of out-of-plane displacements for the relaxation of the S_π state has been emphasized recently at DFT/MRCI level of theory, together with the enhancement

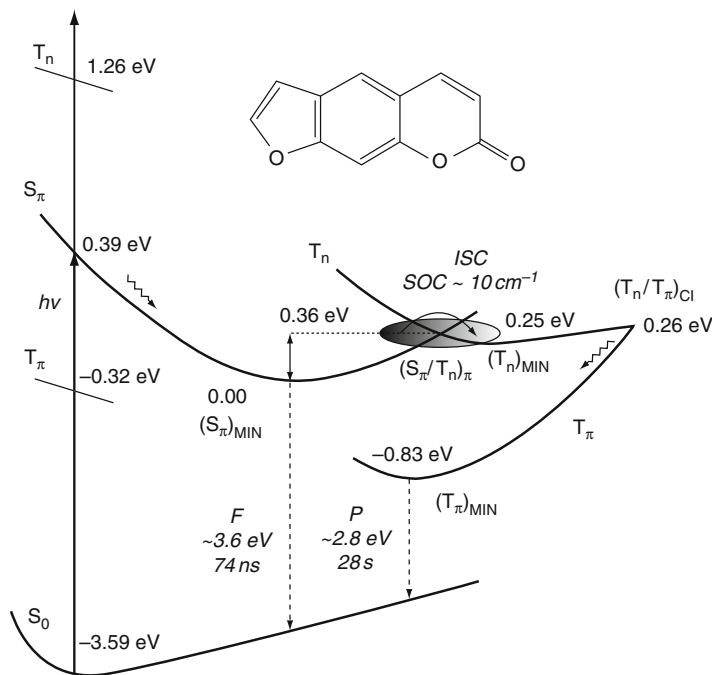


Fig. 18 Scheme of the gas-phase photochemistry of psoralen based on quantum-chemical CASPT2 calculations. Energies are referred to $(S_{\pi})_{\text{MIN}}$

of the SOC terms between low-lying S_{π} and T_{π} states by vibronic coupling effects involving out-of-plane modes, as an alternative mechanism for efficient ISC.

How to Compute Conical Intersections

There is hardly any good recipe to make the calculation of CIs more efficient, in particular those relating the ground and excited states that are usually the most relevant features for the photochemistry of the systems. Experience orientates toward certain types of CIs common in organic compounds, like the so-named ethane- or methanamine-like CIs, taking place by the stretching and twisting of $C=C$ and $N=C$ bonds, respectively, or situations of $C=O$ stretching, in both cases creating an almost diradical species. In many other cases, one has to let the computational algorithm to find the right path. One initial approach is trying to find the minimum energy CI (MECI), considered to be the most relevant for the photochemical process. Technically, it is easier to look for MECIs, because one can simply search for the lowest-energy crossing point fulfilling the nonadiabaticity rules. Calculation of nonadiabatic terms, the so-called branching space, is complex. Time-consuming projected-gradient or Barry's phase algorithms have to be used to guarantee the true CI nature. In practical cases, there are some implementations

available at a high level of theory, like that at the MRCI level in the COLUMBUS code (see Lischka et al. 2004).

In many cases, however, it is safe enough just to compute the minimum energy crossing point (MECP), discarding the calculation of nonadiabatic coupling, because the obtained solution already defines the crossing region accurately enough. Even when the MECI (or MECP) is found, nothing guarantees that this is going to be the relevant structure from the photochemical viewpoint. It might happen that the MECI is placed in a region that it cannot be easily accessed from the main energy decay paths of the molecule. Therefore, the accessibility of the region is the key point making a CI (or a seam of CIs) relevant for the photochemistry. The next example can properly illustrate this idea. In the same manner as described for thymine, the MEP from the FC region on the bright $^1(\pi\pi^*\text{HL})$ state was computed for guanine at the CASPT2//CASSCF (14e⁻,12MOs)/ANO-S C,N,O[3s2p1d]/H[2s1p] level of calculation (see Fig. 1, top), leading, through a barrierless path, to a crossing with the ground state at the end of the MEP. Independently, a MECI search was carried out to find the lowest-energy CI (Fig. 19, bottom, right), which provided a geometry somewhat different from that obtained at the end of the MEP. To solve the problem, a CASSCF path was computed along the seam of CIs (Fig. 19, bottom). The points of this profile have been computed to belong to the seam, that is, both the ground and the $^1(\pi\pi^*)$ states are degenerated, and by successive optimizations in which the radius of the hypersphere has been increased from the initial MECI structure. It is clear that the MECI energy is lower than any other point of the seam, but also that a barrier separates that structure from a CI which is higher in energy but accessible from the end of the MEP (the actual source of the energy) and whose structure is quite similar to that of the end of the MEP (Fig. 19, bottom, left). That CI is the relevant funnel for the photochemistry of guanine.

Other serious problems can be found when computing PEHs crossings, although it is as general issue in quantum chemistry. Determination of molecular structures is one of the most time-consuming steps in computational chemistry. It is therefore a typical strategy to use lower-level methodologies to locate singular-point geometries (minima, transition states, conical intersections . . .) and correct then the energetics by employing higher-level methods. Implicit in this procedure is the problem of the differential correlation effects. We shall check this issue in the case of the location of conical intersections (or any other hypersurface crossing). A usual protocol employed for that purpose is named CASPT2//CASSCF, meaning that geometries have been determined at the CASSCF level and, at those structures, the energies have been computed at the CASPT2 level. CASSCF includes the static or long-range correlation effects and provides wave functions which describe qualitatively the electronic states and form the reference for subsequent calculations. The remaining short-range correlation effects, dynamical electron correlation, are included by means of a second-order perturbative treatment, CASPT2. Depending on the nature of the state, the effects of the static and dynamical correlation energies differ, affecting directly to the description of the hypersurface. Figure 20 summarizes what we have called the differential correlation effects on the location of surface crossing within the CASPT2//CASSCF protocol. It is important to recall that this problem

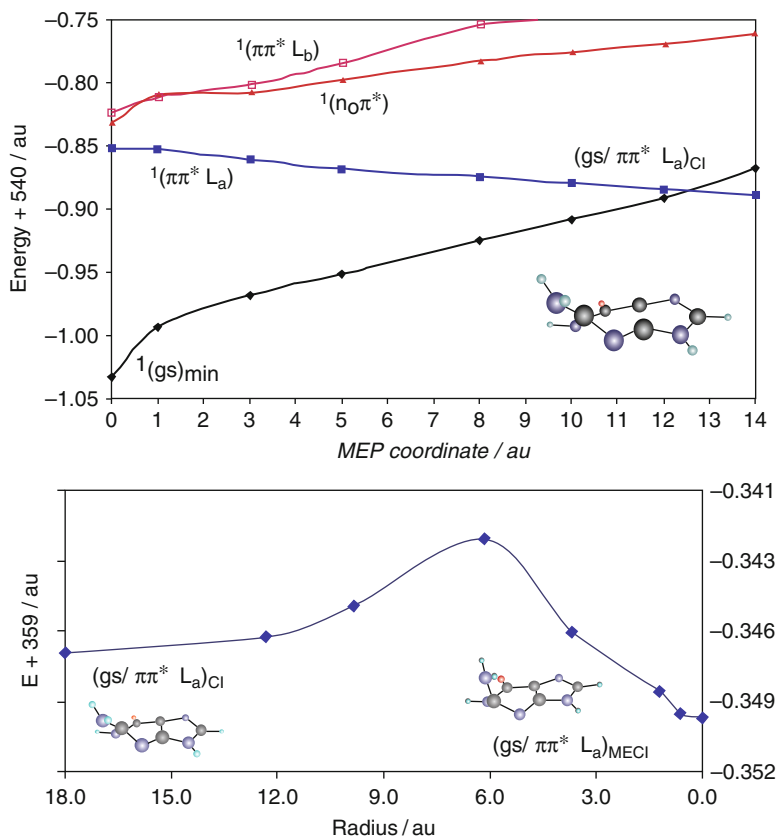


Fig. 19 CASPT2 low-lying singlet states along the MEP from the FC region on the $1(\pi\pi^*HL)$ state in guanine (*top*) and seam of CIs (*bottom*) computed from the MECI structure (*bottom, right*) to the CI close to the end of the MEP (*bottom, left*)

is general in quantum chemistry if two levels of theory are being used to compute geometries and energies, e.g., CCSD(T)//CCSD or MRCI//CASSCF.

Turning to our example, two main situations do actually occur. In cases where the PEHs computed at the CASSCF and CASPT2//CASSCF levels behave approximately parallel (Case A), the CASSCF optimized geometries will be in general correct, despite they have been computed at a lower level of theory. It means that the dynamical correlation contributions are quite regular and similar for the two states in ample regions of the PEH. In these situations, two-step computational strategies like CASPT2//CASSCF can be confidently applied. When dynamical correlation is markedly different for the states considered and varies significantly along the PEH of interest, geometry optimization has to be carried out at the highest correlated level (Case B). Otherwise, the uneven contributions of dynamical correlation may lead to unphysical crossings and interactions between the two electronic states.

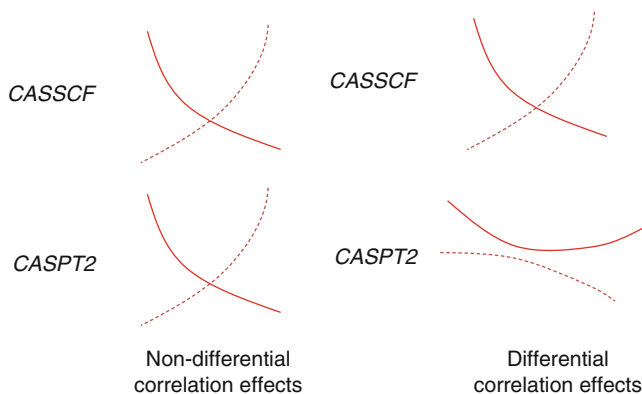


Fig. 20 Different effects of correlation within the CASPT2//CASSCF protocol

The crossing can be then placed at a different geometry or it may never occur! Normally, states of ionic (in the valence-bound sense) nature have larger amounts of dynamical correlation in contrast to covalent states, in which the static correlation is able to recover large fractions of the total correlation energy. When both types of states are involved in a crossing, it is clear that a CASSCF description will give an unbalanced picture of the hypersurfaces and, obviously, of the crossing, which can be recovered only by inclusion of the dynamical correlation, with CASPT2 for instance, even in the searching of states crossing. In difficult cases, it can be convenient to map a grid of CASPT2 points in the region of the CASSCF CI to find the degeneracy at the highest level, or, if possible, to compute the CI with CASPT2, more specifically, with MS-CASPT2, considering that CASPT2 provides non-orthogonal and therefore mixed solutions. This is a tricky issue that was studied recently and requires careful consideration.

Figure 21 displays the result of CASPT2/MS-CASPT2/ANO-L Li [4s2p]/F [4s3p1d] calculations on the lowest-lying $^1\Sigma^+$ states of the LiF molecule along its dissociation path (see Serrano-Andrés and Merchán 2005). The lowest state has ionic and covalent character at short and long distances, respectively, because the lowest dissociation channel leads to the ground states of the neutral atoms, whereas the opposite occurs for the second state. At intermediate distances, the lowest ionic and next covalent states interact and change their character smoothly thanks to the multiconfigurational treatment. According to the non-crossing rule for states of the same spin and spatial symmetry that holds true for diatomic molecules, the two curves cannot become degenerate and instead an avoided crossing should take place. In the sample calculations, the results of three different active spaces and two levels of calculation, CASPT2 and MS-CASPT2, have been displayed. In cases A and B, the CASPT2 curves (dotted lines) incorrectly cross or touch. The reason is that in a single-state perturbation theory each root is not orthogonal to the others.

Neglecting the state interaction leads to unphysical crossing situations. On the contrary, the multistate treatment of MS-CASPT2 (solid lines) makes the two states

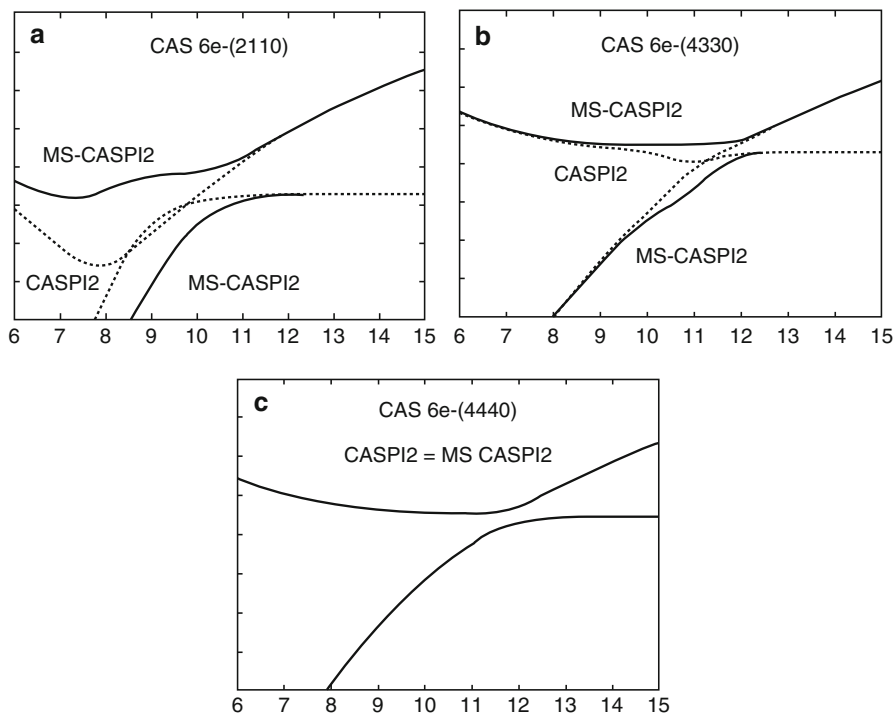


Fig. 21 CASPT2 (dotted lines) and MS-CASPT2 (solid lines) calculations on the lowest-lying $^1\Sigma^+$ states of the LiF molecule

to interact. The corresponding orthogonal solutions display therefore a “correct” avoided crossing. In case A the employed AS is the minimal one (one 2s orbital from Li and three 2p orbitals from F). Although the MS-CASPT2 solution yields a formally correct avoided crossing, the state interaction is overestimated, giving rise to a large energy splitting at too short distances (as compared to higher-level MS-CASPT2 or MRCI calculations) because of the too large value of the off-diagonal elements in the multistate effective Hamiltonian. A better treatment is obtained for enlarged active spaces in B and, especially, C, in which the increased space allows for angular correlation in the AS and the CASPT2 and MS-CASPT2 solutions match, undoubtedly because the initial CASPT2 states were including most of the interaction effects. Unfortunately in most cases the enlargement of the active space will not be possible. A good strategy is testing different spaces to check the stability of the results. The right solution seems to be between the CASPT2 and the MS-CASPT2 ones. In the worst case, the CASPT2 result is generally not so bad after all.

Table 5 compiles the results of CASPT2 and MS-CASPT2 calculations on the two lowest singlet states of the C_{2h} ethylene dimer using a cc-pVDZ basis set computed at the MECP point upon the increase of the AS (see Serrano-Andrés and Merchán 2005). In order to provide a proper multistate treatment, the off-diagonal

Table 5 Computed MS-CASPT2 energies and off-diagonal Hamiltonian matrix elements for the MECP point of the $1^1A_g/2^1A_g$ states in the C_{2h} ethylene dimer^a

	ΔE_{CASPT2}	$-H_{12}$	$-H_{21}$	$-(H_{21} + H_{12})$	$\Delta E_{\text{MS-CASPT2}}$
Active spaces	(kcal mol ⁻¹)	(kcal mol ⁻¹)	(kcal mol ⁻¹)	(kcal mol ⁻¹)	(kcal mol ⁻¹)
2,020, 4e ⁻	0.41	1.96	2.44	2.20	4.42
3,030, 4e ⁻	2.08	2.09	2.89	2.49	5.40
4,040, 4e ⁻	3.35	2.60	5.12	3.86	8.41
4,040, 8e ⁻	2.76	2.51	5.48	4.00	8.46
4,242, 12e ⁻	2.13	1.27	0.03	0.65	2.49

^aThe MS-CASPT2 calculations comprise two roots at the CASPT2 (2,020,4e⁻) optimized $1^1A_g/2^1A_g$ MECP geometry

elements of the effective Hamiltonian have to be small and similar, otherwise the matrix symmetrization will be erroneous, whereas the energy split undergone by the original states once diagonalized the matrix will be excessive. Indeed, although the CASPT2 energy differences between the states remained always small (<3.35 kcal mol⁻¹), the splitting largely increases at the MS-CASPT2 level, especially when the off-diagonal elements H_{12} and H_{21} most differ. It is only with the largest active space (angular $\sigma - \pi$ correlation has been included in the active space in this case, 4,242) that the elements become smaller and the MS-CASPT2 reduces. Values close to 2 kcal mol⁻¹ can be considered small enough to identify such feature as a CI point within the accuracy of the method. What happens when a CI search is performed blindly using MS-CASPT2 solutions? Typically, it occurs that the obtained solutions are too separated and, apparently, the closest-energy solutions correspond to an avoided crossing. This is probably a wrong answer derived from that fact that the active space is too small and the MS-CASPT2 method is overshooting the interaction of the CASPT2 solutions. If it is not possible to increase the AS (may be MS-RASPT2 is a solution), it is better trusting in the CASPT2 energies.

Setting the Path for Dynamics

Mapping the PEHs of the molecular systems is not but the initial stage of the calculations. Once the topography of the hypersurfaces is obtained by solving the time-independent Schrödinger equation, the evolution of the nuclei on that potential must be described. Studying the reaction dynamics (RD) of the system implies considering time dependence and the statistical distribution on a sample of molecules. Several procedures can be followed, from the most expensive approach, solving the time-dependent Schrödinger equation, typically along specific reaction coordinates, to the semiclassical approaches, in which quantum-chemical information is used to solve the classical Newton equations. When properly averaged, the obtained data yield, in principle, predictive data for reaction rates, state lifetimes, and population distribution. We will not discuss dynamical calculations here, but will simply try to emphasize that the quality of a RD calculation, no matter how sophisticated, strongly relies on the quality of the underlying quantum-chemical description of the PEHs.

In other words, the dynamics is always as good as the quantum chemistry employed to describe the PEH. Expensive RD calculations tend to use analytical potentials fitted with quantum-chemical parameters, a technique filled with problems. The other approach is known as direct or “on the fly” reaction dynamics, meaning that the quantum-chemical information is obtained at each of the points of a trajectory and passed to the RD step. The basic difficulty stays: how good is the description of the PEHs and how this affects to the RD results?

In the study of thymine described above, it was found that when using lower levels of calculation, for instance, a 6-31G* basis set or a small active space, the CASSCF MEPs run in wrong regions of the PEHs, something that only occurred in thymine and uracil, not in purine bases. The MEP FC¹ ($\pi\pi^*$) leads in this case to the high-energy planar minimum and not to the low-lying CI with the ground state, an outcome obtained only when increasing the quality of the calculation by using ANO-type basis sets, larger active spaces or CASPT2 gradients. It is obvious that the RD description will be strongly affected by the quality of the underlying quantum-chemical description. As an illustration, Fig. 22 includes two sample trajectories computed in the guanine molecule on the lowest-energy ¹($\pi\pi^*$) state starting near the FC region. The calculations used a Verlet-type algorithm for the classical part (which gives the solution of the equations of motion, i.e., how position, velocity, and acceleration change over time) and CASSCF(12,10) or CASPT2(12,10)/6-31G(d,p) quantum-chemical energies and gradients computed “on the fly” at intervals of 5 fs. In this preliminary study, most of the computed trajectories starting close to the FC region were shown to follow a similar path than the MEP, that is, leading to the region of CI with the ground state, and this is done in times close to 100 fs, explaining the ultrafast lifetime measured for the molecule in molecular beams (see Serrano-Andrés et al. 2008). From that point, the complications continue with the selection of the surface hopping algorithm for the population transfer, the diabatic description of the crossing regions, the re-evaluation of the reaction step, etc. (see, just as one of the many examples the use of the NEWTON-X reaction dynamics approach (Barbatti et al. 2007)).

Photoinduced Reactions in Bimolecular Systems

The theoretical description of photoinduced bimolecular reactions requires specific strategies and procedures that will be illustrated here with two examples. First, the cytosine dimer, which forms an excimer in the ground state and it is also able to react yielding a photoadduct in the ground state. Second, the energy transfer reaction between psoralen and molecular oxygen, in which the lowest-energy triplet states are the protagonists.

An Application Example of Reactivity: Cytosine Dimer

One intriguing aspect of UV-irradiated DNA is the appearance of red-shifted long-lived emissive states not found in base monomers, whereas the DNA absorption spectra closely resemble that of the monomers. This phenomenon is called excimer

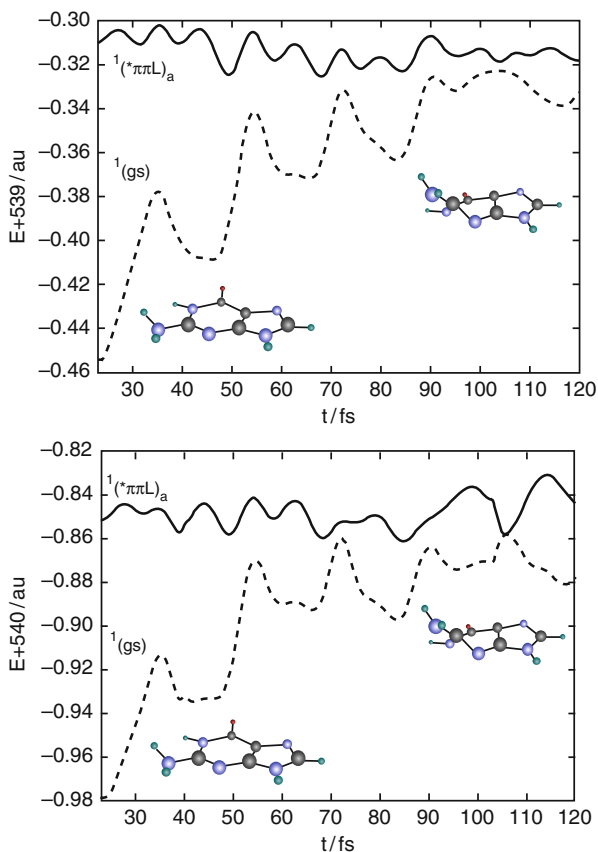


Fig. 22 Sample CASSCF (*top*) and CASPT2 (*bottom*) semiclassical trajectories run on the lowest-energy $^1(\pi\pi^*)$ state of guanine from the FC region

fluorescence, reflecting the relevant role assumed to be played by the corresponding excited dimer (excimer) of the biopolymer, given the similarity between the emission from dinucleotides and polynucleotides (Eisinger and Shulman 1968). Recent experimental results in the light of time- and wavelength-resolved fluorescent techniques using 80 picoseconds (ps) excitation pulses make readily apparent the longer-decay components and red-shifted emission that it was assumed to arise from excimer formation. Because of the slow rate of energy relaxation, these long-lived states associated to excimer-like states have been suggested as the precursors of the DNA photolesions, including photodimers (Schreier et al. 2007).

Classically, an excimer (homodimer) or exciplex (heterodimer) is defined as a dimer system which is bound in the excited state and dissociative in the ground state (see Fig. 23). The existence of excimers between DNA nucleobases was proven theoretically by means of CASPT2(12,12)/ANO-S C,N,O [3s2p1d]/H [2s1p] calculations (see Olaso-González et al. 2006). Potential energy curves (PECs)

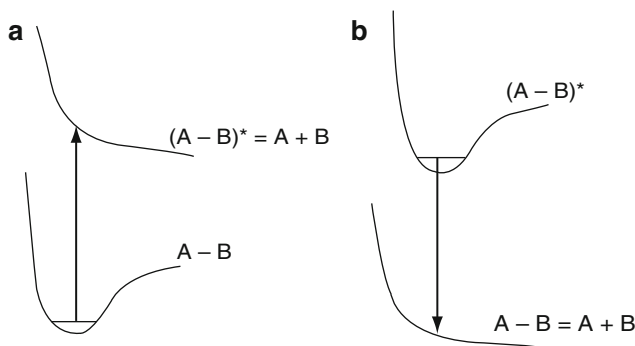


Fig. 23 (a) The nonexcited molecule (or a set of molecules; notice that A and B can be both atoms or molecules) is in a bound state; once excited, the molecules evolve to a dissociative state. (b) An excimer is a combination of atoms (or molecules) such that it is bound only in the excited state, and then it dissociates as soon as the system relaxes

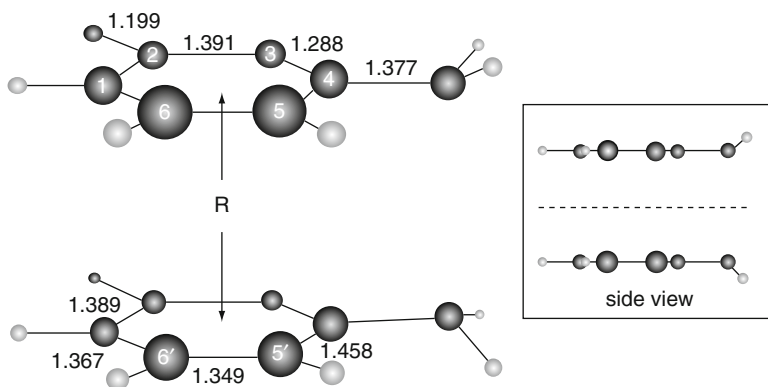


Fig. 24 Labeling for the cytosine dimer. Bond distances (in Å) correspond to the ground-state equilibrium geometry of the monomer. The homodimer system displays C_s symmetry, being the mirror symmetry plane represented by a *dashed line* in the side-view inset

with respect to the intermolecular separation (R) of two cytosine molecule kept at the ground-state geometries were built (see Fig. 24). The structure allows for an effective and natural interaction of two cytosine molecules in the biologically relevant *cis-syn* stereoisomer. The ground-state PEC is repulsive at the CASSCF level, whereas the lowest singlet excited state is weakly bound. In contrast, at the CASPT2 level (see Fig. 25), when dynamic correlation is taken into account, the ground-state S_0 and the three lowest singlet excited states (S_1 , S_2 , and S_3) have well defined minima with binding energies of a few tenths of an eV.

The Basis Set Superposition Error (BSSE) is corrected by means of the counterpoise procedure, an important aspect that will be discussed later. In the figure energies are referred to two ground-state cytosine molecules separated about 20 a.u. In the asymptotic limit, S_1 and S_2 become degenerate, which is consistent with the

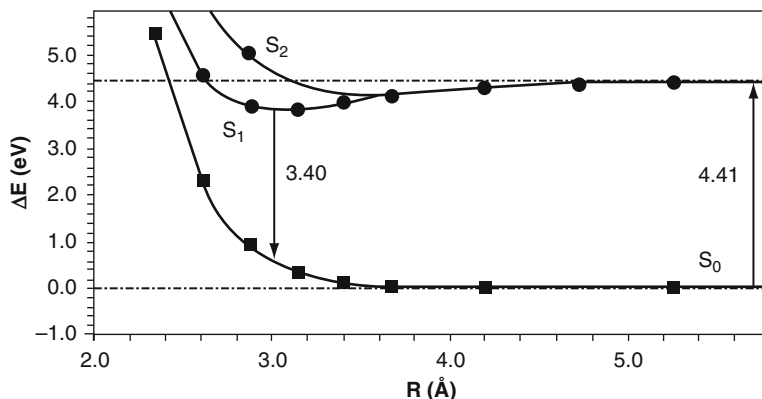


Fig. 25 BSSE-corrected CASPT2 (12,12) potential energy curves built with respect to the distance cytosine-cytosine arranged in a face-to-face disposition

fact that they are related to the equivalent situations $C + C^*$ and $C^* + C$, where C and C^* stand for the ground-state cytosine and its lowest singlet excited state, respectively. The absorption $S_0 \rightarrow S_1$ calculated at 20 a.u. (4.41 eV) corresponds to the monomer absorption. On the other hand, the vertical emission from the minimum of S_1 is calculated at 3.40 eV, and it can be considered the source of emission observed in DNA or oligomer samples. In particular, the fluorescence maximum was reported in aqueous solution for the dimer $d(C)_2$ and the 15-mer $d(C)_{15}$ at 3.22 eV, considerably red-shifted as compared to that of the monomer (3.96 eV).

Since that theoretical determination, the existence of excimers in DNA and their role as precursors for phenomena like the red-shifted DNA emission, charge transports in the DNA strand, the formation of nucleobase photoadducts, and, in general, the control of the decay dynamics of excited DNA has been determined both theoretically and experimentally. Crespo-Hernández et al. (2004), for instance, have shown by using femtosecond transient absorption spectroscopy that excimers are formed in high yields in a variety of synthetic DNA oligonucleotides and conclude that excited-state dynamics of A · T DNA is controlled by base stacking. CASPT2 calculations explained also the dynamics of the adenine dimer by the presence of the excimers.

As an illustration of bimolecular reactions, the formation of pyrimidine nucleobase adducts in the excited state will be presented here (see Roca-Sanjuán et al. 2008b; Serrano-Pérez et al. 2008d). Among the possible photoreactions that pyrimidine (Pyr) bases of nucleic acids may undergo upon ultraviolet (UV) irradiation, cyclobutane thymine dimers ($T < > T$ or CBT) formed by intrastrand adjacent thymine bases (see Fig. 26) constitute one of the major photoinduced lesions, particularly in cellular DNA, in spite of the direct repair by a light-activated mechanism (DNA photolyase) in which abnormal bonds are cleaved. Unrepaired or misrepaired thymine or cytosine dimers and the resultant mutations

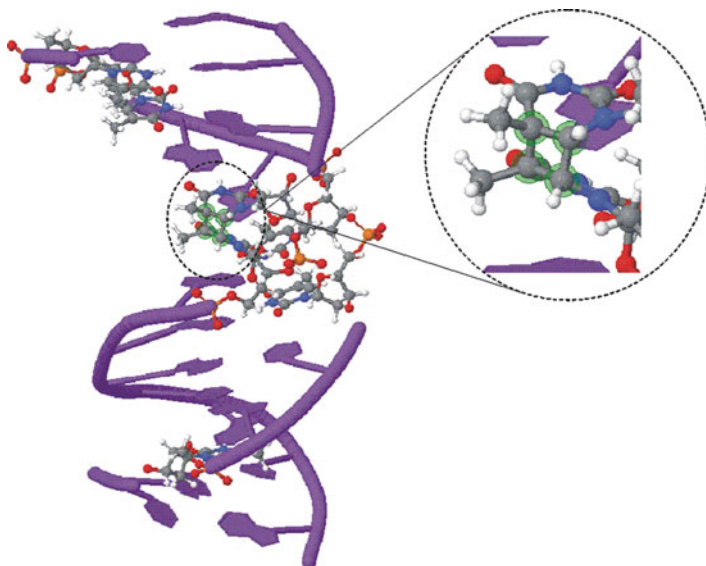


Fig. 26 Photocycloaddition between two adjacent molecules of thymine is a particular class of perycyclic [2 + 2] reaction allowed photochemically and not thermally. It is one of the most usual lesion on DNA. In this picture, taken from a structure obtained in the Protein Data Bank and visualized by Jmol program, the cyclobutane ring is highlighted

may well lead to the development of skin cancer. The major of Pyr < > Pyr photoproducts. The relatively smaller degree of flexibility of A-DNA compared to B-DNA to achieve the right orientations that become prone to react has been related to the greater resistance of A-DNA to Pyr < > Pyr formation.

Dimerization seems to occur only for thymine residues that are already in a reactive arrangement at the instant of excitation because the rate of formation by favorably oriented thymine pairs is much faster than the rate of orientation change. The dimerization reaction occurs both in the singlet and triplet manifolds (see Bosca et al. 2006; Cadet et al. 1992). Focusing on the singlet states, a CI must drive the ultrafast formation of the cycloadduct. Figure 27 displays an scheme of the energy levels of the cytosine dimer in its lowest-lying singlet states based on CASPT2(12,12)/ANO-S C,N,O[3s2p1d]/H[2s1p] results (BSSE corrected). The lines connecting the different solutions represent the different evolution paths followed by the system.

Initially, upon absorption of UV-light radiation in the 4.4 eV region, the system can undergo two types of processes. Nearly unstacked pairs will localize its excitation in one of the monomers and evolve in a ultrafast way toward the ground state via the CI of the monomer, located at 3.6 eV. Other pairs will form an excimer through the stacking interaction of the π clouds of the nucleobases, displaying much larger lifetimes, as it has been already determined experimentally. The binding interaction in the excimer will depend on the conformational arrangement of the

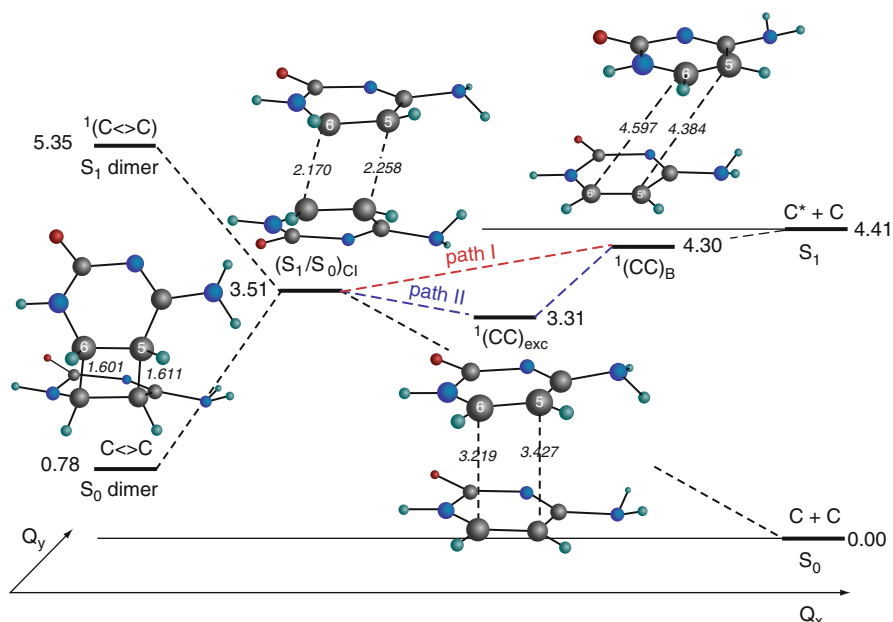


Fig. 27 Scheme based on CASPT2 calculations of the dimerization photochemistry of the cytosine dimer after UV-light absorption

nucleobase pair. Whereas the most common DNA confirmation, B-DNA, displays a very small interaction, strongly overlap situations like the face-to-face parallel arrangement yields the largest binding energy. The cytosine dimer has an excimer structure even lower than the CI leading the system to the formation of the ground state photoadduct CBT. This peculiarity and the presence of both S_0/S_1 CIs, that of the monomer and the one related to CBT, almost isoenergetic and therefore competitive, makes the yield of formation of CBT adducts much smaller than in the case of the thymine dimer, where the CI of the adduct is the lowest and easily accessible feature.

An Application Example of Energy Transfer: Psoralen + O_2

Photomedicine applies the principles of photobiology, photochemistry, and photo-physics to the diagnosis and therapy of diseases. One of the most active research areas in this field is photodynamic therapy (PDT), in which the affected living tissue is treated with a combination of a photosensitizer, activated by UV light, and molecular oxygen in its ground triplet state, $^3\Sigma_g^-$. Oxygen is present in the cellular environment ready to transform into singlet oxygen 1O_2 ($^1\Delta_g$), which is a strong electrophilic species that reacts with different compounds including some components of the cellular membrane causing cell death by apoptosis. An energy transfer (ET) process triggered by electronic coupling between a molecule in an excited state, the donor (D^*), and a molecule, the acceptor (A) or quencher within a

collision complex, is the mechanism through what the reaction takes place, a process that strongly depends on the inter-fragment distance. In general, ET processes, at large separation between the moieties (20–30 Å or even larger), the electronic coupling arises from the Coulomb interaction between electronic transitions that, under the dipole approximation, reduces to the known Förster's dipole–dipole coupling. The process is actually a nonradiative transfer of energy occurring whenever the emission spectrum of D overlaps with the absorption spectrum of A (although no intermediate photon takes part on it). It is the electric field around D*, behaving like a field generated by a classical oscillating dipole, the cause of the excitation of A. At larger separations than Förster's, fluorescence resonance ET (with photon emission by D* and subsequent absorption by A) becomes more efficient than excitation ET. At shorter interfragment distances, however, the so-called Dexter exchange coupling predominates, arising from the exchange integrals that account for the indistinguishability of the electrons in many-electron wave functions. This factor decreases steeply with separation. If the interaction is assumed weak and overlap between D* and A wave functions is produced, Fermi's Golden Rule for coupled transitions can be applied. Such processes have been studied theoretically in depth in recent years, in particular for singlet–singlet ET processes implying an exchange of electrons of the same spin but different energies, that is, the spin state of each fragment is conserved. In PDT, the actual mechanism is, on the other hand, an intermolecular triplet–triplet energy transfer (TET), that is, a process of exchanging both spin and energy between a pair or molecules or molecular fragments. These reactions are commonly used to efficiently populate the triplet states of many organic molecules.

TET processes can be therefore understood as two simultaneous ETs with spin exchange between the interacting fragments and it is similar to the Dexter coupling for singlet–singlet ET, in particular because, as it depends on an electron exchange mechanism, it only takes place at short donor–acceptor distances (<10 Å). In TET the Förster's mechanism will not contribute, because at short distances the dipole approximation breaks down and because the transitions are dipole forbidden. The electronic coupling is not the only key factor that determines the efficiency of the ET process, but also the resonance condition, that is, the energy available in the donor must be at least equal or higher than that required to populate the excited state of the acceptor. If this is the case, the process is usually controlled by diffusion and described as exothermic. In the opposite situation, that is, if the energy of the acceptor is lower than that of the donor, the process becomes thermally activated and lies in the endothermic region. That means that there is an energy barrier whose height will depend on the nature of the acceptor, either classical (for rigid systems) or nonclassical (flexible systems which might find conformations for efficient, non-vertical TET), with a corresponding larger or smaller, respectively, decay in the process rate.

Besides the reaction with DNA nucleobases, psoralen can also interact with molecular oxygen to exert its photosensitizing action. Singlet oxygen and other reactive species of oxygen induce photooxidation of lipids and are considered responsible for cell membrane damaging effects, causing also the appearance of

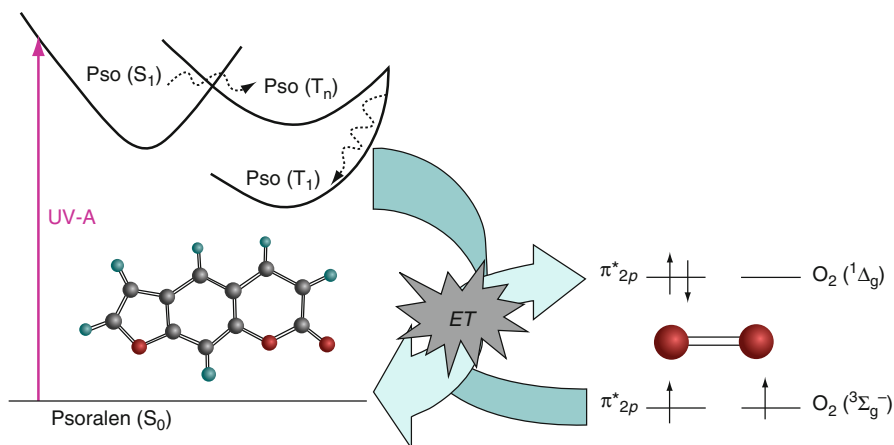
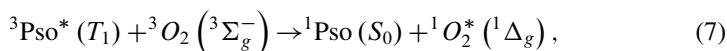


Fig. 28 Scheme of the oxygen-dependent PUVA mechanism

erythema and pigmentation activity in the human skin. In particular, the TET process taking place between psoralen and molecular oxygen is (see also Fig. 28):



where activated psoralen behaves as a donor in its lowest triplet state, and triplet ground-state oxygen is the acceptor. The lowest excited singlet state of molecular oxygen (${}^1\Delta_g$) is located at 0.97 eV. Furocoumarins (psoralen, 8-MOP, 5-MOP, khellin, TMP, and 3-CPS) have their lowest-lying triplet T_1 state energy at least 1.4 eV higher than the oxygen singlet state, what makes the TET exothermic and diffusion-controlled, with molecular oxygen behaving as a rigid, classical acceptor.

A recent study (see Serrano-Pérez et al. 2009) performed employing the CASPT2/ANO-LC,O[4s3p1d]/H[2s1p] methodology estimated the electronic coupling at some specific disposition of the moieties (psoralen and O_2). Looking for an appropriate arrangement yielding the most effective TET process is nontrivial and, in general, not even relevant, in particular in diffusion-controlled systems which may form a collision complex at short distances. It is important, however, to estimate reaction rates and lifetimes at different intermolecular distances. Furthermore, $\text{M}-\text{O}_2$ interaction potentials are very weak, and the potential surfaces are generally characterized by multiple shallow minima. Hence, it is necessary to consider different orientations when approaching M and O_2 through a basic interfragment coordinate, here the distance R. Previous studies on systems composed by two ethylene molecules (Et–Et), by the methaniminium cation and ethylene (MetN^+-Et), and by ethylene and molecular oxygen ($\text{Et}-\text{O}_2$), show that the face-to-face (FF) arrangement is the most appropriate orientation (see Fig. 29).

The geometries of both psoralen and molecular oxygen were kept fixed at the CASSCF optimized triplet excited (T_1) state structure and the triplet ground

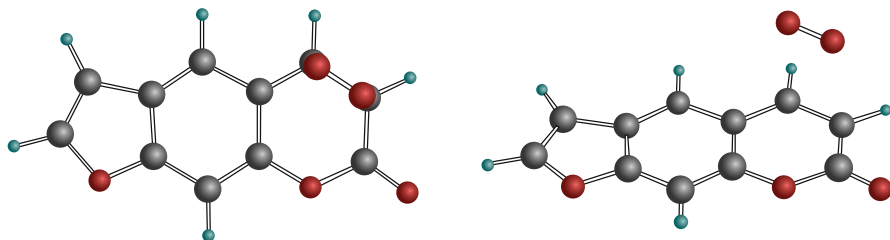


Fig. 29 Psoralen- O_2 supermolecule. The molecular oxygen was placed parallel to the reactive pyrone double bond of psoralen

($^3\Sigma_g^-$) state experimental geometry, respectively. The active space employed was 14 electrons/11 orbitals (8/7 located in the furocoumarin and 6/4 located in O_2). The active space was validated after comparing the results with previous findings in the isolated psoralen and control calculations on the oxygen molecule with larger active spaces and basis sets. The four lowest singlet states and the three lowest triplet states of the supermolecule were computed. The molecule behaves as a donor in its triplet state, and it is capable to transfer its energy to molecular oxygen in its triplet ground state to generate the singlet ground state furocoumarin and excited singlet oxygen ($^1\Delta_g$). The energy of the triplet state of the furocoumarin is much higher than the energy of the oxygen $^1\Delta_g$ state (computed 1.09 eV, experimental 0.97 eV), and the process falls clearly into the exothermic regime, expected to be controlled by diffusion. Figure 30 displays the potential energy curves of the lowest-lying singlet and triplet states of the supermolecule psoralen- O_2 in a FF arrangement with respect to the $C_3=C_4$ bond of the pyrone ring (the reactive bond in the triplet state, where the spin density is basically located at different intermolecular distances). The states of the supermolecule protagonist of the TET are 4^1A (T_1 of psoralen + $^3\Sigma_g^-$ of O_2), as initial energy level, and 1^1A (S_0 of psoralen + $^1\Delta_g$ of O_2), as the final outcome of the process in which both moieties have changed spin and energy. Within the present approach, the electronic coupling (H') is obtained as half the difference $|\epsilon_\infty - \epsilon_i|$, where ϵ_∞ and ϵ_i are the energy gaps between the states 4^1A and 1^1A at infinite distance (at 10 Å in the current computation) and at the different interfragment distances, respectively. In this way, the coupling represents the perturbation introduced in each state due to the interaction within the dimer. It is important to notice how advantageous that definition is, because the results are in practice independent of the number of states considered. Notice that the BSSE problem does not affect the results of the coupling, which uses energy differences between states computed at the same geometry. As a result, the BSSE correction vanishes when the energy gap is computed. In other cases, known as non-vertical TET systems, the geometry of the acceptor can be distorted to better accept the photosensitization from the donor. In such situations, the search for regions of state degeneracy with enhanced effectiveness for the ET transfer may be convenient. This is not the case for the present example, in which the acceptor is a rigid oxygen molecule and the process is basically diffusion controlled.

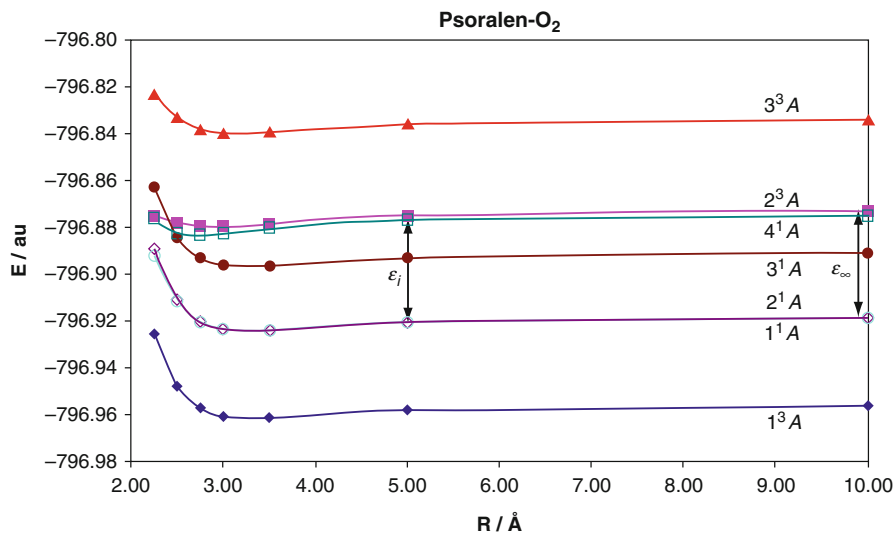


Fig. 30 Potential energy curves of the low-lying excited states of the supermolecule psoralen-molecular oxygen along the interfragment distance (R). The energy coupling H' is obtained as half of the energy difference $|\epsilon_\infty - \epsilon_i|$ between the initial 4^1A (T_1 of psoralen and ${}^3\Sigma_g^-$ of O_2) and final 1^1A (S_0 of psoralen and ${}^1\Delta_g$ of O_2) states of the supermolecule at infinite distance (ϵ_∞ , zero coupling situation) and at each of the distances (ϵ_i)

Table 6 ET analysis of Psoralen- O_2 system along the distance between the two moieties

$R/\text{\AA}$	H'/eV	$k_{\text{TET}}/\text{s}^{-1}$	$\tau_{\text{TET}}/\text{s}$
10.00	0.0000	—	—
5.00	0.0012	1.43×10^9	6.98×10^{-10}
3.50	0.0038	1.41×10^{10}	7.10×10^{-11}
3.00	0.0357	1.21×10^{12}	8.23×10^{-13}
2.75	0.0870	7.23×10^{12}	1.38×10^{-13}
2.50	0.1966	3.69×10^{13}	2.71×10^{-14}
2.25	0.3833	1.40×10^{14}	7.13×10^{-15}

Table 6 compiles the values of the electronic coupling computed at the different internuclear distances, related with the ET rates and lifetimes. In the weak coupling regime in which the electronic interaction is smaller than the vibrational reorganization energy, the rate for triplet–triplet energy transfer (k_{TET}), and the corresponding lifetime (τ_{TET}), between the donor and the acceptor can be estimated using Fermi's Golden Rule:

$$k_{\text{TET}} = \frac{1}{\tau_{\text{TET}}} = \frac{4\pi^2}{h} \left| \langle \psi_i | \hat{H} | \psi_j \rangle \right|^2 \rho_E = \frac{4\pi^2}{h} (H')^2 \rho_E, \quad (8)$$

where the matrix element of the Hamiltonian, H' , is the electronic part of the energy transfer (i.e., the electronic coupling) and ρ_E is the density of vibrational states in the

initial and final states and their spectral overlap. The inverse of the rate is the lifetime of energy transfer. To obtain the TET rates, we have taken values of $\rho_E = 200 \text{ eV}^{-1}$ and $(4\pi^2/h) = 9.554 \times 10^{15} \text{ eV}^{-1} \text{ s}^{-1}$. This order of magnitude for the value of the density of states was used previously as a good estimation in systems of this size.

The whole process of generation of singlet oxygen from psoralen does not only depend on the efficiency of the TET from the triplet state of the photosensitizer, but also on the rate of formation of the triplet state itself. As shown previously, in psoralen the crucial step to populate the triplet manifold in the gas phase is the intersystem crossing (ISC) process between the initially populated singlet $S_\pi (\pi\pi^*)$ state and the lowest-lying triplet $T_n (n\pi^*)$ state. The latter state evolves subsequently toward the lowest triplet $T_1 (\pi\pi^*)$ state via a corresponding (and essentially barrierless and ultrafast) internal conversion (IC).

In a similar manner as for Eq. 8, the estimation of the rate constant, here for nonradiative ISC (k_{ISC}), can be obtained as:

$$k_{\text{ISC}} = \frac{4\pi^2}{h} |H_{\text{SO}}|^2 \rho_E, \quad (9)$$

where H_{SO} stands for the spin-orbit coupling terms for the nonradiative transition $S_\pi (\pi\pi^*) \rightarrow T_n (n\pi^*)$. An estimated value of 200 eV^{-1} will be employed for ρ_E as used for psoralen in studies explicitly computing vibronic factors.

The ISC nonradiative process corresponds to the transfer $S_\pi (\pi\pi^*) \rightarrow T_n (n\pi^*)$ in each furocoumarin donor, a process which will take place more efficiently in the region of the (S_π/T_n) STC singlet-triplet crossing, and that subsequently will give rise to an ultrafast population of the lowest $T_\pi (\pi\pi^*)$ by internal conversion in the triplet manifold. The spin-orbit coupling factors near the crossing point between S_π and T_n , along the LIIC performed in a previous study, is 9.2 cm^{-1} (see Fig. 17). As a result, the k_{ISC} rate computed was $3.58 \times 10^9 \text{ s}^{-1}$. Additionally, to the H_{SO} strength, the presence of energy barriers in the potential energy hypersurfaces may strongly affect the value of the rate constants, which can be corrected using the Arrhenius exponential term in the framework of the transition state theory. As a qualitative estimation of these effects, a corrected ISC rate (k'_{ISC}) can be obtained from:

$$k'_{\text{ISC}} = k_{\text{ISC}} e^{-\Delta E/RT}, \quad (10)$$

where k_{ISC} is that computed from Eq. 9, ΔE is the energy of the barrier from the initial electronic singlet to the triplet state, R is the ideal gas constant, and T the temperature (298 K). In this particular context, ΔE will be estimated as the energy difference between the singlet $S_\pi (\pi\pi^*)$ state, populated at the Franck–Condon geometry, and the triplet $T_n (n\pi^*)$ state in the computed crossing point with the singlet state $(S_\pi/T_n)_X$, considering that the energy obtained initially to populate S_π will be available, at least, to surmount the barrier and populate T_n .

In a previous section, we considered as a barrier the energy difference between the minimum of the S_π state and the crossing point $(S_\pi/T_n)_X$, which is 0.36 eV, but the present approach accounts for the excess vibrational energy available in the

system from the initially populated S_π state (at the Franck–Condon geometry) to the crossing point (0.03 eV), which seems to be more relevant in this context, where the interest focuses on rate constants. The results were $1.11 \times 10^9 \text{ s}^{-1}$ and 0.90 ns for k'_{ISC} and τ'_{ISC} , respectively.

Within this methodology, we can compute the photosensitizing effectiveness, regarding the generation of singlet oxygen, of other furocoumarins and other photosensitizers in general, analyzing the effectiveness of populating T_π as well. Since molecules (donors) that share the same basic structure are known to possess almost the same electronic coupling (against a common acceptor), the first part of this study may well be more interesting if different families of compounds are compared. On the contrary, we should rely on the different way to populate the lowest-lying triplet excited state. In studies performed on the family of furocoumarins (psoralen, 8-MOP, 5-MOP, khellin, TMP, and 3-CPS), the latter, 3-carbethoxypsoralen, was found with the largest ISC rate.

Practical Aspects

The Basis Set Superposition Error (BSSE) for Excited States

The use of finite basis sets derives in a specific defect of the quantum-chemical calculation known as the Basis Set Superposition Error (BSSE). The majority of the contribution to the energy of a system comes from the internal electrons. If the basis set of an atom is deficient in the core region, a molecular method recovers a large amount of energy correcting this deficient area with the basis set of the other atoms. The BSSE is therefore related with the improper inclusion of the correlation energy in a quantum-chemical calculation. Although present in all practical cases, as BSSE strongly depends on the separation between the different centers, its effects only become crucial in dimers and aggregates, that is, when energies at large internuclear distances are compared to those at short distances. In general, the result of ignoring BSSE is both a shortening of bond lengths and an increasing of bond energies because the net effect is an increase of the energy in absolute value. This error is a purely mathematical artifact owing to the fact that the supermolecule possesses a larger basis set than the isolated monomers and as a result the potential energy surface is altered. There are some methods to correct this error like the counterpoise (CP) or the Chemical Hamiltonian Approach (CHA) procedures.

In the previous study on the formation of the excimer of the cytosinedimer, the correction of the BSSE effect was necessary. Figure 31 displays the same PECs as in Fig. 25, but excluding the BSSE correction. All states seem now to be bound, including the ground state, and compared to the BSSE-corrected results, the vertical emission takes place at 3.19 eV instead of 3.40 eV. Table 6 collects the numerical results on the system.

The CASPT2 ground-state binding energy in the uncorrected result is substantial, 0.62 eV, but the system becomes unbound by -0.15 eV when the BSSE is included, that is, the ground state dimer at 3.416 \AA is 0.15 eV above the sum of two ground-state monomers. The CP-BSSE corrections seem to be large. With inclusion of the BSSE, both S_1 and S_2 are bound ($CP - E_b$ is positive). Because of the cancelation

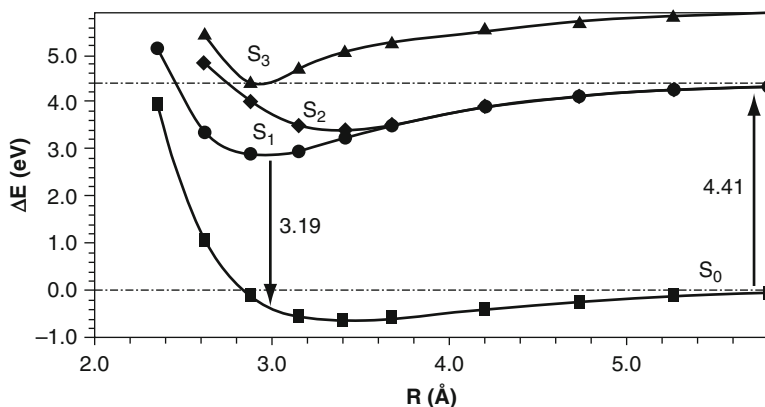


Fig. 31 CASPT2 (12,12) potential energy curves built with respect the distance cytosine–cytosine arranged in a face-to-face way without the inclusion of the CP correction for BSSE

Table 7 Binding energies (E_b), basis set superposition error (BSSE) obtained through the counterpoise method (CP-BSSE), and the corrected binding energy (CP- E_b), computed at the CASPT2 level

State	Geometry	R	E_b	CP-BSSE	CP- E_b
S_0	S_0	3.416	0.62	0.77	-0.15
S_1^a	S_1	2.954	1.51	0.97	0.54
S_0	S_1	2.954	0.29	0.97	-0.68
S_2	S_2	3.376	0.99	0.74	0.25

^aTheCASPT2 vertical emission (fluorescence), including the CP-BSSE correction leadsto 3.19 eV, as the result of (4.41 eV - 0.54 eV - 0.68 eV)

of BSSE corrections, the vertical emission remains like the direct CASPT2 result once that the BSSE has been taken into account.

The corrected minimum for S_1 is obtained at $R = 3.076 \text{ \AA}$, with a vertical emission of 3.40 eV and a binding energy CP- E_b of 0.58 eV. Thus, our best estimate has a difference of 0.18 eV with respect to the emission maximum datum obtained experimentally. It is worth recalling that the computed vertical transition does not have experimental counterpart, and for a truly correct comparison with experiment, vibrational resolution of the band should be computed in order to determine the band maximum (Table 7).

The CP procedure has been followed here to take into account the BSSE, a strategy that is known to slightly overestimate the effect. The method corrects the energy at each geometry by subtracting from the full energy of the supermolecule twice the difference between the energy of the fragment using its own basis set and the energy of the fragment using the full basis, but placing the basis functions of the opposite fragment in ghost centers, that is, at the same position as the atoms but with no atoms on them. This subtracted energy is what is known as the BSSE in the CP correction procedure. Obviously the error will decrease when the distance between the fragments decreases, therefore at large internuclear distances the BSSE vanishes. The uncorrected binding energies will then be larger than they should

because they are computed as differences between the energy of the minimum of the supermolecule (with large and stabilizing BSSE effects) and the summed energy of the independent fragments (with no BSSE). The CP method was developed for ground states, and it can be directly extended to excited states only in those cases in which the excitation is clearly localized in one of the fragments. For delocalized situations, it is not possible to define a consistent procedure. As observed in Table 6, however, BSSE values are the same for ground and excited states at a common geometry, for instance here, 0.97 eV both for the S_0 and S_1 states at the S_1 geometry. This is a systematic behavior, in particular for methods like CASPT2, in which a previous common set of State-Average MOs has been used for both states. Therefore, it can be concluded that it is safe to assign the same BSSE correction for ground and excited states at the same geometry, considering also that the inter-fragment distance is the most important parameter to modulate the extent of the BSSE. Let's finally add that ANO basis sets, large and flexible, produce larger BSSE effects than lower quality basis sets.

Computation of Electronic Couplings

The calculation of the electronic coupling matrix element H' is the crucial part in the determination of ET rates and lifetimes. The extent of the coupling controls the energy transfer process, specifically the passage from one state to another, and it can be taken as a measure of the efficiency of the ET process. Different procedures to estimate the ET coupling have been developed based on diabatic localized dimer calculations, monomer transition densities or transition dipole moments, and a supermolecule ansatz of the dimer, whereas generalization of such approaches to determine TET couplings are also available. From all procedures, an energy gap based method such as the supermolecule dimer approach, in which the value of the coupling is obtained as half of the splitting or perturbation between the interacting states, has been shown to be convenient and accurate, and it will be valid in the weak regime where the interaction is not strong. It is clear that its accuracy strongly relies on the quality of the quantum-chemical method used to perform the electronic structure calculations, something guaranteed by the highly reliable and accurate CASPT2 method. It has also the additional advantage that their results are not affected by the BSSE because they are obtained as energy differences at the different internuclear distances.

Conclusions

Quantum Mechanics helps in explaining natural phenomena from theoretical physics (particles, strings) and theoretical chemistry (chemical reactions, intermolecular forces) to the most complex theoretical biology. The complexity level increases as the simplicity of models decrease due to the growing number of variables to be dealt with and the difficulty in simulating the environment. The challenge lies in the ascertainment that life takes place into hierarchically structured matter (macromolecules, cells, tissues, organs, and entities), and it requires the

action of several physical properties of its constituent elements on the whole: the interactions among them and with the environment.

Interaction of light with tissues or, in general, living or nonliving entities is a complex phenomenon which lies on the borders among physics, chemistry, and biology. Specifically, the quantum chemistry of the excited state has experienced an outstanding development in the last decades. Improved algorithms and computational strategies, and obviously faster computers, have contributed to increasing the accuracy of the theoretical description of the photochemical phenomena. Although quantitative determinations in realistic biological systems are still far into the future, medium-size systems and their processes are already at reach of quantum-chemical models. Even more, state-of-the-art experimental techniques, such those related to femtosecond lasers, cannot be properly interpreted without the support of accurate calculations, in a way that a constructive interplay between theory and experiment shows particularly rewarding.

We are witnesses of how the walls which blocked our understanding of photochemical and photophysical processes are being slowly demolished. Step by step larger systems and more complex processes can be studied with accuracy. In fact, it has been theory which has provided modern photochemistry with the proper framework formed by the potential energy hypersurfaces, their interaction through conical intersections and states coupling, and the molecular evolution along such potentials. Unfortunately there are not many methods able to compute properly excited states, and the available ones are complex and sometimes of limited applicability. Multiconfigurational approaches (MRCI, MRPT, and MRCC) give the most general and unbiased description of all types of excitations and situations, either valence, Rydberg or anionic states, bound and dissociated situations, involving open and closed-shell ground states. Among the available approaches, the CASPT2 method has proved to better balance applicability and accuracy at a relatively low computational cost, although new developments and refinements are soon to come, like RASPT2. Single configurational methods (CC, TDDFT, CIS . . .) must be used with caution because they do not give a balanced treatment of all situations in excited-state chemistry, for instance, for degenerate cases like conical intersections. Apart from having proper methods, they must be provided with the required tools, like geometry optimizers, algorithms to search for conical intersections, to compute state couplings or trajectories for the nuclei. One-electron basis sets must be also selected with caution. The simultaneous calculation of valence and Rydberg states is necessary in those regions where both types of states are present, and properly solving valence-Rydberg mixing requires specific strategies. The calculation of anionic excited states requires their own cautions, such as checking the convergence of the resonance solutions by means of different stabilization methods. On the other hand, when dealing with finite one-electron functions and comparing fragment energies with bonded situations, the problem of BSSE arises. In such a situation, the BSSE corrections may well be very important and the phenomenon should be taken into account.

In summary, the excited state quantum chemistry cannot be undertaken routinely because of its complexity. Some of the tools available for the ground state are not

even developed yet for the excited state. In order to get accurate and predictive results, one has to calibrate carefully the computational procedure and determine the requirements in each situation. In this manner, an accuracy of 0.2–0.1 eV is currently expected for the excitation energies, almost an order of magnitude larger than previously. Undoubtedly, the development and refinement of quantum-chemical methods (the software) and computers (the hardware) are making easier the study of the interaction radiation-matter, a still unfinished task with fascinating challenges in the foreseeable future.

Acknowledgments The authors thank to all the members of the Quantum Chemistry for the Excited State (QCEXVal) group from the University of Valencia for their valuable contributions and to all their coworkers along the years. Financial support is acknowledged from projects CTQ2007-61260, CTQ2010-14892, and CSD2007-0010 Consolider-Ingenio in Molecular Nanoscience of the Spanish MICINN/FEDER.

Bibliography

- Almlöf, J., & Taylor, P. R. (1986). General contraction of Gaussian basis sets. I. Atomic natural orbitals for first- and second-row atoms. *The Journal of Chemical Physics*, *86*, 4070–4077.
- Almlöf, J., & Taylor, P. R. (1991). Atomic Natural Orbital (ANO) basis sets for quantum chemical calculations. *Advances in Quantum Chemistry*, *22*, 301–373.
- Andersson, K., Malmqvist, P.-Å., & Roos, B. O. (1992). Second-order perturbation theory with a complete active space self-consistent field reference function. *The Journal of Chemical Physics*, *96*, 1218–1226.
- Atchity, G. J., Xantheas, S. S., & Ruedenberg, K. (1991). Potential-energy surfaces near intersections. *The Journal of Chemical Physics*, *95*, 1862–1876.
- Atkins, P. W., & Friedman, R. S. (1997). *Molecular quantum mechanics* (3rd ed.). Oxford: Oxford University Press.
- Barbatti, M., Granucci, G., Persico, M., Ruckenbauer, M., Vazdar, M., Eckert-Maksić, M., & Lischka, H. (2007). The on-the-fly surface-hopping program system Newton-X: Application to ab initio simulation of the nonadiabatic photodynamics of benchmark systems. *Journal of Photochemistry and Photobiology A: Chemistry*, *190*, 228–240.
- Bearpark, M. J., & Robb, M. A. (2007). *Conical intersection species as reactive intermediates*. Hoboken: Wiley.
- Ben-Nun, M., & Martínez, T. J. (2000). Photodynamics of ethylene: Ab initio studies of conical intersections. *Chemical Physics*, *259*, 237–248.
- Bosca, F., Lhiaubet-Vallet, V., Cuquerella, M. C., Castell, J. V., & Miranda, M. A. (2006). The triplet energy of thymine in DNA. *Journal of the American Chemical Society*, *128*, 6318–6319.
- Buenker, R. J., Hirsch, G., Li, Y., Gu, J., Alekseyev, A. B., Liebermann, H., & Kimura, M. (2000). *Ab initio calculations of excited state potential functions* (pp. 135–168). Chichester: Wiley.
- Cadet, J., Anselmino, C., Douki, T., & Voituriez, L. (1992). New trends in photobiology: Photochemistry of nucleic acids in cells. *Journal of Photochemistry and Photobiology, B: Biology*, *15*, 277–298.
- Celani, P., & Werner, H.-J. (2003). Analytical energy gradients for internally contracted second-order multireference perturbation theory. *The Journal of Chemical Physics*, *119*, 5044–5057.
- Crespo-Hernández, C., Cohen, B., Hare, P., & Kohler, B. (2004). Ultrafast excited-state dynamics in nucleic acids. *Chemical Reviews*, *104*, 1977–2019.
- Domcke, W., Sobolewski, A. L., & Woywood, C. (1993). Internal conversion funnel in benzene and pyrazine: Adiabatic and diabatic representation. *Chemical Physics Letters*, *203*, 220–226.

- Domcke, W., Yarkony, D. R., & Köppel, H. (Eds.). (2004). *Conical intersections. Electronic structure, dynamics and spectroscopy*. River Edge: World Scientific.
- Eisinger, J., & Shulman, R. G. (1968). Excited electronic states of DNA. *Science*, *161*, 1311–1319.
- Finley, J., Malmqvist, P.-Å., Roos, B. O., & Serrano-Andrés, L. (1998). The multi-state CASPT2 method. *Chemical Physics Letters*, *288*, 299–306.
- Finley, J., Malmqvist, P.-Å., Roos, B. O., & Serrano-Andrés, L. (1999). The multi-state CASPT2 method. *Chemical Physics Letters*, *288*, 299–306.
- Foresman, J. B., Head-Gordon, M., Pople, J. A., & Frisch, M. J. (1992). Toward a systematic molecular orbital theory for excited states. *The Journal of Physical Chemistry*, *96*, 135–149.
- Forsberg, N., & Malmqvist, P.-Å. (1997). Multiconfiguration perturbation theory with imaginary level shift. *Chemical Physics Letters*, *274*, 196–204.
- Ghigo, G., Roos, B. O., & Malmqvist, P.-Å. (2004). A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chemical Physics Letters*, *396*, 142–149.
- Goddard, W. A., & Hunt, W. J. (1974). The Rydberg nature and assignments of excited states of the water molecule. *Chemical Physics Letters*, *24*, 464–471.
- Grein, F. (2009). Coupled cluster and density functional studies on geometries and energies of excited C_{2v} states of ozone. *The Journal of Chemical Physics*, *130*, 124118.
- Grimme, S., & Waletzke, M. (1999). A combination of Kohn–Sham density functional theory and multi-reference configuration interaction methods. *The Journal of Chemical Physics*, *111*, 5645.
- Helgaker, T., Jørgensen, J., & Olsen, J. (2004). *Molecular electronic-structure theory*. Chichester: Wiley.
- Herzberg, G., & Longuet-Higgins, L. C. (1963). Intersection of potential energy surfaces in polyatomic molecules. *Discussions of the Faraday Society*, *35*, 77–82.
- Kleinschmidt, M., Tatchen, J., & Marian, C. M. (2001). Spin-orbit coupling of DFT/MRCI wave-functions: Method, test calculations, and application to thiophene. *Journal of Computational Chemistry*, *23*, 824–833.
- Koch, H., & Jørgensen, P. (1990). Coupled cluster response functions. *The Journal of Chemical Physics*, *93*, 3333–3344.
- Kowalski, K. (2005). Completely renormalized EOMCCSD(T) method employing independent optimization of the cluster product terms. *Chemical Physics Letters*, *411*, 306–310.
- Lischka, H., Dallos, M., Szalay, P. G., Yarkoni, D. R., & Shepard, R. (2004). Analytic evaluation of nonadiabatic coupling terms at the MR-CI level. I. Formalism. *The Journal of Chemical Physics*, *120*, 7322–7329.
- Malmqvist, P.-Å., & Roos, B. O. (1989). The CASSCF state interaction method. *Chemical Physics Letters*, *155*, 189–194.
- Malmqvist, P.-Å., Rendell, A., & Roos, B. O. (1990). The restricted active space self-consistent-field method, implemented with a split graph unitary group approach. *The Journal of Physical Chemistry*, *94*, 5477–5482.
- Malmqvist, P.-Å., Pierloot, K., Shahi, A. R. M., Cramer, C. J., & Gagliardi, L. (2008). The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO_2 and Cu_2O_2 systems. *The Journal of Chemical Physics*, *128*, 204109.
- Martin, J. M. L., Taylor, P. R., & Lee, T. J. (1997). The harmonic frequencies of benzene. A case for atomic natural orbital basis sets. *Chemical Physics Letters*, *275*, 414–422.
- Merchán, M., & Serrano-Andrés, L. (2005). *Ab initio methods for excited states*. Amsterdam: Elsevier.
- Merchán, M., Serrano-Andrés, L., Robb, M., & Blancafort, L. (2005). Triplet-state formation along the ultrafast decay of excited singlet cytosine. *Journal of the American Chemical Society*, *127*, 1820–1825.
- Merchán, M., González-Luque, R., Climent, T., Serrano-Andrés, L., Rodríguez, E., Reguero, M., & Peláez, D. (2006). Unified model for the ultrafast decay of pyrimidine nucleobases. *The Journal of Physical Chemistry B*, *110*, 26471–26476.

- Müller, T. H., Dallos, M., & Lischka, M. (1999). The ethylene 1^1B_{1u} V state revisited. *The Journal of Chemical Physics*, *110*, 7176–7184.
- Oddershede, J. (1987). Propagator methods. *Accounts of Chemical Physics*, *69*, 201–239.
- Olaso-González, G., Roca-Sanjuán, D., Serrano-Andrés, L., & Merchán, M. (2006). Toward the understanding of DNA fluorescence: The singlet excimer of cytosine. *The Journal of Chemical Physics*, *125*, 231102.
- Olsen, J., Roos, B. O., Jørgensen, P., & Jensen, H. J. A. (1988). Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces. *The Journal of Chemical Physics*, *89*, 2185–2192.
- Paschoal, D., Costa, M. F., Junqueira, G. M. A., & Santos, H. F. D. (2009). Ab initio calculation of electric properties for the BH, CO, CS and N₂ molecules. *Journal of Molecular Structure: THEOCHEM*, *913*, 200–206.
- Paterson, M. J., Bearpark, M. J., Robb, M. A., & Blancafort, L. (2004). The curvature of the conical intersection seam: An approximate second-order analysis. *The Journal of Chemical Physics*, *121*, 11562–11571.
- Peach, M. J. G., Benfield, P., Helgaker, T., & Tozer, D. J. (2008). Excitation energies in density functional theory: An evaluation and a diagnostic test. *The Journal of Chemical Physics*, *128*, 044118.
- Pou-Américo, R., Serrano-Andrés, L., Merchán, M., Ortí, E., & Forsberg, N. (2000). A theoretical determination of the low-lying electronic states of the *p*-benzosemiquinone radical anion. *Journal of the American Chemical Society*, *122*, 6067–6077.
- Pérez-Hernández, G., González, L., & Serrano-Andrés, L. (2008). Rydberg or valence? The long-standing question in the UV absorption spectrum of 1,1'-bicyclohexylidene. *ChemPhysChem*, *9*, 2544–2549.
- Robb, M. A., Bernardi, F., & Olivucci, M. (1996). Conical intersections as a mechanistic feature of organic-photochemistry. *Pure and Applied Chemistry*, *67*, 783–789.
- Roca-Sanjuán, D., Rubio, M., Merchán, M., & Serrano-Andrés, L. (2006). Ab initio determination of the ionization potentials of DNA and RNA nucleobases. *The Journal of Chemical Physics*, *125*, 084302.
- Roca-Sanjuán, D., Merchán, M., Serrano-Andrés, L., & Rubio, M. (2008a). Ab initio determination of the electron affinities of DNA and RNA nucleobases. *The Journal of Chemical Physics*, *129*, 095104.
- Roca-Sanjuán, D., Olaso-González, G., González-Ramírez, I., Serrano-Andrés, L., & Merchán, M. (2008b). Molecular basis of DNA photodimerization: Intrinsic production of cyclobutane cytosine dimers. *Journal of the American Chemical Society*, *130*, 10768–10779.
- Roos, B. O., Taylor, P. R., & Siegbahn, P. E. M. (1979). A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chemical Physics*, *48*, 157–173.
- Roos, B. O., Andersson, K., Fülischer, M. P., Serrano-Andrés, L., Pierloot, K., Merchán, M., & Molina, V. (1996). Applications of level shift corrected perturbation theory in electronic spectroscopy. *Journal of Molecular Structure: THEOCHEM*, *388*, 257–276.
- Roos, B. O., Malmqvist, P.-Å., Molina, V., Serrano-Andrés, L., & Merchán, M. (2002). Theoretical characterization of the lowest-energy absorption band of pyrrole. *The Journal of Chemical Physics*, *116*, 7526–7537.
- Roos, B. O., Lindh, R., Malmqvist, P.-Å., & Widmark, P.-O. (2005). New relativistic ANO basis sets for actinide atoms. *Chemical Physics Letters*, *409*, 295–299.
- Roothaan, C. C. J. (1951). New developments in molecular orbital theory. *Reviews of Modern Physics*, *23*, 69–89.
- Rubio, M., Serrano-Andrés, L., & Merchán, M. (2008). Excited states of the water molecule: Analysis of the valence and Rydberg character. *The Journal of Chemical Physics*, *128*, 104305.
- Schreier, W. J., Schrader, T. E., Soller, F. O., Gilch, P., Crespo-Hernández, C. E., Swaminathan, V. N., Carell, T., Zinth, W., & Kohler, B. (2007). Thymine dimerization in DNA is an ultrafast photoreaction. *Science*, *315*, 625–629.

- Serrano-Andrés, L., & Merchán, M. (2005). Quantum chemistry of the excited state: 2005 overview. *Journal of Molecular Structure: THEOCHEM*, 729, 99–108.
- Serrano-Andrés, L., & Merchán, M. (2009). Are the five natural DNA/RNA base monomers a good choice from natural selection? A photochemical perspective. *Journal of Photochemistry and Photobiology, C: Photochemistry Reviews*, 10, 21–32.
- Serrano-Andrés, L., Merchán, M., & Lindh, R. (2005). Computation of conical intersections by using perturbation techniques. *The Journal of Chemical Physics*, 122, 104107.
- Serrano-Andrés, L., Merchán, M., & Borin, A. C. (2006). Adenine and 2-aminopurine: Paradigms of modern theoretical photochemistry. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 8691–8696.
- Serrano-Andrés, L., Merchán, M., & Borin, A. C. (2008). A three-state model for the photophysics of guanine. *Journal of the American Chemical Society*, 130, 2473–2484.
- Serrano-Pérez, J., Serrano-Andrés, L., & Merchán, M. (2006). A theoretical insight into the photophysics of psoralen. *The Journal of Chemical Physics*, 124, 124502.
- Serrano-Pérez, J., Serrano-Andrés, L., & Merchán, M. (2007a). Quantum chemical study on the population of the lowest triplet state of psoralen. *Chemical Physics Letters*, 434, 107–110.
- Serrano-Pérez, J., González-Luque, R., Serrano-Andrés, L., & Merchán, M. (2007b). On the intrinsic population of the lowest triplet state of thymine. *The Journal of Physical Chemistry*, 111, 11880–11883.
- Serrano-Pérez, J., Merchán, M., & Serrano-Andrés, L. (2008a). Photosensitization and phototherapy with furocoumarins: A quantum-chemical study. *Chemical Physics*, 347, 422–435.
- Serrano-Pérez, J., González-Luque, R., Merchán, M., & Serrano-Andrés, L. (2008b). The family of furocoumarins: Looking for the best photosensitizer for phototherapy. *Journal of Photochemistry and Photobiology, A: Chemistry*, 199, 34–41.
- Serrano-Pérez, J., Merchán, M., & Serrano-Andrés, L. (2008c). Photoreactivity of furocoumarins and DNA in PUVA therapy: Formation of psoralen-thymine adducts. *The Journal of Physical Chemistry B*, 112, 14002–14010.
- Serrano-Pérez, J., González-Ramírez, I., Coto, P., Serrano-Andrés, L., & Merchán, M. (2008d). Theoretical insight into the intrinsic ultrafast formation of cyclobutane pyrimidine dimers in UV-irradiated DNA: Thymine versus cytosine. *The Journal of Physical Chemistry B*, 112, 14096–14098.
- Serrano-Pérez, J., Olaso-González, G., Merchán, M., & Serrano-Andrés, L. (2009). Singlet oxygen generation in PUVA therapy studied using electronic structure calculations. *Chemical Physics*, 360, 85–96.
- Shepard, R., Lischka, H., Szalay, P. G., Kovar, T., & Ernzerhof, M. (1992). A general multireference configuration interaction gradient program. *The Journal of Chemical Physics*, 96, 2085–2098.
- Sicilia, F., Bearpark, M. J., Blancafort, L., & Robb, M. A. (2007). An analytical second-order description of the S0/S1 intersection seam: Fulvene revisited. *Theoretical Chemistry Accounts*, 118, 241–251.
- Strickler, S. J., & Berg, R. A. (1962). Relationship between absorption intensity and fluorescence lifetime of molecules. *The Journal of Chemical Physics*, 37, 814–822.
- Szabo, A., & Ostlund, N. S. (1996). *Modern quantum chemistry: Introduction to advanced electronic structure theory*. Mineola: Dover Publications.
- Szymczak, J. J., Barbatti, M., Soo Hoo, J. T., Adkins, J. A., Windus, T.-L., Nachtigallova, D., & Lischka, H. (2009). Photodynamics simulations of thymine: Relaxation into the first excited singlet state. *The Journal of Physical Chemistry, A*, 113, 12686–12693.
- Taylor, P. R. (2007). *Lecture notes in quantum chemistry: European summer school*. Berlin: Springer.
- Teller, E. (1937). The crossing of potential surfaces. *The Journal of Physical Chemistry*, 41, 109–116.
- Widmark, P.-O., Malmqvist, P.-Å., & Roos, B. O. (1990). Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. *Theoretica Chimica Acta*, 77, 291–306.

Gerald Monard and Jean-Louis Rivail

Contents

Introduction: Importance of Solvent Effects in Chemistry	728
The Supermolecule Model	728
Continuum Models	729
The Full Quantum Approaches of Liquids	732
The Quantum Mechanical/Molecular Mechanical (QM/MM) Models	733
Other Statistical Approaches of the Solvent	734
Reference Interaction Site Model (RISM)	734
Langevin Dipole Model (LD)	735
The Future of Studies in the Liquid State	735
Bibliography	735

Abstract

The properties of a molecule may change quite substantially when passing from the isolated state to a solution, and computational chemistry requires the possibility of taking into account the effects of a solvent on molecular properties. These changes are mainly due to long range interactions, and electrostatics involving a large number of solvent molecules play the major role in the phenomenon and free energy changes have to be evaluated. Statistical calculations by means of usual Monte Carlo or molecular dynamics coupled with a full quantum chemical description of a sample representative of the solution is still out of reach for standard molecular modeling computations nowadays. Nevertheless, several simplified approaches are available to evaluate the free energy changes which appear when an isolated molecule, as described by standard quantum computations, undergoes the influence of a solvent and to predict the changes in

G. Monard (✉) • J.-L. Rivail

Theoretical Chemistry and Biochemistry Group SRSMC, Nancy-University CNRS Boulevard des Aiguillettes, Vandoeuvre-les-Nancy, France

e-mail: gerald.monard@cbt.uhp-nancy.fr; Jean-Louis.Rivail@cbt.uhp-nancy.fr

the molecular properties which are the consequences of solvation. In this chapter, we develop the principles of the most usual methods that a computational chemist can find in standard codes or can implement more or less easily to approach the solvent effects in quantum chemistry investigations.

Introduction: Importance of Solvent Effects in Chemistry

Most of the chemical reactions as well as experimental structure determinations are performed in solutions. On the contrary, usual quantum chemical computations usually deal with isolated chemical species. This may lead to erroneous conclusions. For instance, the addition of bromine to an ethylenic hydrocarbon is known for having a different mechanism in the gas phase and in solution. In this example, the velocity constant vary by a factor of 10^{10} when going from carbon tetrachloride to water as a solvent (Reichardt 1979), although the mechanism is the same. These features are confirmed by appropriate quantum chemical computations which show that the transition state of ethylene-bromine would be dissymmetric and 55 kcal/mol above the van der Waals complex in the case of the isolated species (Yamabe et al. 1988), while with a simple simulation of the solvent effect one finds a symmetric transition state lying 30.79 kcal/mol above the van der Waals complex in a non dipolar solvent and 0.02 kcal/mol in water (Assfeld 1994).

A full quantum mechanical treatment of a sample containing the solute and a number of solvent molecules large enough to simulate the solution is still out of reach of usual quantum chemical computations not only because the system of interest should be pretty large but also because the free energy computations would require a statistical treatment on millions of configurations of this system. One is therefore led, in general, to use model systems, simplified expressions of interatomic interactions allowing either statistical simulations or the derivation of molecular distribution functions or some semi-empirical approaches. In this vade mecum, we summarized different ways of modeling solvent effects in quantum chemistry. This includes the supermolecule model, the continuum and semi-macroscopic models, the full quantum description of liquids, the QM/MM methods (see Fig. 1), and the use of analytical molecular distribution functions (Hirata 2003). All these approaches have their limits, and one is led to choose one or the other according to the kind of problem being addressed.

The Supermolecule Model

In some cases, in particular that of aqueous solutions, the specific interactions may be modeled by means of a supermolecule made of an aggregate of the solute and a limited number of solvent molecules treated as an isolated single molecular system at the desired quantum chemical level. This approach may give useful information on strong intermolecular interactions (Catak et al. 2009; Harb et al. 2004) but cannot model correctly the effect of the solvent because the long-range electrostatic

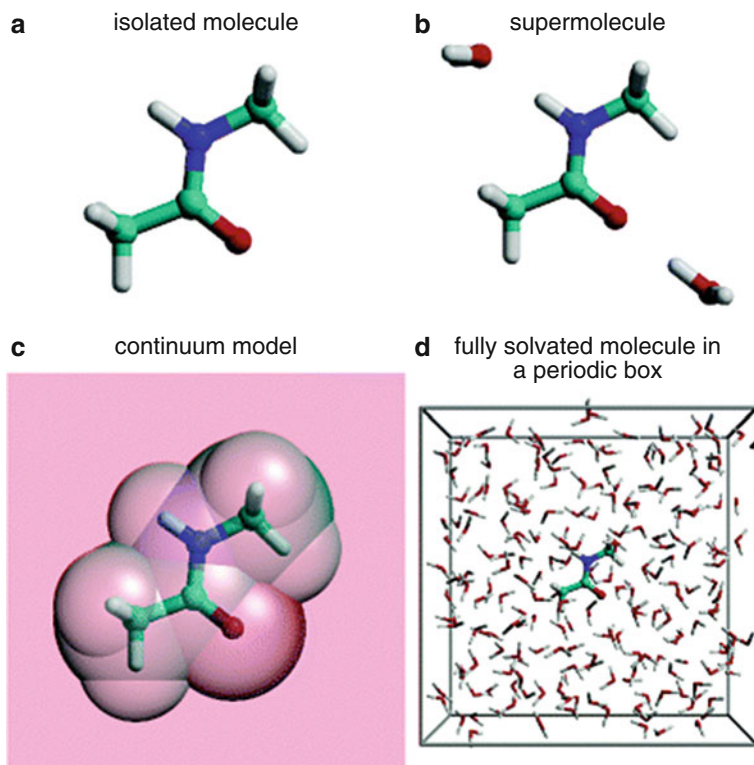


Fig. 1 The different approaches for modeling a solute molecule in a solvent. From *top left to bottom right*: (a) the isolated molecule (gas phase); (b) the supermolecule (solute with a few solvent molecules); (c) the continuum approach (the molecule is inserted in a cavity surrounded by a dielectric continuum); (d) the fully solvated system (the solute is inserted in a periodic box filled with solvent molecules)

interactions, which play an important role in a liquid (Buckingham 1967, 1978), are not taken into account. An efficient way to model these long-range electrostatic interactions is to use a continuum model.

Continuum Models

These models are by far the most widely used in current computational chemistry nowadays, and several review articles give an overview of them (Rivail and Rinaldi 1996; Tomasi 2004; Tomasi et al. 2005). They are based upon the consideration of the solvent as a continuous medium characterized by its macroscopic properties: bulk dielectric permittivity, possibly surface tension. The solute is assumed to occupy a cavity surrounded by this continuum. The free energy of interaction

between this molecule and its surrounding may be analyzed as the sum of three contributions:

- The free energy variation of the system when the cavity is created in the bulk of the liquid. This can be considered as an increase of its free area and is a positive free energy variation.
- The dispersion-repulsion energy arising between a polarizable solute and the solvent molecules of finite size, which is always a negative contribution.
- The electrostatic free energy arising from the polarization of the solvent by the charge distribution of the solute. The result of this polarization is a non zero electrostatic potential within the cavity which perturbs the charge distribution of the solute, leading to changes in the solute's structure, compared with the isolated molecule (Luque et al. 2003).

The earlier attempt to approach the electrostatic contribution to the free energy of solvation is due to Kirkwood (1934). This model is based on a multipole expansion of the charge distribution of the solute at the center of a spherical cavity surrounded by a continuum represented by the dielectric permittivity of the solvent. When this expansion is limited to rank 1 which corresponds to a pure dipole μ , one finds the Onsager model (Onsager 1936) in which the electrostatic contribution to the free energy of solvation by a solvent of dielectric constant ε of a molecule having a dipole moment μ in a cavity of radius a takes the expression:

$$\Delta G = -\frac{1}{2a^3} \frac{2(\varepsilon - 1)}{2\varepsilon + 1} \mu^2. \quad (1)$$

Onsager has shown, in the case of a pure liquid made of dipolar, polarizable molecules, that the dielectric constant of the liquid is fairly well predicted when the volume of the cavity is equal to the apparent molecular volume in the liquid. This finding gives a hint to the estimate of the volume of the cavity which is a parameter in the Kirkwood model. In the case of a non neutral solute, the charge q corresponds to the multipole of rank 0 and the corresponding term in the multipole expansion is the well-known Born formula (Born 1920) giving the free energy of solvation of charged solutes:

$$\Delta G = -\frac{1}{a} \frac{\varepsilon - 1}{\varepsilon} q^2. \quad (2)$$

This approach which has the advantage of giving an analytical formula of the free energy of solvation has been extended to ellipsoidal (Rinaldi and Rivail 1973) and spheroidal cavities (Rivail and Terryn 1982) and has been applied long ago to quantum chemical computations, at various levels of approximation (Rinaldi et al. 1983). As the perturbation to the wavefunction depends on the charge distribution of the solute which is defined by the wavefunction to be computed, the computational scheme requires self consistency combined with the usual self

consistent computational algorithms for quantum chemistry. This is frequently referred as the Self Consistent Reaction Field (SCRF) where the solute is polarized by the solvent which is, in return, polarized by the solute. Nowadays, this model has been extended to cavities of any shape using a multicentric multipole expansion on all the atoms of the solute (Rinaldi et al. 2004). In all these approaches, the first and second derivatives of energy are computed allowing a full geometry optimization and the determination of the stationary points in the potential energy surface and the computation of the harmonic vibrational frequencies (Rinaldi et al. 2006).

A closely related approach, using the same kind of molecular cavity has been developed by Tomasi et al. (1981). In this model, the multipole expansion is replaced by a numerical computation of the electrostatic potential inside the cavity, due to the polarization of the boundary by the solute. This model also has been improved along the years. It has various versions all known under the acronym of PCM (Cancès et al. 1997) which are implemented in the Gaussian suites of programs, and a closely related model is available in the Jaguar package (Marten et al. 1996; Tannor et al. 1994).

Rather, different approaches are based on the so-called generalized Born model in which the charge distribution of the solute is represented by point charges on all atoms and the solute–solvent interaction is defined by a sum of the Born contributions of each atom in which the reciprocal of the sphere radius in Eq. 2 is replaced by a parameterized empirical formula. The most elaborated applications of this model is found in the series of SM x models of Cramer and Truhlar ($x = 1$ to 8 nowadays, corresponding to successive improvements) (Cramer and Truhlar 1991, 2008).

Finally, to simplify the electrostatic algorithm some methods start with a dielectric permittivity $\varepsilon = \infty$ corresponding to a conductor like solvent and then correct the result either by the ratio $2(\varepsilon - 1) / (2\varepsilon + 1)$ of the Onsager factor of the actual dielectric permittivity of the solvent to the limiting value when $\varepsilon = \infty$, or by the ratio of the Born factors: $(\varepsilon - 1) / \varepsilon$. It is noteworthy that the difference may not be negligible for low dielectric constants. This approximation has first been introduced in the COSMO model (Klamt and Schüürmann 1993). A version of the PCM model, known as C-PCM (Barone and Cossi 1998) uses the same approximation to evaluate the electrostatic solvation term.

The importance of any kind of solvents, including the non polar ones has to be emphasized because some people claim that the properties in a non polar solvent should be close to the gas state. The experimental results mentioned in (Reichardt 1979) show that it is not true, and this can also be seen in the case of the Born or Onsager models. If one considers a solvent such as cyclohexane with a dielectric permittivity close to 2, one sees that the electrostatic free energy of solvation is for the former and the latter model respectively 50 % and 40 % of the limiting value at infinite permittivity. Similarly, one also notices that for high dielectric permittivities, these terms do not vary noticeably when this constant varies, justifying the COSMO-type models in particular in the case of water and highly polar solvents.

The quantitative use of continuum models requires care. The absolute values of free energies of solvation depend on many parameters. The electrostatic contribution

mainly depends on the shape and volume of the cavity. The most widely used shapes are defined by the “molecular surface” (Pascual-Ahuir and Silla 1990; Silla et al. 1991) defined after the atomic radii (Bondi 1964) multiplied by a factor of the order of 1.3. A test for the choice of this parameter can be the comparison of the volume of the cavity with the apparent molecular volume of the solute (when known) in the liquid state or in the solution like in the Onsager’s theory of the dielectric permittivity of pure polar liquids.

The non electrostatic contributions are far more difficult to model. Usually, they are evaluated by empirical formulae which, in the most favorable cases, introduce some error bars. The models parameterized on a large number of experimental data, such as SM8 (Cramer and Truhlar 1991, 2008), are probably rather reliable although one may deal with a case which is an exception. Conversely, variations of solvation energy for systems in which the non electrostatic terms are assumed not to vary significantly, such as conformational equilibria, may be approached rather safely by means of these computational efficient models.

The Full Quantum Approaches of Liquids

A reliable computation of free energies in a molecular system requires the use of statistical mechanics on a sample which can be considered as macroscopically representative. In the case of solutions, this requires more than one hundred solvent molecules and the computation of an enormous number of different configurations. Full quantum chemical studies are now possible thanks to some specific algorithms but they are still lengthy. An efficient approach uses a standard semi-empirical code, reparameterized to account properly for intermolecular energies (Bernal-Uruchurtu and Ruiz-López 2000; Bernal-Uruchurtu et al. 2000), coupled with a linear scaling divide and conquer algorithm (Dixon and Merz 1996, 1997), which allows the simulation of a set of 216 water molecules in a quite affordable computer time (Monard et al. 2005). The use of fast DFT methods can now replace the semi-empirical ones (Hu et al. 2007) and one may anticipate that in the near future, the theoretical and algorithmic improvements will make possible the use of non empirical treatments. The use of plane wave basis sets (Todorova et al. 2006; VandeVondele et al. 2005) is another attempt to improve the efficiency of the codes.

In contrast with the previous approaches in which the energy and the forces calculated at each configuration of the system are computed within the usual Born-Oppenheimer approximation (BOMD), the Car-Parrinello unified approach for molecular dynamics and DFT (Car and Parrinello 1985) is an efficient alternative to the simulation of macroscopic samples although it requires rather large computational facilities and time. The CPMD package is available for such computations which require some expertise to be ran safely (Marx and Hutter 2000). A comparison of CPMD and BOMD simulations may be useful to determine the best computational procedures (Kuo et al. 2006). Car-Parrinello computations are widely applied to the simulation of liquids, mainly water (Izvekov and Voth 2002) but other liquids are also considered (Bakó et al. 2006). All these full QM approaches are, for

the time being, mainly limited to the understanding of the structural properties of pure liquids.

The Quantum Mechanical/Molecular Mechanical (QM/MM) Models

Chemical studies usually deal with a solute which can be a single molecule or a molecular complex or transition state in a chemical reaction. In such systems, the role of the solvent is mainly a physical perturbation which can be simulated at a lower theoretical level than that required for the study of the subsystem of chemical interest. The success of continuum models confirms this statement. In order to describe the solution at the molecular level and to perform full statistical mechanics computations on a model of macroscopic sample, one may set up some computationally efficient approaches by limiting the quantum chemical study to the solute and using one of the usual classical force-fields to represent the solvent molecules. The computation of the statistical averages can be done by means of either Monte Carlo or molecular dynamics algorithms. The so-called QM/MM models are now widely used in such chemical studies.

The Hamiltonian is a sum of three terms: H_{QM} , the usual quantum Hamiltonian of the solute, H_{MM} , the classical Hamiltonian corresponding to the configuration of the solvent $H_{QM/MM}$, which corresponds to the interaction energy between the classical solvent molecules and the quantum solute (Monard et al. 2003). The first two Hamiltonians are well defined by the choice of the quantum methodology and the classical force field chosen to model the liquid. The definition of the interaction Hamiltonian is far more difficult. It is usually made of a sum of van der Waals interaction energies between the classical atoms of the solvent and the quantum atoms of the solute, the long-range electrostatic interactions between the classical solvent molecules and the electron distribution and the nuclei of the solute and the possible specific interactions (hydrogen bonds). This is the most delicate part of the methodology, which requires specific parameterizations for each kind of quantum chemical method used as well as each classical force field. A large number of papers have been published on this problem (Freindorf and Gao 1996; Freindorf et al. 2005; Giese and York 2007; Luque et al. 2000; Riccardi et al. 2004). For specific problems, useful methodologies have been set up to determine the most appropriate parameters (Martín et al. 2002).

Once the methodology has been defined, running a computation requires the usual molecular dynamics or Monte Carlo codes, but the necessity of taking into account a large number of configurations makes any improvements which accelerate the process very valuable. One of the time consuming step is the computation of the long-range electrostatic interactions for which some algorithms, such as the use of Ewald sum (Nam et al. 2005) or periodic boundary conditions (Laino et al. 2006), may become very efficient to compute accurate free energies (Gao et al. 2005).

Various methods and levels of approximation can be used for the QM part of the studies, from semi-empirical (Cummins and Gready 1997; Geerke et al. 2008),

empirical Valence Bond (Sumner and Iyengar 2008) to correlated levels (Kongsted et al. 2003; Woods et al. 2008). QM/MM simulations also work with the Car-Parrinello methodology (Laio et al. 2002).

QM/MM methods have now reached a good level of achievement and have opened the way to refined studies of molecular structure and reactivity in the liquid state (Hu and Yang 2009). However, may be one of the last remaining challenges for QM/MM methods is to correctly and efficiently handle reactivity problems where one or more solvent molecules are directly implied in the chemical mechanism. The usual solution found in literature is to choose among all solvent molecules the unique molecule which will react with the solute. Some solutions which take into account the dynamical nature of the first solvation shell have been suggested so far (Kerdcharoen and Morokuma 2002, 2003; Tongraar et al. 1998; Tongraar and Rode 2004), but they have not been applied to reactivity problems yet.

Other Statistical Approaches of the Solvent

Reference Interaction Site Model (RISM)

Integral equations of the liquid state provide us, in principle, with a rigorous statistical description of the liquid, although, to solve them, one is forced to use several approximations, in particular various closure conditions to get pair correlation functions (Hansen and McDonald 1976). These approaches have been successfully used to describe simple liquids. In the case of polyatomic molecules, the difficulty increases rapidly with the number of sites and the nature of the interatomic potential. Nonetheless, a Reference Interaction Site Model (RISM) has been proposed to extend the treatment to polyatomic liquids (Chandler and Andersen 1972) and has been successfully applied to several problems dealing with molecular solutions (Hirata and Rossky 1982). Some routines are now available to solve the RISM equations. These equations have also been coupled with a quantum computation of the solute (Ten-no et al. 1993). The interatomic potentials have the standard Lennard-Jones expression in which the electrostatic interactions are developed on a basis of point charges. Therefore, the electronic distribution of the solute is limited to effective partial charges q_λ located on the atoms (denoted by a greek index here λ). The Hartree-Fock equations are modified by introducing a perturbative potential on each atom λ of the solute:

$$V_\lambda = \rho \sum_{j \in S} q_j \int_0^\infty r^{-1} g_{\lambda j}(r, q_\lambda) 4\pi r^2 dr, \quad (3)$$

where ρ is the solvent density and $g_{\lambda j}(r, q_\lambda)$ is the radial distribution function of the j th site of solvent molecules S relative to the λ th site of the solute molecule. This function is given by the RISM equations. The RISM-SCF cycle simultaneously optimizes the pair correlation functions and the electronic wavefunctions. This

approach has been successfully applied to the study of solvent effects on the electronic spectra of carbonyl compounds (Ten-no et al. 1994).

Langevin Dipole Model (LD)

Between a macroscopic description of the solvent represented by the continuum models and a microscopic approach where all solvent atoms are explicitly represented, there exists an alternative semi-macroscopic approach: the Langevin Dipole model. This model treats the solvent (usually water) explicitly but in a simplified way. It represents the time average polarization of the solvent molecules by a cubic grid of polarizable dipoles (Warshel 1991; Warshel and Russel 1984). According to the Langevin equation (Langevin 1905), the electrostatic field of the solute reorients and polarizes the grid point dipoles. While this method has been successfully applied to the computation of the free energy of solvation of small neutral and ionic molecular solutes at a quantum level (Florián and Warshel 1997, 1999; Kongsted et al. 2009), the Langevin Dipole model has been mostly used to represent the electrostatic interactions between water as a solvent and proteins through the Protein Dipoles/Langevin Dipoles (PDL) method (Roca et al. 2007; Warshel et al. 2006).

The Future of Studies in the Liquid State

The number of selected papers dealing with various approaches of the liquid state and solutions clearly indicate that taking into account the influence of the solvent in quantum chemical studies will soon become quite usual. The large variety of methods allows us to adapt the level of modeling to the accuracy required to get useful chemical information. The methodology is still expected to improve in the near future. In particular, the full quantum treatment of a solution which takes advantage of the advances in quantum modeling of pure liquids seems to be at hand, especially when one can use different quantum levels for the solute and for the solvent. Like in the case of QM/MM the key feature is the proper treatment of the interaction between both subsystems but one can be confident to the quantum chemists to find the proper solutions to this problem.

Bibliography

- Assfeld, X. (1994). PhD dissertation, Université Henri Poincaré, Nancy.
- Bakó, I., Hutter, J., & Pálincás, G. (2006). Car-Parrinello molecular dynamics simulation of liquid formic acid. *Journal of Physical Chemistry A*, 110, 2188.
- Barone, V., & Cossi, M. (1998). Quantum calculation of molecular energies and energy gradients in solution by a conductor solvent model. *Journal of Physical Chemistry A*, 102, 1995.
- Bernal-Uruchurtu, M. I., & Ruiz-López, M. F. (2000). Basic ideas for the correction of semiempirical methods describing H-bonded systems. *Chemical Physics Letters*, 330, 118.

- Bernal-Uruchurtu, M. I., Martins-Costa, M. T. C., Millot, C., & Ruiz-López, M. F. (2000). Improving description of hydrogen bonds at the semiempirical level: Water-water interactions as test case. *Journal of Computational Chemistry*, *21*, 572.
- Bondi, A. (1964). van der Waals volumes and radii. *Journal of Physical Chemistry*, *68*, 441.
- Born, M. (1920). Volumen und hydrationswärme der Ionen. *Zeitschrift für Physik*, *1*, 45.
- Buckingham, A. (1967). Permanent and induced molecular moments and long-range intermolecular forces. *Advances in Chemical Physics*, *12*, 107.
- Buckingham, A. (1978). Basic theory of intermolecular forces: Applications to small molecules. In B. Pullman (Ed.), *Intermolecular interactions, from diatomics to biopolymers* (Vol. 1, p. 1). New York: Wiley.
- Cancès, E., Menucci, B., & Tomasi, J. (1997). A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *Journal of Chemical Physics*, *107*, 3032.
- Car, R., & Parrinello, M. (1985). Unified approach for molecular dynamics and density-functional theory. *Physical Review Letters*, *55*, 2471.
- Catak, S., Monard, G., Aviyente, V., & Ruiz-López, M. F. (2009). Deamidation of asparagine residues: Direct hydrolysis versus succinimide-mediated deamidation mechanisms. *Journal of Physical Chemistry A*, *113*, 1111.
- Chandler, D., & Andersen, H. C. (1972). Optimized cluster expansions for classical fluids. II. Theory of molecular liquids. *Journal of Chemical Physics*, *57*, 1930.
- Cramer, C. J., & Truhlar, D. G. (1991). Molecular orbital theory calculations of aqueous solvation effects on chemical equilibria. *Journal of the American Chemical Society*, *113*, 8552.
- Cramer, C. J., & Truhlar, D. G. (2008). A universal approach to solvation modeling. *Accounts of Chemical Research*, *41*, 760.
- Cummins, P. L., & Gready, J. E. (1997). Coupled semiempirical molecular orbital and molecular mechanics model (QM/MM) for organic molecules in aqueous solution. *Journal of Computational Chemistry*, *18*, 1496.
- Dixon, S. L., & Merz, K. M., Jr. (1996). Semiempirical molecular orbital calculations with linear system size scaling. *Journal of Chemical Physics*, *104*, 6643.
- Dixon, S. L., & Merz, K. M., Jr. (1997). Fast, accurate semiempirical molecular orbital calculations for macromolecules. *Journal of Chemical Physics*, *107*, 879.
- Florián, J., & Warshel, A. (1997). Langevin dipoles model for ab initio calculations of chemical processes in solution: Parametrization and application to hydration free energies of neutral and ionic solutes and conformational analysis in aqueous solution. *Journal of Physical Chemistry B*, *101*, 5583.
- Florián, J., & Warshel, A. (1999). Calculations of hydration entropies of hydrophobic, polar, and ionic solutes in the framework of the langevin dipoles solvation model. *Journal of Physical Chemistry B*, *103*, 10282.
- Freindorf, M., & Gao, J. (1996). Optimization of the Lennard-Jones parameters for a combined ab initio quantum mechanical and molecular mechanical potential using the 3-21G basis set. *Journal of Computational Chemistry*, *17*, 386.
- Freindorf, M., Shao, Y., Furlani, T. R., & Kong, J. (2005). Lennard-Jones parameters for the combined QM/MM method using the B3LYP/6-31G*/AMBER potential. *Journal of Computational Chemistry*, *26*, 1270.
- Gao, D., Svoronos, P., Wong, P. K., Maddalena, D., Hwang, J., & Walker, H. (2005). pK_a of acetate in water: A computational study. *Journal of Physical Chemistry A*, *109*, 10776.
- Geerke, D. P., Thiel, S., Thiel, W., & van Gasteren, W. F. (2008). QM-MM interactions in simulations of liquid water using combined semi-empirical/classical Hamiltonians. *Physical Chemistry Chemical Physics*, *10*, 297.
- Giese, T. J., & York, D. M. (2007). Charge-dependent model for many-body polarization, exchange, and dispersion interactions in hybrid quantum mechanical/molecular mechanical calculations. *Journal of Chemical Physics*, *127*, 194101.
- Hansen, J. P., & McDonald, I. R. (1976). *Theory of simple liquids*. London: Academic.

- Harb, W., Bernal-Uruchurtu, M., & Ruiz-López, M. (2004). An improved semiempirical method for hydrated systems. *Theoretical Chemistry Accounts*, *112*, 204.
- Hirata, F. (2003). *Molecular theory of solvation*. Dordrecht: Kluwer.
- Hirata, F., & Rossky, P. J. (1982). Application of an extended RISM equation to dipolar and quadrupolar fluids. *Journal of Chemical Physics*, *77*, 509.
- Hu, H., & Yang, W. (2009). Development and application of ab initio QM/MM methods for mechanistic simulation of reactions in solution and in enzymes. *Journal of Molecular Structure: THEOCHEM*, *898*, 17.
- Hu, H., Lu, Z., Elsner, M., Hermans, J., & Yang, W. (2007). Simulating water with the self-consistent-charge density functional tight binding method: From molecular clusters to the liquid state. *Journal of Physical Chemistry A*, *111*, 5685.
- Izvekov, S., & Voth, G. A. (2002). Car-Parrinellomoleculardynamicsimulation of liquid water: New results. *Journal of Chemical Physics*, *116*, 10372.
- Kerdcharoen, T., & Morokuma, K. (2002). ONIOM-XS: An extension of the ONIOM method for molecular simulation in condensed phase. *Chemical Physics Letters*, *355*, 257.
- Kerdcharoen, T., & Morokuma, K. (2003). Combined quantum mechanics and molecular mechanics simulation of Ca^{2+} /ammonia solution based on the ONIOM-XS method: Octahedral coordination and implication to biology. *Journal of Chemical Physics*, *118*, 8856.
- Kirkwood, J. (1934). Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *Journal of Chemical Physics*, *2*, 351.
- Klamt, A., & Schüürmann, G. (1993). COSMO – A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of the Chemical Society, Perkin Transactions*, *2*, 799.
- Kongsted, J., Osted, A., & Mikkelsen, K. V. (2003). Coupled cluster/molecular mechanics method: Implementation and application to liquid water. *Journal of Physical Chemistry A*, *107*, 2578.
- Kongsted, J., Söderhjelm, P., & Ryde, U. (2009). How accurate are continuum solvation models for drug-like molecules? *The Journal of Computer-Aided Molecular Design*, *23*, 395.
- Kuo, I.-F. W., Mundy, C. J., McGrath, M. J., & Siepmann, J. I. (2006). Time-dependent properties of liquid water: A comparison of Car-Parrinello and Born-Oppenheimer molecular dynamics simulations. *Journal of Chemical Theory and Computation*, *2*, 1274.
- Laino, T., Mohamed, F., Laio, A., & Parrinello, M. (2006). An efficient linear-Scaling electrostatic coupling for treating periodic boundary conditions in QM/MM simulations. *Journal of Chemical Theory and Computation*, *2*, 1370.
- Laio, A., VandeVondele, J., & Rothlisberger, U. (2002). A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *Journal of Chemical Physics*, *116*, 6941.
- Langevin, P. (1905). Magnétisme et Théorie des électrons. *Annales de Chimie-Physique*, *8*, 70.
- Luque, F. J., Reuter, N., Cartier, A., & Ruiz-López, M. F. (2000). Calibration of the quantum/classical hamiltonian in semiempirical QM/MM AM1 and PM3 methods. *Journal of Physical Chemistry A*, *104*, 10923.
- Luque, F., Curutchet, C., Muñoz-Muriedas, J., Bidon-Chanal, A., Sorietas, I., Morreale, A., Gelpi, J. L., & Orozco, M. (2003). Continuum solvation models: Dissecting the free energy of solvation. *Physical Chemistry Chemical Physics*, *5*, 3827.
- Marten, B., Kim, K., Cortis, C., Friesner, R. A., Murphy, R. B., Rignalda, M. N., Sitkoff, D., & Honig, B. (1996). New model for calculation of solvation free energies: Correction of self-consistent reaction field continuum dielectric theory for short-range hydrogen-bonding effects. *Journal of Physical Chemistry*, *100*, 11775.
- Martín, M. E., Aguilar, M. A., Chalmet, S., & Ruiz-López, M. F. (2002). An iterative procedure to determine Lennard-Jones parameters for their use in quantum mechanics/molecular mechanics liquid state simulations. *Chemical Physics*, *284*, 607.
- Marx, D., & Hutter, J. (2000). Ab initio molecular dynamics: Theory and Implementation. In J. Grotendorst (Ed.), *Modern methods and algorithms of quantum chemistry* (p. 301). Jülich: John von Neumann Institute for Computing, NIC Series.

- Miertus, S., Scrocco, E., & Tomasi, J. (1981). Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects. *Chemical Physics*, 55, 117.
- Monard, G., Prat-Resina, X., González-Lafont, A., & Lluch, J. M. (2003). Determination of enzymatic reaction pathways using QM/MM methods. *International Journal of Quantum Chemistry*, 93, 229.
- Monard, G., Bernal-Uruchurtu, M. I., van der Vaart, A., Merz, K. M., Jr., & Ruiz-López, M. F. (2005). Simulation of liquid water using semiempirical hamiltonians and the divide and conquer approach. *Journal of Physical Chemistry A*, 109, 3425.
- Nam, K., Gao, J., & York, D. M. (2005). An efficient linear-scaling ewald method for long-range electrostatic interactions in combined QM/MM calculations. *Journal of Chemical Theory and Computation*, 1, 2.
- Onsager, L. (1936). Electric moments of molecules in liquids. *Journal of the American Chemical Society*, 58, 1486.
- Pascual-Ahuir, J. L., & Silla, E. J. (1990). GEPOL: An improved description of molecular surfaces. I. Building the spherical surface set. *Journal of Computational Chemistry*, 11, 1047.
- Reichardt, C. (1979). *Solvent effects in organic chemistry*. Weinheim/New York: Verlag Chemie.
- Riccardi, D., Li, G., & Cui, Q. (2004). Importance of van der Waals interactions in QM/MM simulations. *Journal of Physical Chemistry B*, 108, 6467.
- Rinaldi, D., & Rivail, J. L. (1973). Polarisabilites moléculaires et effet diélectrique de milieu í l'état liquide. Étude théorique de la molécule d'eau et de ses dimères. *Theoretica Chimica Acta*, 32, 57.
- Rinaldi, D., Ruiz-López, M. F., & Rivail, J. L. (1983). Ab initio SCF calculations on electrostatically solvated molecules using a deformable three axes ellipsoidal cavity. *Journal of Chemical Physics*, 78, 834.
- Rinaldi, D., Bouchy, A., Rivail, J. L., & Dillet, V. (2004). A self-consistent reaction field model of solvation using distributed multipoles. I. Energy and energy derivatives. *Journal of Chemical Physics*, 120, 2343.
- Rinaldi, D., Bouchy, A., & Rivail, J. L. (2006). A self-consistent reaction field model of solvation using distributed multipoles. II: Second energy derivatives and application to vibrational spectra. *Theoretical Chemistry Accounts*, 116, 664.
- Rivail, J., & Rinaldi, D. (1996). Liquid-state quantum chemistry: Computational applications of the polarizable continuum models. In J. Leszczynski (Ed.), *Computational chemistry, review of current trends* (p. 139). Singapore: World Scientific Publishing.
- Rivail, J. L., & Terryn, B. (1982). Free-energy of an electric charge distribution separated from an infinite dielectric medium by a 3 axes ellipsoidal cavity – Application to the study of molecular solvation. *Journal of Chemical Physics*, 79, 1.
- Roca, M., Messer, B., & Warshel, A. (2007). Electrostatic contributions to protein stability and folding energy. *FEBS Letters*, 581, 2065.
- Silla, E. J., Tuñón, I., & Pascual-Ahuir, J. L. (1991). GEPOL: An improved description of molecular surfaces II. Computing the molecular area and volume. *Journal of Computational Chemistry*, 12, 1077.
- Sumner, I., & Iyengar, S. S. (2008). Combining quantum wavepacket ab initio molecular dynamics with QM/MM and QM/QM techniques: Implementation blending ONIOM and empirical valence bond theory. *Journal of Chemical Physics*, 129, 054109.
- Tannor, D. J., Marten, B., Friesner, R. M. R. A., Sitkoff, D., Nicholls, A., Rignalda, M., Goddard, W., & Honig, B. (1994). Accurate first principles calculation of molecular charge distributions and solvation energies from Ab Initio quantum mechanics and continuum dielectric theory. *Journal of American Chemical Society*, 116, 11875.
- Ten-no, S., Hirata, F., & Kato, S. (1993). A hybrid approach for the solvent effect on the electronic structure of a solute based on the RISM and Hartree-Fock equations. *Chemical Physics Letters*, 214, 391.

- Ten-no, S., Hirata, F., & Kato, S. (1994). Reference interaction site model self-consistent field study for solvation effect on carbonyl compounds in aqueous solution. *Journal of Chemical Physics*, *100*, 7443.
- Todorova, T., Seitsonen, A. P., Hutter, J., Kuo, I.-F. W., & Mundy, C. J. (2006). Molecular dynamics simulation of liquid water: Hybrid density functionals. *Journal of Physical Chemistry B*, *110*, 3685.
- Tomasi, J. (2004). Thirty years of continuum solvation chemistry: A review, and prospects for the near future. *Theoretical Chemistry Accounts*, *112*, 184.
- Tomasi, J., Mennucci, B., & Cammi, R. (2005). Quantum mechanical continuum solvation models. *Chemical Reviews*, *105*, 2999.
- Tongraar, A., & Rode, B. M. (2004). Dynamical properties of water molecules in the hydration shells of Na^+ and K^+ : Ab initio QM/MM molecular dynamics simulations. *Chemical Physics Letters*, *385*, 378.
- Tongraar, A., Liedl, K. R., & Rode, B. M. (1998). Born-Oppenheimer ab initio QM/MM dynamics simulations of Na^+ and K^+ in water: From structure making to structure breaking effects. *Journal of Physical Chemistry A*, *102*, 10340.
- VandeVondele, J., Krack, M., Mohamed, F., Parrinello, M., Chassaing, T., & Hutter, J. (2005). Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Computer Physics Communications*, *167*, 103.
- Warshel, A. (1991). *Computer modeling of chemical reactions in enzymes and solutions*. New York: Wiley.
- Warshel, A., & Russel, S. T. (1984). Calculations of electrostatic interactions in biological systems and in solutions. *Quarterly Review of Biophysics*, *17*, 283.
- Warshel, A., Sharma, P. K., Kato, M., & Parson, W. W. (2006). Modeling electrostatic effects in proteins. *Biochimica et Biophysica Acta*, *1764*, 1647.
- Woods, C. J., Manby, F. R., & Mulholland, A. J. (2008). An efficient method for the calculation of quantum mechanics/molecular mechanics free energies. *Journal of Chemical Physics*, *128*, 014109.
- Yamabe, S., Minato, T., & Inagaki, S. (1988). Ab initio structures of transition-states in electrophilic addition-reactions of molecular halogens with ethene. *Journal of the Chemical Society, Chemical Communications*, *532*, 35.

Miroslav Medved', Šimon Budzák, Wojciech Bartkowiak, and Heribert Reis

Contents

Introduction	743
Definition of Macroscopic and Microscopic Quantities	744
Solute–Solvent Interactions	750
Interaction-Induced Properties	751
Distance Dependence	753
Many-Body Effects	754
Splitting Schemes	756
Implicit Solvent Methods	759
Polarizable Continuum Model (PCM) Approaches	761
Multipole Expansion (MPE) Method	763
Explicit Molecular Models	764
Supermolecular (SM) Calculations	767
QM/Langevin Dipole/Monte Carlo (QM/LD/MC) Method	769
Rigorous Local Field (RLF) Method	771
Discrete Reaction Field (DRF) Method	773
Polarizable Embedding (PE)	777
Frozen-Density Embedding (FDE)	779
Distributed Multipole Moments and (Hyper)polarizabilities	782

M. Medved' • Š. Budzák (✉)

Department of Chemistry, Faculty of Natural Sciences, Matej Bel University, Banská Bystrica,
Slovak Republic

e-mail: Miroslav.Medved@umb.sk; Simon.Budzak@umb.sk

W. Bartkowiak

Department of Physical and Quantum Chemistry, Wrocław University of Technology, Wrocław,
Poland

e-mail: wojciech.bartkowiak@pwr.wroc.pl

H. Reis

Institute of Biology, Medicinal Chemistry and Biotechnology, National Hellenic Research
Foundation, Athens, Greece

e-mail: hreis@eie.gr

Two Case Studies: Liquid Water and <i>p</i> -Nitroaniline in 1,4-Dioxane	786
Water	786
EFISH of <i>p</i> -Nitroaniline in 1,4-Dioxane	788
Bibliography	789

Abstract

Theoretical background and applications of methods for the calculation of solvent effects on molecular electric properties are reviewed. Macroscopic linear and nonlinear susceptibilities are defined, and their relationship to microscopic properties, i.e., (hyper)polarizabilities, is described. The role of specific intermolecular interactions in property calculations is demonstrated in terms of interaction-induced properties. Two categories of models for the description of solvent effects are presented. The first category comprises continuum models, in which the solvent is described as a homogeneous medium with the solute molecule located inside a cavity. Within this category, the polarizable continuum model and the multipole expansion method are described in more detail. The second group of approaches is based on discrete solvent models in which the solvent molecules are explicitly considered. Selected representative methods such as the supermolecular approach, polarizable embedding, and frozen-density embedding are presented. The possibility to combine explicit and implicit methods is demonstrated on the discrete reaction field and the Langevin dipole approaches. Finally, two illustrative examples (liquid water and *p*-nitroaniline in 1,4-dioxane) of the application of the presented methods for (non)linear property calculations are given.

Glossary

AIM	Atoms in molecules approach
ALDA	Adiabatic local density approximation
BSSE	Basis set superposition error
CAMM	Cumulative atomic multipole moments
COSMO-RS	Conductor-like screening model for real solvents
DIM	Dipole interaction model
DMA	Distributed multipole analysis
DRF	Discrete reaction field
DSA	Differential shell approach
EFISH	Electric field induced second harmonic generation
FDE	Frozen density embedding
FF	Force field
FF/FDE	Finite field/frozen density embedding
GSAM	Global search algorithm of minima
HRS	Hyper Rayleigh scattering

IDRI	Intensity dependent refractive index
KS	Kohn Sham
LIM	Localized dipole interaction model
MBP	Many-body corrected molecular polarizabilities
MC	Monte Carlo
MM	Molecular mechanics
MPE	Multipole expansion method
ND-FF	Numerical derivative using finite fields
NDO	Neglect of differential overlap
NLO	Nonlinear optical (properties)
pABA	<i>p</i> -aminobenzoic acid
pNA	<i>p</i> -nitroaniline
QM	Quantum mechanics, quantum mechanical
QM/LD/MC	Quantum mechanics/Langevin dipole/Monte Carlo
QM/MC	Quantum mechanics/Monte Carlo
QM/MM	Quantum mechanics/Molecular mechanics
PCM	Polarizable continuum model
PE	Polarizable embedding
PPM	Polarizability per molecule
RI	Resolution of identity
RLF	Rigorous local field
SAPT	Symmetry adapted perturbation theory
SHG	Second harmonic generation
SM	Supermolecule, supermolecular
SMx	Solvation model
SOPPA	Second order polarization propagator approximation
SCRf	Self-consistent reaction field
TCNE	Tetracyanoethylene
THG	Third harmonic generation

Introduction

The electronic and optical response properties of materials represent important characteristics playing a crucial role in many practical applications in (opto)electronics (e.g., optical information processing, telecommunication, data storage, FETs, photovoltaic cells, LEDs), chemical analysis, and bioimaging (e.g., multiphoton fluorescence microscopy). In order to understand the response of materials in various (non)linear optical and electro-optical processes, the concepts relating the macroscopic response properties (susceptibilities) to molecular (hyper)polarizabilities are introduced. Computational techniques for the calculation of electric properties of isolated molecules are already quite advanced, as shown, e.g., by Jaszuński et al. (2015) in this handbook. However, for most applications as well as for the experimental determination of these properties, condensed phases are much

more relevant. Accurate methods to compute the corresponding properties of condensed phases, the electric susceptibilities, are less developed than those for isolated species. The effect of the environment on molecular electric properties is correspondingly to some extent unknown. For solid phases, such as molecular crystals or surfaces, it is often a good approximation to consider a rigid structure of the phase, possibly undergoing harmonic vibrations. The dynamic nature of the liquid phase is much more pronounced, but its effect on the electric properties is difficult to model accurately. Thus, it is still not well understood and is under active investigation.

In this chapter, we present a few selected models which can be applied to obtain electric susceptibilities of liquids, or the electric properties of a molecule in solution, which is not the same, as we will show below. The selection is far from complete and was made mainly to present the many different aspects of the problem, which may have to be considered for a successful calculation.

We start our overview with introducing the applied terminology, notation, and definition of macroscopic and microscopic electric properties. Then, we describe the most important aspects of solute–solvent interactions together with the concept of interaction-induced properties which are relevant for the solvent models outlined in subsequent sections. The models for the description of solvent effects can be divided into two groups. In continuum models the solvent is described as a structureless homogeneous medium that is characterized by its macroscopic dielectric properties. The solute molecule is located inside a cavity in this medium, and its properties are treated at a chosen quantum-mechanical (QM) level of theory. In these models, the averaging over different solvent configurations as well as molecular polarization effects is implicitly taken into account. Approaches which are physically more appropriate rely on discrete solvent models and are presented in section “Explicit Molecular Models,” together with selected combined discrete/continuum methods.

Definition of Macroscopic and Microscopic Quantities

Any attempt to connect macroscopic quantities with microscopic ones requires a clear definition of the corresponding properties. We consider here only electric properties, and the macroscopic quantity generally used to describe the effect of external electric fields (be it static or dynamic, e.g., in the form of light) is the polarization $\underline{P}(t)$, as it appears in the usual Maxwell equations of electromagnetism, while its microscopic counterpart is the electric dipole moment $\underline{\mu}(t)$. For weak fields, both quantities may be written as an expansion with respect to the corresponding field strengths. The treatment of the microscopic dipole moment is shown in detail in Jaszuński et al. (2015) in this handbook; thus, we will concentrate here on the macroscopic polarization. Although the treatment is essentially analogous, there are some differences, which are often the cause of some confusion.

The macroscopic susceptibility tensors with components $\chi_{\alpha\alpha_1,\dots,\alpha_n}^{(n)}(-\omega_\sigma; \omega_1, \dots, \omega_n)$ can be defined via the expansion terms of the polarization with respect to the macroscopic field, $P_\alpha^{(n)}(t) \sim [E(t)]^n$ as

$$P_{\alpha}^{(n)}(t) = \epsilon_0 \sum_{\omega_1 \dots \omega_n} \chi_{\alpha\alpha_1 \dots \alpha_n}^{(n)}(-\omega_{\sigma}; \omega_1, \dots, \omega_n) E_{\alpha_1}(t) \dots E_{\alpha_n}(t) \quad (1)$$

where $\omega_{\sigma} = \sum_i \omega_i$, the sum over frequencies runs over all distinctive sets of frequencies $\omega_1, \dots, \omega_n$, and the permittivity of vacuum ϵ_0 occurs in the SI system of units. For the Cartesian indices $\alpha_i = x, y, z$, the Einstein summation convention is applied. A similar expression for the dipole moment is employed to define the molecular (hyper)polarizabilities (see Eq. 11.84 in Jaszuński et al. 2015). The latter, however, is written as a Taylor expansion instead of the power expansion used in Eq. 1. Thus, the corresponding equation defining the (hyper)polarizabilities $\alpha^{(n)}$ via the field expansion terms of the dipole moment $\mu_{\alpha}^{(n)}(t) \sim [F(t)]^n$ is

$$\mu_{\alpha}^{(n)}(t) = \frac{1}{n!} \sum_{\omega_1 \dots \omega_n} \alpha_{\alpha\alpha_1 \dots \alpha_n}^{(n)}(-\omega_{\sigma}; \omega_1, \dots, \omega_n) F_{\alpha_1}(t) \dots F_{\alpha_n}(t) \quad (2)$$

with $\alpha^{(n)} = \alpha, \beta, \gamma$ for $n = 1, 2, 3$. This leads to the appearance of factors $n!$ in equations expressing the susceptibilities in terms of (hyper)polarizabilities. In a condensed phase the field $\underline{F}(t)$ in Eq. 2 may not be the macroscopic field $\underline{E}(t)$ used in Eq. 1. In fact, it turns out that a much more convenient choice is the so-called *local* (or *exciting*) field, which is effective on a molecule and which may differ substantially from $\underline{E}(t)$.

The macroscopic or Maxwell field may be written as a superposition of a homogeneous static and a dynamic, monochromatic field:

$$\underline{E}(t) = \underline{E}^0 + \underline{E}^{\omega} \cos(\omega t). \quad (3)$$

This is not the most general expression to describe all nonlinear optical effects, but it covers many cases of experimental interest.

With such a field, the expansion terms $\underline{P}^{(i)}(t)$ can be expressed as a Fourier series:

$$\underline{P}^{(i)}(t) = \underline{P}_0^{(i)} + \underline{P}_{\omega}^{(i)} \cos(\omega t) + \underline{P}_{2\omega}^{(i)} \cos(2\omega t) + \underline{P}_{3\omega}^{(i)} \cos(3\omega t) + \dots \quad (4)$$

Inserting Eq. 3 in Eq. 1 and employing trigonometric identities, the expansion terms in the Fourier series can be connected with the susceptibilities. This leads to

$$P_{0,\alpha}^{(1)} = \epsilon_0 \chi_{\alpha\beta}^{(1)}(0; 0) E_{\beta}^0 \quad (5)$$

$$P_{0,\alpha}^{(2)} = \epsilon_0 \left[\chi_{\alpha\beta\gamma}^{(2)}(0; 0, 0) E_{\beta}^0 E_{\gamma}^0 + \frac{1}{2} \chi_{\alpha\beta\gamma}^{(2)}(0; -\omega, \omega) E_{\beta}^{\omega} E_{\gamma}^{\omega} \right]$$

$$P_{0,\alpha}^{(3)} = \epsilon_0 \left[\chi_{\alpha\beta\gamma\delta}^{(3)}(0; 0, 0, 0) E_{\beta}^0 E_{\gamma}^0 E_{\delta}^0 + \frac{3}{2} \chi_{\alpha\beta\gamma\delta}^{(3)}(0; -\omega, \omega, 0) E_{\beta}^{\omega} E_{\gamma}^{\omega} E_{\delta}^0 \right] \quad (6)$$

$$P_{\omega,\alpha}^{(1)} = \epsilon_0 \chi_{\alpha\beta}^{(1)}(-\omega; \omega) E_{\beta}^{\omega} \quad (7)$$

$$P_{\omega,\alpha}^{(2)} = 2\epsilon_0 \chi_{\alpha\beta\gamma}^{(2)}(-\omega; \omega, 0) E_\beta^\omega E_\gamma^0 \quad (8)$$

$$P_{\omega,\alpha}^{(3)} = \epsilon_0 \left[3\chi_{\alpha\beta\gamma\delta}^{(3)}(-\omega; \omega, 0, 0) E_\beta^\omega E_\gamma^0 E_\delta^0 + \frac{3}{4}\chi_{\alpha\beta\gamma\delta}^{(3)}(-\omega; \omega, -\omega, \omega) E_\beta^\omega E_\gamma^\omega E_\delta^\omega \right] \quad (9)$$

$$P_{2\omega,\alpha}^{(2)} = \frac{1}{2}\epsilon_0 \chi_{\alpha\beta\gamma}^{(2)}(-2\omega; \omega, \omega) E_\beta^\omega E_\gamma^\omega \quad (10)$$

$$P_{2\omega,\alpha}^{(2)} = \frac{3}{2}\epsilon_0 \chi_{\alpha\beta\gamma\delta}^{(3)}(-2\omega; \omega, \omega, 0) E_\beta^\omega E_\gamma^\omega E_\delta^0 \quad (11)$$

$$P_{3\omega,\alpha}^{(3)} = \frac{1}{4}\epsilon_0 \chi_{\alpha\beta\gamma\delta}^{(3)}(-3\omega; \omega, \omega, \omega) E_\beta^\omega E_\gamma^\omega E_\delta^\omega \quad (12)$$

These expansions may be compared with the analogous expressions for the molecular dipole moment in Jaszúński et al. (2015). The linear susceptibility $\underline{\chi}^{(1)}$ is related to the relative permittivity tensor by $\underline{\epsilon} = \underline{\chi}^{(1)} + \underline{I}$, where \underline{I} is the 3×3 identity tensor. Another related quantity is the refractive index: If α is a direction parallel to a principal axis of the indicatrix of the material, the refractive index $n_\alpha(\omega)$ is given by $n_\alpha(\omega) = (\epsilon_\alpha(\omega))^{1/2}$. For an isotropic liquid, for which $\chi_{\alpha\beta}^{(1)} = \delta_{\alpha\beta}\chi^{(1)}$ and $\epsilon_{\alpha\beta} = \delta_{\alpha\beta}\epsilon$, these relationships simplify to $\epsilon = \chi^{(1)} + 1$ and $n(\omega) = (\epsilon(\omega))^{1/2}$.

We note that the conventions used here to define the (hyper)polarizabilities and susceptibilities are probably the most popular in the field of computational chemistry. There are other conventions used, for example, by several experimentalist groups but also by theoreticians (for more details, see Reis (2006) and references therein).

The connection between the polarization of a liquid composed of n_I components and the induced dipole moments of the molecules of component I is given by

$$P_{\omega,\alpha} = \epsilon_0 \sum_{I=1}^{n_I} N_I \langle \mu_\alpha^I(\omega) \rangle, \quad (13)$$

where $\langle \dots \rangle$ denotes averaging over the sample, and N_I is the particle density for the component I . Different models can be invoked to approximate the property on the right-hand side, e.g., a continuum model, a discrete model, or mixtures of the two. Some of these methods will be described in more detail later in this chapter. To clarify the connection between microscopic and macroscopic electric properties, we will take here the view of the liquid as composed of discrete, polarizable molecules, interacting via electrostatic multipole moments. This is essentially the picture applied in the rigorous local field (RLF) method (Janssen et al. 1999). In a typical RLF computation, the liquid is modeled by a molecular simulation method, e.g., a Monte Carlo or molecular dynamics calculation, which yields a series of different configurations. In a second step, the susceptibilities are computed for (usually) a subset of these configurations, and the simulated macroscopic responses are finally obtained by averaging over the individual responses.

In order to connect the macroscopic polarization with microscopic properties of *isolated molecules*, it is essential to realize that the microscopically effective fields in Eq. 2 are *local fields*, which differ from the macroscopic fields \underline{E} in Eq. 1. In a liquid subject to externally applied fields, the total local field on molecule k consists of two contributions, i.e., $\underline{F}_k^{\text{tot}} = \underline{F}_k^{(0)} + \underline{F}_k^{\text{ext}}$, where the first contribution $\underline{F}_k^{(0)}$ is due to electric fields arising from all the surrounding molecules in the liquid without any externally applied field, which, in turn, is considered by the second term and is described below. The *permanent local field* $\underline{F}_k^{(0)}$ is essentially a static field and may be computed, for example, by summing over the field contributions of the surrounding molecules:

$$F_{\alpha,k}^{(0)} = \sum_{l \neq k} \left[T_{\alpha,k;l}^{10} q_l + T_{\alpha,k;\beta,l}^{11} \mu_{\beta,l}^{\text{sol}} + \dots \right] \quad (14)$$

where $q, \mu_{\beta,l}^{\text{sol}}$ are (components of) the charge (if nonzero) and dipole of molecule l , respectively, and $T_{\alpha,k;l}^{10}, T_{\alpha,k;\beta,l}^{11}$ are components of the corresponding dipole–charge and dipole–dipole propagator tensors. In general, a propagator tensor $T_{\alpha_1 \dots \alpha_n, k; \beta_1 \dots \beta_m, l}^{nm}$ relates the negative of the n -th order derivative of the Coulomb potential $\Phi(\underline{r}_k)$ at molecule k to the m -th order multipole of molecule l . The dipole and higher multipole moments are the total moments, i.e., they depend again on the local field $\underline{F}_k^{(0)}$, e.g.,

$$\mu_{\alpha,k}^{\text{sol}} = \mu_{\alpha,k} + \alpha_{\alpha\beta;k} F_{\beta,k}^{(0)} + \dots \quad (15)$$

where $\mu_{\alpha,k}$ and $\alpha_{\alpha\beta,k}$ are components of the permanent dipole and (dipole–dipole) polarizability, respectively, of the *isolated* molecule.

The permanent local field can be quite large for dipolar solutes and/or dipolar or multipolar solvents, and may also be inhomogeneous over the molecule, especially for larger molecules, thus a converged description in terms of the one-point multipole description used above may require the consideration of multipoles higher than dipole and of field gradients of first or higher order, in addition to the homogeneous field. One way to alleviate this problem is to use distributed multipole moments and polarizabilities instead of the one-point properties. Some approaches will be discussed in section “Explicit Molecular Models.” A potentially more accurate procedure is to compute quantum–mechanically the properties of the solute embedded in the field due to surrounding molecules, as in the discrete reaction field (DRF), the polarizable embedding (PE), as well as the frozen-density embedding (FDE) approaches described later.

The properties of the molecules under the influence of the permanent local field have been called *solute* properties (Wortmann and Bishop 1998). The application of an external field gives rise then to the second contribution to the local field $\underline{F}_k^{\text{ext}}(\omega)$. It can be given in terms either of the applied field $\underline{E}^{(0)}(\omega)$ or of the macroscopic field $\underline{E}(\omega)$ and the thereby induced dipole moments $\underline{\mu}^{\text{ind}}(\omega)$ by

$$\underline{F}_k^{\text{ext}}(\omega) = \underline{E}^{(0)}(\omega) + \sum_{l \neq k} \underline{T}_{kl}^{11} \cdot \underline{\mu}_l^{\text{ind}}(\omega) \quad (16)$$

$$= \underline{E}(\omega) + \sum_{l \neq k} \underline{L}_{kl}^{11} \cdot \underline{\mu}_l^{\text{ind}}(\omega), \quad (17)$$

where \underline{T}_{kl}^{11} is the dipole–dipole field propagator tensor, which contains a term which is only *conditionally convergent* in the long-wavelength limit, while \underline{L}_{kl}^{11} is the (dipole–dipole) Lorentz factor tensor containing only the absolutely convergent terms, and the conditionally convergent term has been combined with $\underline{E}^{(0)}$ to yield the macroscopic field \underline{E} . For more details, see, e.g., Boettcher and Bordewijk (1978) and Munn (1988).

Again, the dipole moments $\underline{\mu}_k^{\text{ind}}(\omega)$ depend on $\underline{F}_k(\omega)$:

$$\begin{aligned} \underline{\mu}_{\alpha,k}^{\text{ind}}(\omega) = & \alpha_{\alpha\beta,k}^{\text{sol}}(\omega) F_{\beta,k}^{\text{ext}}(\omega) + K_{12}^{(2)} \frac{1}{2} \beta_{\alpha\beta\gamma,k}^{\text{sol}}(-\omega; \omega_1, \omega_2) F_{\beta,k}^{\text{ext}}(\omega_1) F_{\gamma,k}^{\text{ext}}(\omega_2) + \\ & K_{123}^{(3)} \frac{1}{6} \gamma_{\alpha\beta\gamma\delta,k}^{\text{sol}}(-\omega; \omega_1, \omega_2, \omega_3) F_{\beta,k}^{\text{ext}}(\omega_1) F_{\gamma,k}^{\text{ext}}(\omega_2) F_{\delta,k}^{\text{ext}}(\omega_3) \dots \end{aligned} \quad (18)$$

where the numerical factors $K_{12\dots n}^{(n)} = K^{(n)}(-\omega; \omega_1, \dots, \omega_n)$ depend on the number of equal and/or zero frequencies among the set of frequencies $\{\omega, \omega_1, \dots, \omega_n\}$. Here, it is assumed that $\underline{E}(\omega)$ is the field of light at optical wavelengths; thus, no dipole moment appears on the right-hand side of the equation. If a static external field is applied, it may cause a change of the liquid *structure*. In the consecutive two-step approach considered here, such a structure changing effect should be taken into account already in the molecular simulation. A computationally less expensive method is to assume that the orientational effect of the static field is well described by the Boltzmann distribution $\sim \exp^{-\underline{\mu} \cdot \underline{F}^{\text{ext}}(0)/kT}$. Assuming in addition that $\underline{\mu} \cdot \underline{F}^{\text{ext}}(0)/kT \ll 1$, the Boltzmann factor can be expanded and only the first nonzero terms retained. This leads, e.g., to the standard expression for the EFISH process given in Eq. 65, which can then be applied in conjunction with a field-free molecular simulation.

The *solute (hyper)polarizabilities* can again be expressed by the properties of the isolated molecule and the permanent local field, i.e.,

$$\alpha_{\alpha\beta,k}^{\text{sol}}(\omega) = \alpha_{\alpha\beta,k}(\omega) + \beta_{\alpha\beta\gamma}(-\omega; \omega, 0) F_{\gamma,k}^{(0)} + \dots \quad (19)$$

$$\beta_{\alpha\beta\gamma,k}^{\text{sol}}(-\omega; \omega_1, \omega_2) = \beta_{\alpha\beta\gamma,k}(-\omega; \omega_1, \omega_2) + \gamma_{\alpha\beta\gamma\delta,k}(-\omega; \omega_1, \omega_2, 0) F_{\delta,k}^{(0)} + \dots \quad (20)$$

$$\begin{aligned} \gamma_{\alpha\beta\gamma\delta,k}^{\text{sol}}(-\omega; \omega_1, \omega_2, \omega_3) = & \gamma_{\alpha\beta\gamma\delta,k}(-\omega; \omega_1, \omega_2, \omega_3) \\ & + \delta_{\alpha\beta\gamma\delta\epsilon,k}(-\omega; \omega_1, \omega_2, \omega_3, 0) F_{\epsilon,k}^{(0)} + \dots \end{aligned} \quad (21)$$

The set of Eqs. 13, 17, and 18 can be solved for the susceptibilities. The result is a set of temperature-independent expressions, which in general constitute the most important contributions, as well as some temperature-dependent contributions (Kasprowicz–Kielich and Kielich 1975; Janssen et al. 1999). The temperature-independent expressions for the diagonal components can be written as

$$\chi_{\alpha\alpha}^{(1)}(\omega) = \frac{1}{\epsilon_0 v} \left\langle \sum_k \alpha_{\alpha\beta,k}^{\text{sol}}(\omega) d_{\beta k,\alpha}(\omega) \right\rangle \quad (22)$$

$$\chi_{\alpha\alpha\alpha}^{(2)}(-\omega; \omega_1, \omega_2) = \frac{1}{2\epsilon_0 v} \left\langle \sum_k d_{\beta k,\alpha}(\omega) \beta_{\beta\gamma\delta,k}^{\text{sol}}(-\omega; \omega_1, \omega_2) d_{\gamma k,\alpha}(\omega_1) d_{\delta k,\alpha}(\omega_2) \right\rangle \quad (23)$$

$$\begin{aligned} & \chi_{\alpha\alpha\alpha\alpha}^{(3)}(-\omega; \omega_1, \omega_2, \omega_3) \\ &= \frac{1}{6\epsilon_0 v} \left\langle \sum_k d_{\beta k,\alpha}(\omega) \gamma_{\beta\gamma\delta\epsilon,k}^{\text{sol}}(-\omega; \omega_1, \omega_2, \omega_3) d_{\gamma k,\alpha}(\omega_1) d_{\delta k,\alpha}(\omega_2) d_{\epsilon k,\alpha}(\omega_3) \right\rangle \end{aligned} \quad (24)$$

where k labels the molecules in the simulation box with volume v . The *local field tensors* $d_{\alpha k,\beta}(\omega)$, which correspond to the *cavity* and *reaction* field factors in the continuum model, are given by

$$d_{\alpha k,\beta}(\omega) = \sum_{k'} [\underline{\mathbb{I}} - \underline{\underline{L}} \cdot \underline{\alpha}^{\text{sol}}(\omega) / (\epsilon_0 v)]_{\alpha k,\beta k'}^{-1}, \quad (25)$$

where $I_{\alpha k,\beta k'} = \delta_{kk'} \delta_{\alpha\beta}$ and $\alpha_{\alpha k,\beta k'}^{\text{sol}}(\omega) = \alpha_{\alpha\beta,k}^{\text{sol}} \delta_{kk'}$ and $L_{\alpha k,\beta k'}$ is an element of the Lorentz factor tensor. In the two-step approach for the calculation of susceptibilities mentioned above, for each (or each n th) configuration of the simulation, the expressions in the brackets are evaluated, and finally an average over all snapshots is performed. Apart from the direct contribution shown in Eq. 24, there is also a so-called cascading contribution to $\chi^{(3)}$, which, however, is often much smaller than the direct part and requires that the molecules have a nonvanishing first hyperpolarizability (Munn 1988; Reis et al. 2006).

The solute properties, defined in Eqs. 19, 20, and 21, are not, in general, experimentally accessible, being defined as electric properties in the absence of an external electric field. Experimentally relevant molecular properties may be defined as derivatives of the total molecular dipole moment with respect to the *macroscopic field* \underline{E} , i.e., as in Eq. 2 with $\underline{F} \equiv \underline{E}$. These have been introduced in the framework of the semiclassical continuum model of Onsager and Lorentz by Wortmann and Bishop (1998) and were called *effective properties*. In the semiclassical Onsager model, the connection between solute properties and effective properties is given by cavity and reaction field factors, e.g., for the polarizability:

$$\alpha_{\beta\beta}^{\text{eff}}(\omega) = f_{\beta\beta}^C \omega F_{\beta\beta}^{R\omega} \alpha_{\beta\beta}^{\text{sol}}(\omega) \quad (26)$$

where, in the *spherical cavity* approximation and with a cavity of volume v ,

$$f_{\beta\beta}^{C\omega} = \frac{3\epsilon(\omega)}{2\epsilon(\omega) + 1} \quad (27)$$

$$F_{\beta\beta}^{R\omega} = \frac{1}{1 - f_{\beta\beta}^{R\omega} \alpha_{\beta\beta}^{\text{sol}}(\omega)} \quad (28)$$

$$f_{\beta\beta}^{R\omega} = \frac{2(\epsilon(\omega) - 1)}{v(2\epsilon(\omega) + 1)} \quad (29)$$

This model can be extended to ellipsoidal cavities, with sometimes greatly improved results (Wortmann and Bishop 1998). For more details about the cavity and reaction field factors, see Boettcher and Bordewijk (1978).

From Eqs. 1 and 13 one may define effective properties in the RLF model as the coefficient of the corresponding term in the field expansion of the averaged dipole moment $\langle \mu_{\alpha}^I(\omega) \rangle$. Thus, they can be computed from Eqs. 22, 23, and 24 by removing the prefactors $1/(n!\epsilon_0 v)$ and restricting the sums over all molecules k to summations over all molecules of kind I , k_I .

To conclude this section, we remark that:

1. In more sophisticated continuum models, such as PCM, the connection between solute and effective properties is not linear any more (see section “Polarizable Continuum Model (PCM) Approaches”).
2. The definition of *effective* properties is not unique; see, e.g., the discrete reaction field model, described in section “Discrete Reaction Field (DRF) Method”.
3. The corresponding relationships of Eq. 26 between the effective and solute *hyperpolarizabilities* given in Wortmann and Bishop (1998) have been criticized and corrected in Munn et al. (2005).

Solute–Solvent Interactions

In this section, we discuss some aspects of intermolecular interactions on electric and optical properties, which can be relevant for the calculations of liquids. The emphasis is put here on effects which are considered only very approximately or not at all in models of liquid presented in later sections. This may help to keep the possible limitations of these approaches in perspective. The picture presented in the previous section rests essentially on the classical model of mutually polarizing charge densities, with or without an external applied electric field. No account is made of the quantum-mechanical effects of dispersion and exchange interactions. Although some of the models described below take these effects into account in a more or less empirical way, e.g., by using Lennard–Jones potentials, a more quantitative knowledge of these effects is desirable. The classical approximation of electrostatic interaction with polarization appears to work quite well in many

circumstances, but it has been shown in some cases that these effects may be strongly changed or even overwhelmed by short-range interaction and/or dispersion effects. In this connection one should also mention the *confinement* effect of charges on electric properties, which are especially pronounced in anions and which have been investigated a long time ago by Fowler, Madden, and coworkers on ionic crystals (see, e.g., Fowler and Madden 1984; Jemmer et al. 1998).

Interaction-Induced Properties

In order to study the effect of intermolecular interactions on electric properties, it is useful to introduce the concept of interaction-induced properties. We consider here only liquids or more generally clusters in which the molecules (or ions) retain essentially their identity and interact via weak intermolecular forces which are much smaller than the intramolecular forces. Then the interaction energy can be defined as the difference between the energy of the whole system and that of the isolated parts. As the electric properties are defined as electric field derivatives of the energy, an interaction-induced electric property P_{int} can naturally be defined as

$$P_{\text{int}} = P^{\text{tot}} - \sum_i P^i \quad (30)$$

where P^{tot} is the property of the total “supersystem” and P^i are the properties of the isolated subsystems. This equation, however, does not allow to determine which part of P_{int} “belongs” to which molecule. This is indeed not possible without further analysis, because many intermolecular interactions are not additive and/or may involve many-body effects. Thus, in order to be able to analyze the solvent effect on property P of solute A , it is necessary to develop a distributing or *splitting* scheme for P_{int} .

There are essentially two different approaches used for the determination of interaction energies and properties, based either on supermolecular calculations or on perturbation theory. The supermolecular (SM) approach relies on computation of the energies and/or properties of the whole system and of the isolated molecules, which yields P_{int} according to Eq. 30. If any approximate representation of the wave function relying on finite basis sets is used for these calculations, the well-known problem of basis set superposition error (BSSE) due to incomplete basis sets has to be taken care of. The BSSE can be quite large for E_{int} if too small basis sets are employed. This may be different for properties, and it was indeed shown in several works (Baranowska et al. 2011; Medveđ et al. 2015) that the BSSE for interaction-induced dipole moment and (hyper)polarizabilities of weakly bound (including hydrogen-bonded) systems is rather small, although it is quite large for the interaction energies. It thus seems that the BSSE has no serious effect in electric property computations provided that reasonable basis sets including diffuse functions are used, although more data may be required to draw a definite conclusion.

An advantage of the SM approach is that it is independent of the strength of the interactions; however, it yields only the total interaction energy or property. A deeper insight into the nature of interaction-induced properties in terms of the type of interaction can be obtained using perturbation theory, e.g., in the *variational-perturbational scheme* of Sokalski et al. (1988). Another popular approach is SAPT, which stands for *symmetry-adapted perturbation theory*, although this is also the generic name for a whole class of short-range perturbation theories (see, e.g., Stone 2013, pp. 110). For reviews of SAPT see Szalewicz (2012) (see also Misquitta 2015 in this handbook). In SAPT, the interaction energy of a dimer AB is evaluated as a double-perturbative expansion through the Møller–Plesset fluctuation potential operators W_A and W_B of the subsystems A and B, respectively, and the interaction potential V . In this manner, the interaction energy is determined *without* computing the total energy of the supersystem and subsystems, it is thus free of the BSSE, and, moreover, a decomposition of the interaction energy into physically meaningful components (i.e., electrostatic, exchange, induction, and dispersion terms) can be performed, e.g., in the simplest truncation of SAPT denoted as SAPT0:

$$E_{\text{int}}^{\text{SAPT0}} = E_{\text{elst}}^{(10)} + E_{\text{exch}}^{(10)} + E_{\text{ind}}^{(20)} + E_{\text{exch-ind}}^{(20)} + E_{\text{disp}}^{(20)} + E_{\text{exch-disp}}^{(20)} \quad (31)$$

where the perturbation term $E^{(nm)}$ is n -th order in V and m -th order in $W_A + W_B$. Thus, electrostatic and exchange interactions are already included in first order in the interaction V , while induction and dispersion first appear through second order. The exchange–induction energy $E_{\text{exch-ind}}^{(20)}$ and the exchange–dispersion energy $E_{\text{exch-disp}}^{(20)}$ terms represent the effect of the antisymmetrization of the first-order induction and dispersion wave functions.

Using the definition of electric moments and (hyper)polarizabilities as field derivatives of the energy, it is straightforward, e.g., by applying *finite field* techniques, to compute interaction-induced electric properties of the individual molecules of the cluster, as well as their decomposition into the different interaction types. For a property P , we thus have in the SAPT0 approximation

$$P_{\text{int}} = P_{\text{elst}}^{(10)} + P_{\text{exch}}^{(10)} + P_{\text{ind}}^{(20)} + P_{\text{exch-ind}}^{(20)} + P_{\text{disp}}^{(20)} + P_{\text{exch-disp}}^{(20)} \quad (32)$$

In Table 1 we show the effect of the type of interaction on electric properties for a few selected dimers. It can be seen that the total induced dipole moment can be well reproduced by considering only electrostatic effects. In the case of polarizability, the differences between the total and electrostatic contributions are more significant. It can also be observed that the latter usually overestimates the former mainly due to the important role of the exchange repulsion effects, particularly in the case of hydrogen-bonded systems. For the first hyperpolarizability, it is difficult to neglect any contribution without further analysis. Even the dispersion contribution, which is in most cases negligible, may sometimes become unusually large; it contributes, e.g., more than 15 % to β_{int} of the pABA-z...benzene complex.

Table 1 The decomposition of the dominant component of interaction-induced properties for several molecular dimers. The total P_{int} value includes all terms of Eq. 32. TCNE = tetracyanoethylene, pABA-c(z) = canonical (zwitterionic) form of *p*-aminobenzoic acid, Bz = benzene. DNA pairs correspond to the B-form. All values are in a.u.

System	Dipole moment					Ref.
	$\mu_{\text{elst}}^{(10)}$	$\mu_{\text{exch}}^{(10)}$	$\mu_{\text{ind}}^{(20)}$	$\mu_{\text{disp}}^{(20)}$	μ_{int}	
HF dimer	-0.13	-0.03	-0.02	0.01	-0.17	Skwara et al. (2008)
HCN dimer	-0.27	0.01	-0.06	0.00	-0.29	Góra et al. (2011)
TCNE···Bz	-0.54	0.29	-0.31	-0.03	-0.58	This work.
pABA-c···Bz	-0.13	0.00	0.03	0.00	-0.11	Medveď et al. (2015)
pABA-z···Bz	-0.92	0.06	0.12	-0.02	-0.78	Ibid.
System	Polarizability					
	$\alpha_{\text{elst}}^{(10)}$	$\alpha_{\text{exch}}^{(10)}$	$\alpha_{\text{ind}}^{(20)}$	$\alpha_{\text{disp}}^{(20)}$	α_{int}	
HF dimer	1.38	-0.83	0.32	0.16	0.96	Skwara et al. (2008)
HCN dimer	4.76	-1.54	1.24	0.32	4.50	Góra et al. (2011)
A···T	4.29	-6.22	3.63	1.24	2.94	Czyźnikowska et al. (2011)
G···C	5.23	-11.12	6.70	1.97	2.78	Ibid.
C···C	-8.22	-0.87	1.46	0.76	-6.89	Ibid.
TCNE···Bz	22.92	-17.65	15.80	2.33	23.43	This work
pABA-c···Bz	-24.82	-1.20	3.90	1.10	-19.50	Medveď et al. (2015)
pABA-z···Bz	-20.80	-5.48	5.96	1.72	-18.64	Ibid.
System	First hyperpolarizability					
	$\beta_{\text{elst}}^{(10)}$	$\beta_{\text{exch}}^{(10)}$	$\beta_{\text{ind}}^{(20)}$	$\beta_{\text{disp}}^{(20)}$	β_{int}	
HF dimer	5.30	-11.41	-0.04	1.56	-4.76	Skwara et al. (2008)
HCN dimer	-15.2	12.6	-21.5	-0.02	-38.9	Góra et al. (2011)
TCNE···Bz	449	-502	881	32	860	This work
pABA-c···Bz	-270	-12	70	4	-184	Medveď et al. (2015)
pABA-z···Bz	-237	-91	7	48	-275	Ibid.

Distance Dependence

If the distance between the entities *A* and *B* in a dimer *AB* is sufficiently large, the interaction Hamiltonian *V* in the SAPT perturbation expression may be represented by its multipole expansion as a sum of terms proportional to inverse powers of the intermolecular distance *R*. This allows to express the individual interactions mentioned in the previous subsection in terms of molecular properties and to study their distance dependence. It turns out that the first-order term $\mu_{\text{elst}}^{(10)}$ is due to the polarization of molecule *A* by the electric field created by the permanent multipole moments on molecules *B* and *vice versa*. The second-order term $\mu_{\text{ind}}^{(20)}$ is caused by two mechanisms. Firstly, an induced moment on *A* leads to first-order polarization of *B*, which in turn induces a dipole on *A*. Secondly, permanent moments of *B* interact with the hyperpolarizability of *A* and *vice versa*. This nonlinear effect requires a nonvanishing first hyperpolarizability on *A* and/or *B*.

Finally, the second-order $\mu_{\text{disp}}^{(20)}$ can be described as coming from the interaction of the polarizability of A with first hyperpolarizability of B and *vice versa*. The term $\mu_{\text{exch}}^{(20)}$ does not appear in the multipole expansion. Indeed it decays very fast with the intersystem distance since it is related to the overlap of the wave functions of A and B.

The influence of the distance dependence on interaction-induced properties can be demonstrated by considering two interacting HF molecules (Skwara et al. 2008) (Fig. 1a). Starting from the intermolecular distance, defined as the distance between the two hydrogen-bonded atoms $R = 2.3\text{--}2.5 \text{ \AA}$ (i.e., $R \approx R_{\text{eq}} + 0.5 \text{ \AA}$), the interaction energy is dominated essentially by E_{elst} , because the other terms cancel each other in this region, although they are individually quite large, about 10–15% of E_{elst} . In the region of $R > R_{\text{eq}} + 1.0 \text{ \AA}$, all terms except E_{elst} can be neglected. As shown in Fig. 1b, c, this holds also for μ_{int} and α_{int} . In both cases, the first-order electrostatic contribution fairly well describes the total interaction-induced values. For the interaction-induced first hyperpolarizability, the case is more complicated, since now $\beta_{\text{elst}}^{(10)}$ has the opposite sign compared to the total MP2 β value for $R > 1.5 \text{ \AA}$. This is caused by a much faster decay of $\beta_{\text{elst}}^{(10)}$ combined with relatively large β_{exch} and $\beta_{\text{exch-ind}}$ contributions.

Many-Body Effects

As already mentioned, some interactions (e.g., electrostatic and second-order dispersion forces) depend only on two interacting partners, which means that they are pairwise additive. However, some effects such as the induction, exchange, and third-order dispersion interactions have *many-body* character, i.e., they are not pairwise additive. Generally, the total property P^{tot} can be expressed as

$$P^{\text{tot}} = \sum_i P^i + \sum_{i < j} \Delta^{ij} P^{(2)} + \sum_{i < j < k} \Delta^{ijk} P^{(3)} + \sum_{i < j < k < l} \Delta^{ijkl} P^{(4)} + \dots \quad (33)$$

where P^i is the property of an isolated molecule i and $\Delta^{ij} P^{(2)}$, $\Delta^{ijk} P^{(3)}$, and $\Delta^{ijkl} P^{(4)}$ are two-, three-, and four-body contributions to the interaction-induced property arising from interaction between molecules i , j , k , and l , respectively. From many-body effects, the three-body interactions are usually most relevant; it has been shown that three-body effects may contribute from a few to several tens of percents of the two-body contributions to the interaction energy (Bukowski and Szalewicz 2001). The role of many-body effects on interaction-induced properties has been investigated, for example, by Skwara et al. (2007), Góra et al. (2011), and Baranowska et al. (2010). For hydrogen-bonded systems, the ratio of the three-body and the sum of two-body contributions ($\Delta^{123} P^{(3)} / \sum \Delta^{ij} P^{(2)}$) to the interaction-induced dipole moment was found to be small, about 5%, because the main contribution to $\Delta\mu$ comes from electrostatic forces, which are pairwise additive. For the dipole polarizability, $\Delta^{123} \alpha^{(3)} / \sum \Delta^{ij} \alpha^{(2)}$ is typically about 10–

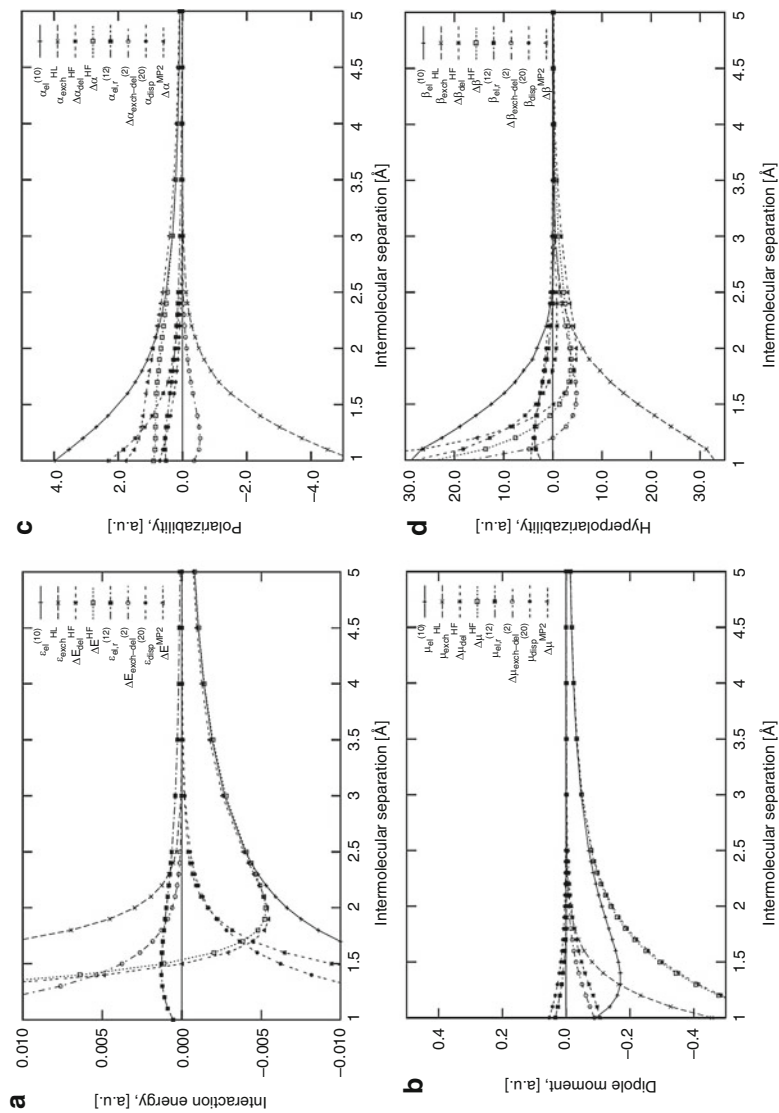


Fig. 1 Several contributions to the interaction energy (a) and interaction-induced properties (b–d) as a function of the intersystem distance for the HF dimer calculated at the MP2/aug-cc-pVTZ level of theory (Adapted from Skwara et al. 2008)

15 %, but for some cases, such as the diformamide trimer, it goes up to several tens of percents, and the many-body contributions may even be larger than the two-body contribution. The dominant component of the first hyperpolarizability was found to be even more sensitive to many-body effects than the polarizability. In this case, $\Delta^{123}\beta^{(3)}/\Sigma\Delta^{ij}\beta^{(2)}$ values typically span from 10 to 20 %. If the number of molecules in the cluster increases, the many-body effects become dominant (e.g., in HCHO... $(\text{HF})_n$) over the two-body effects (Baranowska et al. 2010). The decomposition of the three-body contribution to interaction-induced (hyper)polarizabilities shows that the nonadditivity arises mainly from the induction and exchange–induction interactions as well as higher-order electron correlation effects (Góra et al. 2011). To conclude, considering only two-body contributions may not be sufficient for a reliable description of the total interaction-induced (hyper)polarizabilities. In addition, it is not straightforward to predict whether nonadditive interactions will lead to an enhancement or a reduction of the two-body contribution.

Splitting Schemes

The previous sections have shown the usefulness of interaction-induced properties derived from approaches based on perturbation theory. As mentioned above, in order to assign molecular contributions to the properties or energies obtained from cluster (or supermolecular) calculations, some kind of splitting scheme has to be applied. The first schemes have been proposed for partitioning the total polarizability (Mikkelsen et al. 1995; Morita and Kato 1999; Jensen et al. 2002). In case of a system composed of N identical particles, the simplest way is to calculate the average polarizability per molecule (PPM), i.e., α/N (see, e.g., Hammond et al. 2009). Another scheme, which is also applicable to many component systems, is the differential shell approach (DSA) (Mikkelsen et al. 1995; Morita and Kato 1999), in which the solute polarizability is defined as the difference between the total supermolecular polarizability and the polarizability of the cluster without the solute molecule. The solute polarizability can thus be obtained in two calculations of the supersystem with and without the solute molecule. This scheme can also be applied for higher-order properties and provides a reasonable procedure to estimate the solute hyperpolarizability.

A more sophisticated treatment applicable for pure as well as many-component liquids is based on the expansion of the total polarizability, α^{tot} , through a cluster sum (Moraldi et al. 1989):

$$\alpha^{\text{tot}} = \sum_i \alpha^i + \sum_{i<j} \Delta^{ij} \alpha^{(2)} + \sum_{i<j<k} \Delta^{ijk} \alpha^{(3)} + \dots \quad (34)$$

where α^i is the polarizability of an isolated molecule i and $\Delta^{ij}\alpha^{(2)}$, $\Delta^{ijk}\alpha^{(3)}$, ... are two-body, three-body, ... contributions to the interaction-induced polarizability,

respectively. The total cluster polarizability can alternatively be expressed as the sum of many-body corrected molecular polarizabilities (MBP):

$$\alpha^{\text{tot}} = \sum_i \alpha_{\text{MB}}^i \quad (35)$$

which are defined in such a way that the two expansions of the total polarizability lead to the same result, i.e.,

$$\alpha_{\text{MB}}^i = \alpha^i + \sum_j W_{ij}^{(2)} \Delta^{ij} \alpha^{(2)} + \sum_{j,k} W_{ijk}^{(3)} \Delta^{ijk} \alpha^{(3)} + \dots \quad (36)$$

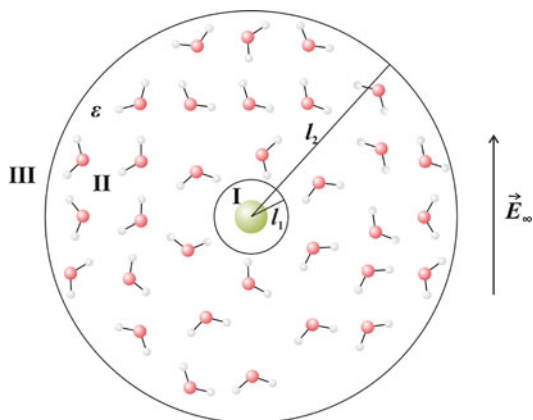
where $W_{ij}^{(2)}$, $W_{ijk}^{(3)}$, ... are weighting factors that divide the two-, three-, ... body contributions among the molecules according to a chosen scheme, which can be based, e.g., on the polarizabilities of the isolated molecules (Jensen et al. 2002)

$$W_{ij,\alpha\beta}^{(2)} = \frac{|\alpha_{\alpha\beta}^i|}{|\alpha_{\alpha\beta}^i| + |\alpha_{\alpha\beta}^j|} \quad (37)$$

and analogously for the higher-order weighting factors. The MBP scheme is in principle also applicable to hyperpolarizabilities. Its major disadvantage is that it is necessary to evaluate many-body contributions to the interaction-induced properties, which can be computationally demanding. Jensen et al. (2002) proposed an efficient scheme for partitioning the cluster polarizability called a localized dipole interaction model (LIM), which is based on the dipole interaction model of Applequist et al. (1972) and relies on decomposition of a so-called relay matrix into individual molecular contributions using the weighting factors as in the MBP approach. An *intramolecular* variant of this model has been used to define distributed polarizabilities and is described in section “Discrete Reaction Field (DRF) Method.” It was shown that for small water clusters the LIM scheme provides results comparable with the more demanding MBP approach. While the DSA scheme was in qualitative agreement with LIM and MBP, the PPM approach was not able to reproduce the anisotropic character of the molecular polarizability.

The appropriateness of the DSA scheme for the splitting of cluster polarizabilities into molecular contributions was also addressed by Morita and Kato (1999). The main drawback of DSA is that the difference of the property values for the supersystem with and without the solute molecule involves dielectric interference effects between the solute and solvent polarization, such as *electrostatic shielding* or *reaction field* coming from the medium polarization. In order to estimate these effects and thus to obtain the *net (intrinsic) solute polarizability* of a molecule in solution, Morita and Kato combined the multilayer dielectric continuum model with the supermolecular approach based on Monte Carlo simulations. In their three-layer model shown in Fig. 2, the solute molecule occupies the central part I, i.e., the cavity,

Fig. 2 Schematic picture of the three-layer model used by Morita and Kato (1999) to derive dielectric characteristics of the solvent enabling to determine the net (intrinsic) solute polarizability of a molecule in solution



whose radius l_1 was estimated from the van der Waals atomic radii of the solute and solvent molecules, similarly as in continuum models. Part II ($l_1 < r < l_2$) corresponds to the solvent with the static and optical dielectric constants ϵ_0^{II} and $\epsilon_\infty^{\text{II}}$, respectively, which may be different from their bulk values. Due to the statistical sampling, ϵ_0^{II} and $\epsilon_\infty^{\text{II}}$ are supposed to be isotropic. The size of the region II should be converged with respect to the target values (typically ca. 6–7 Å). Part III ($r > l_2$) is the external region, i.e., vacuum. From the boundary conditions for the electrostatic potential that must hold for the system in an external electric field E_∞ , Morita and Kato derived a pair of equations for spherical solutes, in which (i) the polarizability of the solvent α^{solv} can be expressed as a function of ϵ_∞ , l_1 , and l_2 and (ii) the difference $\Delta\alpha$ between the polarizability of the supersystem α^{tot} and that of the solvent is an explicit function of ϵ_∞ , l_1 , l_2 , and α , which is the intrinsic polarizability of the solute. Since α^{solv} and α^{tot} can be obtained from the QM/MC calculations, the optical dielectric constant ϵ_∞ can be determined from the first equation, and afterward the second equation can be utilized to obtain the intrinsic solute polarizability.

In Table 2 we show calculated isotropic polarizabilities for several solutes in water and liquid argon. The main conclusions that can be drawn are:

1. The intrinsic polarizability values α of the studied solutes are close to the difference $\Delta\alpha$, which means that the dielectric shielding or reaction field effects do not have large impact on the present systems.
2. A decrease of the molecular polarizabilities is observed when going from the gas phase to the solvent. Typically, the polarizabilities decrease by about 15 % for neutral systems; however, for the Cl^- anion it is more than 35 %.

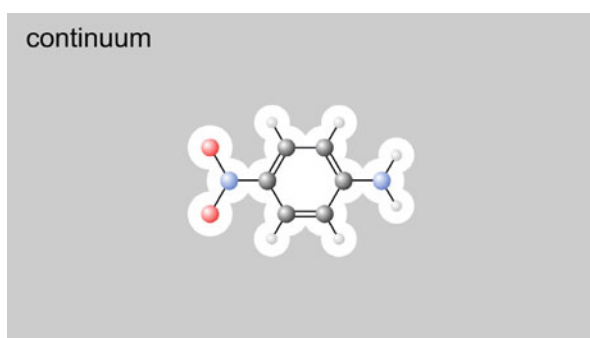
Although the validity of the first conclusion requires justification for a wider spectrum of solvents, the good performance of the differential shell approach (DSA) is encouraging. The second conclusion is in contrast with the PCM findings, since the continuum models usually predict significant *enhancement* of the molecular

Table 2 Comparison of the gas phase, DSA, and net polarizabilities (in a.u.) of Ne, Ar, CH₄, Cl⁻, and H₂O in water and liquid argon (Taken from Morita and Kato 1999)

Solute	Water				Liquid argon			
	α^{gas}	$\Delta\alpha^{\text{DSA}}$	α^{net}	$\alpha^{\text{net}}/\alpha^{\text{gas}}$	α^{gas}	$\Delta\alpha^{\text{DSA}}$	α^{net}	$\alpha^{\text{net}}/\alpha^{\text{gas}}$
Ne	2.358	1.923	1.945	0.825	2.357	2.160	2.160	0.917
Ar	10.586	9.188	9.227	0.872	10.582	10.368	10.334	0.977
CH ₄	16.002	13.947	14.003	0.875	16.000	15.904	115.839	0.990
Cl ⁻	31.036	19.563	19.677	0.634	30.940	25.510	25.406	0.821
H ₂ O	8.530	7.177	7.214	0.846	8.529	7.801	7.778	0.912

polarizability in solution (see previous subchapter). This qualitative disagreement has its roots in the fact that the PCM treatment leads to overpolarization, since it is not compensated (among other effects) for the exchange valence repulsion effects. The supermolecular approach of Morita and Kato, on the other hand, takes into account all interactions, as described by the level of theory applied. Analysis of the α^{tot} values (based on evaluation of a hypothetical polarizability α_{exch} of the solute molecule where only the solute–solvent exchange interaction is involved) suggests that the exchange interactions play indeed an important role in reducing the solute polarizability. For instance, the reduction of α^{gas} for Ne and H₂O in liquid water due to the exchange interactions was estimated to be about 47 % and 63 %, respectively, of the total reduction ($\alpha^{\text{gas}} - \alpha$). The remainder is attributed to other contributions such as the induction, dispersion, and charge transfer interactions. Nevertheless, the exchange interaction appears to be a dominant mechanism of the reduction as was shown also in Mennucci et al. (1998); Serr and Netz (2006).

Implicit Solvent Methods



Although there are several approaches which treat the molecules in a liquid explicitly, they scale very unfavorably with the size of the system. Consider a small dipolar organic molecule such as benzoic acid in aqueous solution. This system may require a cutoff radius of about 12 Å for a reasonably accurate description of

the long-range electrostatic dipole–dipole interactions. A sphere of this size contains about 300 molecules, which would be a sizable computationally demanding task. In addition, the calculation would have to be repeated for many different configurations which may occur in the liquid for a realistic description. It turns out that such a detailed description may not be necessary; in many cases the *average* influence of the solvent on the properties of a solute may be well described by just a few macroscopic parameters. This is the underlying principle of the continuum (or implicit solvent) models. Here the solvent is treated as a uniform, polarizable continuum, described by the dielectric constant ϵ and refractive index $n(\omega)$, while the solute surrounded by the continuum is treated on a molecular level.

The creation of a cavity in the medium for the solute requires work related to the destabilization of the system. This effect is counteracted by favorable electrostatic interactions between the molecular charge density and the polarizable continuum, which in turn acts back on the (polarizable) molecular charge density. Additional contributions to the interaction energy are due to dispersion and repulsion terms, which would need to be considered for a full description. The solvation free energy of the solute–solvent system may thus be written as a sum of three components:

$$\Delta G_{\text{solv}} = \Delta G_{\text{cav}} + \Delta G_{\text{elst}} + \Delta G_{\text{disp-rep}} \quad (38)$$

Not all terms are considered in each model; in most of the current implicit solvation models, only the electrostatic term is explicitly considered. The cavitation and dispersion–repulsion terms may be computed separately, using empirical parameters and/or simplifying assumptions (Amovilli and Mennucci 1997); such an approximate approach still needs to be validated by comparison with the fully quantum-mechanical studies at a molecular level.

As mentioned above, the polarization interaction between continuum and solute is not linear; thus, the energy has to be calculated iteratively, leading to the so-called self-consistent reaction field (SCRF) models. Many different SCRF models have been implemented in publicly available program packages, differing, for example, in the description of the cavity, the treatment of the electrostatic solute–continuum interactions, etc. Popular models include the polarizable continuum model (PCM) (Tomasi et al. 2005); the conductor-like screening model for real solvents (COSMO-RS) (Klamt 2011); the solvation models (SM x), where x provides information about the model (e.g., $x = 5.2$ denotes a fifth-generation model using Mulliken charges) (Marenich et al. 2007); and the multipole expansion method (MPE) (Rinaldi and Rivail 1973; Mikkelsen et al. 1994a). They generally perform very well in calculations of the solvation free energies of neutral molecules, with absolute mean errors lower than 1 kcal/mol (Klamt et al. 2009). This is a nice performance for such computationally quite cheap approaches leading to the spread of continuum models into many other branches of computational chemistry. Another attractive advantage of continuum models is that they have been implemented in conjunction with virtually all levels of computational chemistry.

Many computations of electric, linear, and nonlinear optical properties employing continuum models used either the PCM or the MPE model; thus, we will only describe these two models in more detail below.

Polarizable Continuum Model (PCM) Approaches

The simplest shape of the cavity for solute is a sphere. This form has the advantage that the electrostatic interaction between solute and the dielectric medium can be calculated analytically, and has been used in some of the very first dielectric continuum theories which were developed to account for solvent effects, such as the dielectric polarization theories of Onsager and Kirkwood (see Boettcher and Bordewijk (1978) for an extensive discussion of these theories). Although a generalization of the analytical solution to ellipsoidal shapes is possible, these are still rather unnatural cavity shapes for many molecules. The PCM tries to employ molecular-shaped cavities, generated by interlocking spheres centered on each nucleus. The radii of these spheres are usually taken to be the van der Waals radii, scaled by a constant (typically 1.1, but see discussion below). As no analytical solution is possible for such shapes, the resulting surface of this cavity is then partitioned into small pieces k , called *tesserae*, each being assigned then a point charge q_k . In this way the total potential due to the polarization of the dielectric medium can be written as a discrete sum over the tesserae k :

$$V(\underline{r}) = \sum_{k=1}^N \frac{q_k}{|\underline{r} - \underline{r}_k|} \quad (39)$$

where \underline{r}_k is a vector to a point of tessera k . The local value of the potential necessary to define q_k depends on all point charges $\underline{q} = (q_1, \dots, q_N)^T$, which may be determined from a matrix equation:

$$\underline{q} = -\underline{K}(\epsilon) \cdot \underline{V}(\underline{r}_k) \quad (40)$$

where $\underline{V}(\underline{r}_k)$ collects the electrostatic potential at the tesserae due to the remaining tesserae. The matrix \underline{K} depends on the cavity geometrical parameters and the dielectric constant of the solvent.

The solvent–solute coupling is then introduced into the Hamiltonian describing the solute by adding to the Hamiltonian of the isolated molecule (\hat{H}_0) the electrostatic potential defined by Eq. 39. The wave function and the free energy of the solute can be finally obtained by minimizing the following functional:

$$G(\Psi) = \langle \Psi | \hat{H}_0 + 1/2\hat{V} | \Psi \rangle \quad (41)$$

Following the definition of electric properties, the electric dipole moment, (hyper)polarizabilities, etc., may be obtained by differentiation of the total energy of the solvated molecule with respect to a static or dynamic external field. According to Eq. 38, each field derivative may consist of different contributions, although generally only the electrostatic term is taken into account self-consistently. Mennucci et al. (1998) investigated the different contributions to the electric properties of several small molecules. They found a reduction of components of polarizability and second hyperpolarizability tensors up to 2 % with respect to the gas-phase values due to the Pauli repulsion term. This can be understood as an effect of constraining solute electron density. As expected, the electrostatic contribution yields much larger effects on the properties mentioned above (up to +16 % for components of α and up to +26 % for components of γ). The overall effect of dispersion was positive, but small (+1–2 %). The contributions of Pauli repulsion and dispersion to β were positive or negative, depending on the component, but the absolute changes were generally larger than for α and γ .

The calculations reported above used water as a solvent, which is very polar. While the electrostatic component is clearly very sensitive to solvent polarity, the effect of the solvent polarity on the remaining terms is not clear a priori. In another study, the contributions of electrostatic, dispersion, and Pauli repulsion on the push–pull molecule *p*-nitroaniline in nonpolar cyclohexane, dichloromethane, as well as water were studied (Budzák, unpublished). It was found that the dispersion and repulsion effects were negligible in comparison to the electrostatic effect of the polar molecule. We reiterate that Pauli repulsion, dispersion, and cavitation energies are usually not computed by default in PCM calculations in many quantum chemistry codes. When dealing with less polar solutes in nonpolar solvents, it may be worthwhile to check the individual importance of those terms on electric properties.

As mentioned above, the cavity surface is usually constructed using van der Waals radii, which are scaled by an empirically determined factor 1.1. This scaling factor is probably appropriate for neutral molecules in the ground state, but may be questionable if it is suitable for solutes with a very diffuse electron density such as anions; see, e.g., Serr and Netz (2006). Similarly, for molecules in electronically excited states, the effect of cavity size should be checked carefully, as the van der Waals radii determined for the electronic ground state may not be appropriate for excited states. As it may be anticipated, electric properties like the polarizability tend to the gas-phase values with increasing cavity size and, in general, increase with decreasing cavity size. For cavities, which are too small, a substantial part of the wave function is not confined anymore by the cavity, and electron charge density escapes the cavity; this is the so-called charge penetration effect.

In section “Definition of Macroscopic and Microscopic Quantities” the concept of *solute* and *effective* properties was described, following the definition of Wortmann and Bishop (1998). Only the latter are the properties actually related to experimentally determined values, while the former are in a way fictive properties as they cannot be determined directly. The distinction carries over to the SCRf models: in a standard PCM computation the *solute* properties are calculated. These may be

multiplied by Onsager–Lorentz cavity factors for the conversion to the effective properties. A more consistent way would be to use the implementation of Cammi et al. (1998), which allows the computation of *effective* (hyper)polarizabilities in the framework of the PCM directly. In their formulation the interaction of an effective dipole moment (μ^{eff}) with the Maxwell electric field \underline{E} is introduced by adding the following term to the total Hamiltonian:

$$\hat{H}' = \hat{\underline{\mu}} \cdot \underline{E} + \sum_{k=1}^N V_k \frac{\partial q_k^{\text{ex}}}{\partial \underline{E}} \cdot \underline{E} \quad (42)$$

where $\hat{\underline{\mu}}$ and V_k are the electronic dipole moment operator of the solute and the electronic electrostatic potential at the cavity surface, respectively. The charges q_k^{ex} describe that part of the solvent polarization which is induced by the external field and are computed using the (static or frequency-dependent) dielectric constant $\epsilon(\omega)$. This term describes the cavity field effect including the response of the solvent to the external field after the creation of the cavity in the solvent and allows the direct calculation of the effective (hyper)polarizabilities of the solvated molecule.

A rigorous treatment of cavity field effects within the PCM has been presented recently by Pipolo et al. (2014), bypassing the use of Onsager–Lorentz factors. The cavity field interactions within the PCM are described by factors reflecting the different cavities created by different molecules. Ferrighi et al. (2006) observed in their study of push–pull phenylpolyenes that PCM cavity field factors are smaller than the Onsager–Lorentz factors; moreover, their values decrease with increasing chain length. Such effects may even substantially change trends obtained in series of homologues, e.g., polyenes with increasing chain lengths.

Multipole Expansion (MPE) Method

The multipole expansion method (Mikkelsen et al. 1994a) is another SCRF method and has been frequently applied for the prediction of solute (hyper)polarizabilities. Its implementation in Dalton (2015) allows the analytical calculation of electric properties using linear, quadratic, and cubic response theory for singlets at HF, DFT, and multiconfigurational SCF (MCSCF) levels of theory, where it is possible to employ both equilibrium and nonequilibrium solvation. In the latter case, the charge distribution of the solvent is not in equilibrium with the charge distribution of the solute, which is relevant if optical external fields are applied, e.g., in the form of light. For (quasi)-static electric fields, on the other hand, equilibrium solvation would be used.

The MPE method is based on the polarization theory of Onsager, as extended by Kirkwood. The cavity in this model is a sphere, and the interaction of the solute with the continuum is described by a multipole expansion of the charge density of the solute (not just the dipole as in the Onsager theory). Thus, the polarization free

energy can be written as

$$\Delta G_{\text{solv}} = \sum_{m,l} g_l(R, \epsilon) |\langle \Psi | \hat{T}_{lm} | \Psi \rangle|^2 \quad (43)$$

where \hat{T}_{lm} are the spherical multipole operators and $g_l(R, \epsilon)$ are functions of the (frequency-dependent) dielectric constant $\epsilon(\omega)$, the cavity radius R , and the rank of the multipole l :

$$g_l(R, \epsilon) = -\frac{1}{2} R^{-(2l+1)} \frac{(l+1)(\epsilon-1)}{l+(l+1)\epsilon} \quad (44)$$

The maximum order l_{max} used in the expansion of Eq. 43 is in principle unlimited, but in many cases, results are converged with $l_{\text{max}} = 6$. Nevertheless, this should be checked in any specific application of the method.

In the quantum-mechanical formulation, the total Hamiltonian for the solute becomes

$$\hat{H} = \hat{H}^{\text{vac}} + \hat{W}(\Psi) \quad (45)$$

$$\hat{W}(\Psi) = 2 \sum_{ml} g_l(\epsilon) \langle \Psi | \hat{T}_{ml} | \Psi \rangle \hat{T}_{ml} \quad (46)$$

As the solvation part depends on the wave function, the corresponding equations to determine energy and wave functions have to be solved iteratively.

Using a spherical cavity, the cavity radius is an even more decisive factor in MPE than in PCM with its molecular-shaped cavity. A nonempirical way to uniquely determine this radius for pure liquids by comparing the results based on the Clausius–Mossotti equation with the SCRF predictions has been suggested in Luo et al. (1997). It turns out that the radius is quite dependent on the quantum-mechanical method applied as well as on the basis set.

A comparison of the performance of the MPE and PCM approaches with explicit models for electric properties is presented in section “Two Case Studies: Liquid Water and *p*-Nitroaniline in 1,4-Dioxane.”

Explicit Molecular Models

In explicit molecular models, the description of a solution as “a molecule embedded in a dielectric continuum” in implicit models is abandoned, and the molecules composing the liquid are treated explicitly, usually at different levels of theory. Even with the most approximate methods, this is clearly not possible for all the molecules composing the liquid, so further approximations are still required to render the

computation feasible. The models we will present below use one of the following strategies:

1. The molecules treated explicitly are contained in a box (or a sphere), which is surrounded by a dielectric continuum (discrete reaction field (DRF), Langevin dipoles (QM/LD/MC)).
2. Periodic boundary conditions are applied, which means that the box filled with molecules is replicated in space (rigorous local field RLF)).
3. A sufficiently large cluster of molecules is taken as a representation of the liquid (supermolecular (SM) approach, polarizable embedding (PE), frozen-density embedding (FDE)).

In addition to the treatment of inter- and intramolecular interactions, explicit models usually also take into account the pronounced *dynamic* nature of a liquid at the molecular level, both by considering the change of the geometric structure of individual molecules and by describing the macroscopic structure as an average over a large number of different individual arrangements of the molecules satisfying certain thermodynamic conditions (usually NVT, NVE, or NPT ensembles). The usual way to treat this problem of statistical mechanics is to apply the methods of molecular simulation, e.g., by Monte Carlo or molecular dynamics calculations, which yield a number of snapshots to be analyzed individually and averaged. This is clearly computationally much more demanding than a continuum model calculation, but is probably also more physically appropriate. Advantages of explicit models over continuum models are, for example, that they are able to describe specific interactions such as hydrogen bonding. In addition, they do not depend on the definition of cavity, which is in some way an artificial construct.

Fully quantum-mechanical simulations are possible, generally at the DFT level; e.g., Car–Parrinello molecular dynamics (CPMD) simulations have been used to model condensed phases including the determination of molecular electric properties (Dyer and Cummings 2006). Those methods are, however, still computationally rather demanding, and therefore a more common and less expensive approach is to combine a fully QM method with force field-based molecular mechanics (MM) methods. In such a hybrid QM/MM method, the central region (usually a solute molecule) is modeled at the QM level, while the environmental effects of the surrounding solvents are computed using MM force fields. The interactions between QM and MM regions are taken into account by including electrostatic and polarization effects as well as short-range repulsion and dispersion interactions. Thus, an effective hybrid QM/MM Hamiltonian is constructed in the following form:

$$\hat{H}_{\text{ef}} = \hat{H}^0 + \hat{H}_{\text{QM/MM}} + \hat{H}_{\text{MM}}, \quad (47)$$

where \hat{H}^0 is the many-body quantum-mechanical Hamiltonian, $\hat{H}_{\text{QM/MM}}$ is Hamiltonian describing the interactions between QM and MM regions (atoms), and \hat{H}_{MM} is defined by classical force fields.

In general, a successful application of explicit QM/MM methods is significantly more involved than a continuum model calculation. This is due to a variety of parameters and variables which need to be decided upon, such as (a) the choice of the MM force field, (b) the type of simulations (molecular dynamics or Monte Carlo simulations), (c) technical details regarding simulations (time step of simulations, number of solute and solvent molecules, number of configurations required for statistical averaging, etc.), and (d) choice of level of QM theory, basis set, and technique for the determination of electrical properties. Hence, the construction of a reliable, but efficient, QM/MM technique for the investigation of the nonlinear optical response of molecular systems in condensed phases requires extended development as well as numerical testing (e.g., calibration of the QM/MM potential).

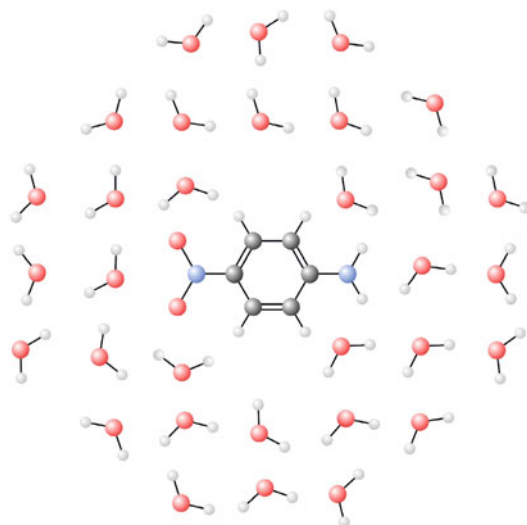
Usually, the determination of the liquid structure occurs separately from the computation of the electric properties, leading to sequential Monte Carlo (or molecular dynamics)/quantum mechanics (S-MC/QM or S-MD/QM) techniques (Canuto et al. 2005; Mata et al. 2009). The S-MC(MD)/QM methodology was successfully used in theoretical studies of the dynamic polarizability of liquid water as well as the polarizability of the fluoride anion (F^-) in water, where it was found that the polarizability of the F^- anion is significantly reduced compared to that of the free anion, as has been found previously for F^- in the crystal (Fowler and Madden 1984).

For a reliable structure determination of the liquid, the simulation box should be large enough to properly account, for example, for long-range electrostatic interactions. This is true even if periodic boundary conditions are applied; if the simulation box is too small, these may lead to artificial periodicity effects. In addition, the solute properties may significantly depend on the local geometry of solvent molecules in the vicinity of the solute. Consequently, particular attention should be paid to an accurate description of possible structural configurations of the first solvent shell. An alternative to the dynamic models may be to generate relevant configurations of the local environment of the solute ("microsolvation"). Marchal et al. 2009 proposed a method called the *global search algorithm of minima* (GSAM). While in dynamic models mean values of properties are obtained by averaging over the simulated configurations, in GSAM one relies on Boltzmann distribution factors. The GSAM method was successfully applied, e.g., in studies of the structure and properties of microhydrated nucleic acids.

A considerable number of papers using explicit models for the computation of macroscopic susceptibilities of liquids have been published. We present below various models which have been quite extensively employed, and/or are in active development. They differ in several aspects, for example, in the treatment of molecular electrostatic and polarization effects, which is clearly an important aspect for the computation of electric susceptibilities and will thus be treated in more details. Another relevant point is whether the method under consideration only allows the simulation of the solvent effect or it also takes the effect of the macroscopic field consistently into account. In the nomenclature of Wortmann

and Bishop (1998) mentioned above, we distinguish methods that stop after the computation of the solute properties, while others go all the way up to the effective properties. Solute properties in the first case may still be converted into susceptibilities using local field factors, which, however, would not be predicted by the theory in question.

Supermolecular (SM) Calculations



A relatively straightforward strategy to compute the NLO properties of aggregates of molecular systems is the supermolecular (SM) approach (Kanis et al. 1994). In this approach, the macroscopic system is approximated by a reasonably large cluster (the supermolecule) of a finite number of molecules, which are usually all described at the same quantum chemical level (the choice of the QM method is only restricted by the size of the cluster). There are no general rules for the structure of the clusters to be used in a SM calculation nor for the number of molecules to be included in the cluster. In the case of crystals, the information from experimental structure determinations (neutron diffraction or X-ray experiments) may be employed, but for liquids it is more customary either to rely on configurations generated by molecular simulations, which are then in some way averaged in order to create a single representative supermolecule, or to optimize chemically reasonable structures generated either by efficient search algorithms (e.g., in the global search algorithm of minima (GSAM) (Marchal et al. 2009)) or simply by using chemical intuition.

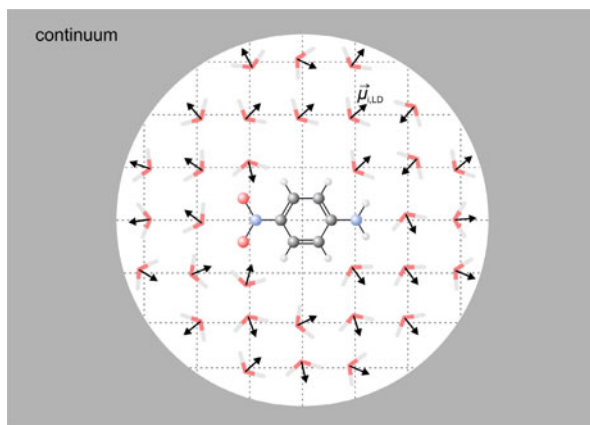
The size of the cluster should in principle be converged with respect to the property of interest. This can be a challenge in the case of NLO properties due to their strong dependence on long-range (e.g., dipole–dipole) electrostatic interactions, which may require the consideration of hundreds of molecules for a converged description. As this is usually computationally too expensive for standard correlated *ab initio* QM methods, one possibility to circumvent this problem is to use a multilayered approach, i.e., to apply different levels of theory for parts of the cluster (e.g., a high-level method for the solute, medium level for the first solvent shell, and a low level for more distant regions). Another possibility is to use a reasonable QM method for a smaller cluster and to combine this calculation with a continuum model. These multilayer approaches finally lead to the QM/MM models described later in this chapter. As for the SM method, it remains to note that the potentially high accuracy of this approach is often counterbalanced by the limited system size which can be treated. In addition one should mention the incapability of the SM method to account for dynamic structural effects.

The SM approach also suffers from the problem of distributing the computed response of the total cluster over the molecules which constitute the SM. See section “Splitting Schemes” for possible solutions based on the concept of interaction-induced properties.

The SM approach allows to study the dependence of the structure and the properties of specific molecular clusters (e.g., clusters with hydrogen or halogen bonds, π – π stacking, etc.), on specific structural parameters, e.g., mutual orientation, distances, etc. In a recent example, Suponitsky and Masunov (2013) considered several modes of stacking between donor–acceptor-substituted benzene and stilbene molecules, with different overlap of the corresponding molecular π systems. Several aspects connected with the type of intermolecular interactions due to the stacking were studied and their effect on the first hyperpolarizability of the aggregate analyzed. In another study using the SM method, a significant enhancement of the first hyperpolarizability of buckybowls in ordered aggregates was demonstrated (Li et al. 2013). Thus, the SM approach can lead to valuable insights in the effect of intermolecular interactions on NLO properties, which may also be very useful in the analysis of the NLO response of liquids.

As described in section “Implicit Solvent Methods,” most continuum models are not able to describe the effect of specific solute/solvent interactions, for example, hydrogen bonds. In particular, they often predict solvatochromic shifts as well as (hyper)polarizabilities of solute molecules dissolved in protic solvents incorrectly. An improved description is often achieved by including a small number of solvent molecules around the solute (e.g., the first solvent shell) in the explicit calculation while describing the rest of the solvent by the continuum model. Using this method, Wang et al. (2003) showed that H bonding has a large effect on the electronic structure and first hyperpolarizability of *p*-nitroaniline (pNA) in protic solvents. For a discussion of the application of this approach to liquid water, see section “Water.”

QM/Langevin Dipole/Monte Carlo (QM/LD/MC) Method



One of the first QM/MM implementations was described in Warshel and Levitt (1976). The authors used the semiempirical MNDO/2 Hamiltonian to describe the QM region. The MM part was represented by a polarizable force field and by Langevin dipoles (LDs) for the solvent molecules. In 1997, a modified and extended quantum-mechanical Langevin dipole/Monte Carlo (QM/LD/MC) technique was presented and applied to calculations of the solvent effect on molecular properties in electronic ground and excited states (Lipiński and Bartkowiak 1997). The new concepts in the classical LD/MC model and their QM/LD/MC extension included (i) the representation of a solute molecule by atomic charges, dipoles, and quadrupoles, (ii) the MC sampling technique for determination of the optimal position and orientation of the solute molecule, (iii) the use of the full Langevin formula for the orientational polarization in the self-consistent calculations, (iv) mutual polarization of the solute and solvent molecule.

In the QM/LD/MC model, the solvent molecules are represented by an ensemble of polarizable Langevin dipoles (LDs), which are distributed among the nodes of a regular cubic lattice with the center of mass of the solute in its center, extending up to a radius R , typically 12–18 Å, with a distance between two neighboring nodes (Δ) of ca. 3 Å. This sphere is then surrounded by a bulk solvent treated as a dielectric continuum with a bulk dielectric constant (ϵ) corresponding to the simulated solvent. The local field on the i -th node results from the charge density distribution of the solute molecule as well as from the electrostatic field generated by all other dipoles:

$$\underline{E}_i^{(n)} = \underline{E}_{i,0} + \underline{E}_{i,LD}^{(n)} \quad (48)$$

where $\underline{E}_{i,0}$ is the field generated by the solute molecule, $\underline{E}_{i,LD}^{(n)}$ is the field produced by other solvent molecules, and the upper index n refers to the n -th iterative step, since the fields are determined in a self-consistent manner reflecting the

mutual orientational as well as electronic polarization. In the LD model, the solvent polarization (expressed by the dipole moment of an i -th solvent molecule) is approximated by the Langevin-type function:

$$\underline{\mu}_{i,LD}^{(n+1)} = \underline{e}_i^{(n)} |\mu_p| [\coth(z_i^{(n)}) - (z_i^{(n)})^{-1}] \quad (49)$$

where $\underline{e}_i^{(n)}$ is the unit vector in the direction of $\underline{E}_i^{(n)}$, μ_p is the permanent dipole moment of the solvent molecule, and

$$z_i^{(n)} = \frac{|\mu_p| |E_i^{(n)}|}{kT} \quad (50)$$

where k is the Boltzmann constant and T denotes absolute temperature. In addition to the permanent dipole moment contribution defined by Eq. 50, the total solvent dipoles involve the electronically induced contribution evaluated as

$$\underline{\mu}_{i,LD,ind}^{(n+1)} = \underline{\alpha}_{LD} \cdot \underline{E}_i^{(n)} \quad (51)$$

where $\underline{\alpha}_{LD}$ is the polarizability of the solvent molecule.

The electrostatic potential generated by the solute on each solvent molecule represented by a LD is calculated using the cumulative atomic multipole moments (CAMP) scheme (Sokalski and Poirier 1983). In this formalism, each atom k of the solute molecule is represented by a scalar net atomic charge (q_k), a vector of atomic dipole ($\underline{\mu}_k$), as well as a tensor of atomic quadrupole (\underline{Q}_k). In addition, the polarization of the solute is modeled by atomic dipole polarizabilities. In the classical calculations, the optimum position and orientation of the solute molecule (treated as a rigid body), placed in a cubic grid of polarizable solvent molecules, is determined using the MC method. In each MC step, permanent and induced dipole moments of each solvent molecules and the electric potential produced by all solvent molecules on each atom of the solute are iteratively calculated.

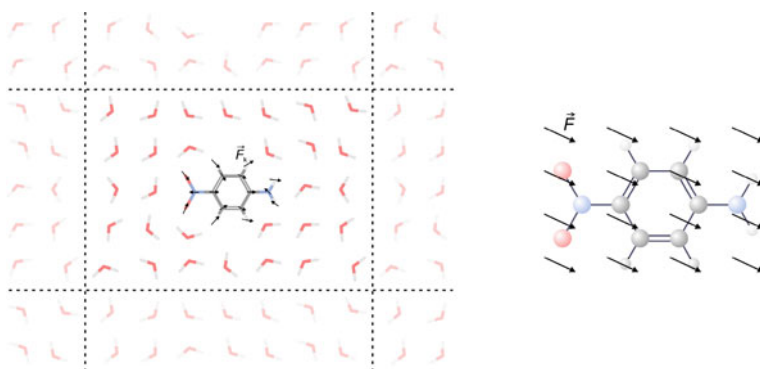
We note that the QM/LD/MC technique is less expensive than a standard classical molecular simulation. However, the QM/LD/MC model was only formulated for two polar solvents: water and chloroform. Hence, this technique is restricted to the description of a polar solute in either one of these two polar solvents. The QM/LD/MC model belongs to the sequential simulation techniques (S-QM/LD/MC). The permanent and induced dipole moments of the solvent molecules obtained in the MC run generate the electrostatic potential over the solute molecule. This potential is then averaged over the MC snapshots. Hence, the total potential V_R is a sum of averaged potentials due to the permanent and induced dipole moments of solvent molecules ($V_R = V_{perm} + V_{ind}$). The effect of the reaction field is as usually introduced into the solute Hamiltonian by means of perturbation operator \hat{V}_R according to the equation

$$\hat{H}_{ef} = \hat{H}^0 + \hat{V}_R \quad (52)$$

where \hat{H}^0 is the unperturbed Hamiltonian of an isolated solute molecule. After the QM calculations, new values of CAMM are obtained, which generate new electrostatic potential ($V_R = V_{\text{perm}} + V_{\text{ind}}$), which is then incorporated into Eq. 52 to give new CAMM. This process is repeated, until self-consistency is reached. In the context of *solute* versus *effective* properties, it should be noted that both of them can in principle be determined by the QM/LD/MC approach (applying the numerical derivative using finite fields (ND-FF) method). While the former refer directly to the properties computed from the solute density corresponding the effective Hamiltonian given by Eq. 52, the latter can be obtained in the same manner as in other combined D/C approaches by multiplying the solute properties with cavity field factors.

The QM calculations based on the QM/LD/MC technique were performed using the semiempirical GRINDOL program which uses an approach based on the neglect of differential overlap (NDO) method (Lipiński 1988). The majority of the investigations using this method were devoted to linear and nonlinear electrical properties of large organic molecules in solutions (Zaleśny et al. 2002).

Rigorous Local Field (RLF) Method



In the rigorous local field (RLF) method, a liquid is considered as an infinitely replicated simulation box, using Born–von Karman periodic boundary conditions to account for long-range electrostatic and polarization effects, very similar in spirit to the treatment of a crystal in solid-state physics. The theoretical basis of the RLF method applied to liquids has been developed by Janssen et al. (1999).

The molecular properties entering the electrostatic calculations are computed by QM methods. The liquid structure is usually obtained beforehand by a classical molecular simulation computation; thus, this method would be accurately described as a sequential MM/QM/ES (electrostatic) method. As long as the simulation box is large enough to avoid artificial periodic artifacts, the electrostatic treatment is probably more accurate, but at the same time computationally more demanding

than the supermolecular method or the continuum model, which considers the liquid either as a simulation box or even just the solute molecule surrounded by a dielectric continuum.

Once the liquid structure is available, it can be used to determine the permanent local field or reaction field according to Eq. 14, which determines the *solute* properties. Using electric properties of the isolated molecules (solute and solvent), computed by ab initio methods, Eqs. 14 and 15 are solved. If the field expansion of the dipole moments (and higher multipole moments) is truncated after the polarizability terms, an analytical solution is possible (Reis et al. 2006). In the few cases where it has been checked, this was found to be a valid approximation, i.e., the contribution of the hyperpolarizabilities to the reaction field was small. The magnitude of the permanent local field depends clearly on the electric properties of the solute and solvent. In the case of pure benzene, for example, the field was considered to be small enough to be neglected completely (Janssen et al. 1999). On the other hand, the contribution of *quadrupole* moments may be substantial, for example, in solutions of *p*-nitroaniline in nonpolar to medium polar solvents (Reis et al. 2006) or for pure acetonitrile (Avramopoulos et al. 2007).

With the solute properties determined, the susceptibilities can finally be computed using Eqs. 22, 23, and 24, possibly supplemented by the temperature-dependent terms mentioned after Eq. 21, as has been done in the simulation of the third-harmonic generation (THG) of signal $\chi(-3\omega; \omega, \omega, \omega)$ of pure benzene in Janssen et al. (1999). We note, however, that according to Kasproicz-Kielich and Kielich (1975), a T -dependent term should only occur for the optical Kerr effect (or intensity-dependent refractive index IDRI) $\chi(-\omega; \omega, \omega, -\omega)$, but not in the case of THG. On the other hand, checks performed in Reis et al. (2001) showed that the effect of these terms is small anyway for liquid benzene.

The theory developed by Janssen et al. (1999) is applicable for the special case of a liquid in an *optical* macroscopic field $\underline{E}(\omega)$. For such a field one may assume that the liquid structure is not affected by the field, which means that the field can be neglected in the molecular simulation step, i.e., the determination of the liquid structure. If all applied fields are optical, the liquid is isotropic, so $\chi^{(2)}$ should be zero, which yields a condition which may be used to check the correctness of the details of the simulation, for example, simulation time or number of configurations.

We note that the formulation of the local field problem in terms of the *macroscopic field* E and Lorentz factor tensors \underline{L} in Eq. 17, instead of the external field E^0 and the dipole-dipole propagator tensors \underline{T}^{11} , takes care automatically of the *cavity field factors*, which need to be considered separately for the calculation of susceptibilities by methods which contain a dielectric continuum as part of the description of the liquid system.

In several applications of RLF, e.g., for acetonitrile (Avramopoulos et al. 2007), water, and benzene (Reis et al. 2001), the molecules considered and their properties were small enough so that the one-point expansion of the electric properties was an accurate approximation. For larger molecules such as *p*-nitroaniline, however, this was not the case. Here, a very simple distributed model called the *submolecule*

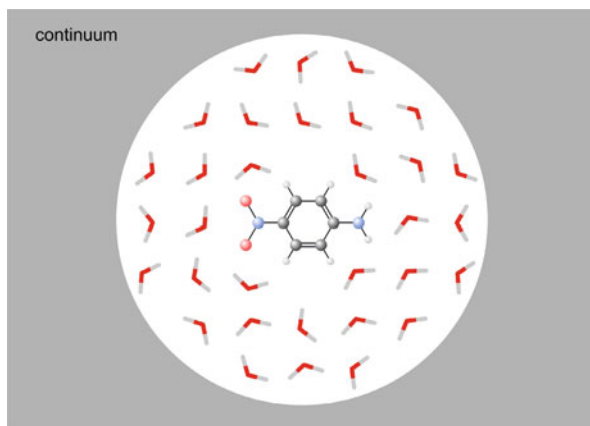
approach was applied. In this model, the molecule was partitioned into n smaller units, and of each electric property P , a part P/n was assigned to each subunit. Thus, for example, if μ is the permanent dipole moment of the molecule, each submolecule carries a dipole moment μ/n . Although quite simple and crude, this distribution scheme was found to be surprisingly successful.

As mentioned above, only for all-optical external fields one may assume that the liquid structure is undisturbed, and a field-free molecular simulation is sufficient to obtain this structure. If an applied field component is *static*, as in an electric field-induced second harmonic generation (EFISH) experiment, the isotropy of the liquid will in general be destroyed. Thus, a static field will influence the liquid structure and needs to be included already in the simulation step. In that case a nonzero $\chi^{(2)}$ should be obtained from Eq. 23. This approach has been used to compute the EFISH susceptibility of liquid acetonitrile in Avramopoulos et al. (2007). The total EFISH signal $\chi_{iiii}^{(3)}(-2\omega; \omega, \omega, 0)$ was computed by summing the contribution due to the orientational effect of the static field via the permanent dipole moments and the “intrinsic” third-order effect due to the second hyperpolarizability γ :

$$\chi_{zzzz}^{(3)}(-2\omega; \omega, \omega, 0) = \frac{\chi_{zzz}^{(2)}(-2\omega; \omega, \omega; E_z)}{3E_z} + \chi_{zzzz}^{(3\gamma)}(-2\omega; \omega, \omega, 0) \quad (53)$$

An alternative to a field-dependent molecular simulation for the description of static field effects is the assumption of a Boltzmann distribution in the field; see Eq. 65 in section “Discrete Reaction Field (DRF) Method.”

Discrete Reaction Field (DRF) Method



The discrete reaction field (DRF) method (Jensen et al. 2003) is based on the model devised by Lorentz to derive the local field in an isotropic liquid. In this model, the liquid is partitioned into two parts. The first part is a sphere around the molecule of interest, which is microscopically large but macroscopically small and in which the

molecules are considered as discrete entities. The second part is the rest of the liquid, which is described as a dielectric continuum with a relative permittivity $\epsilon(\omega)$. If an external electric field is applied to the liquid, the local field acting on the molecule in the center of the sphere (the solute) can be expressed as

$$\underline{F}^{\text{loc}} = \underline{E}^{\text{mac}} + \frac{4\pi}{3} \underline{P} + \underline{F}^{\text{disc}} \quad (54)$$

where the first two terms comprise the macroscopic contribution due to the macroscopic field and the polarization of the sphere and the third term is the discrete contribution from the molecules inside the sphere. Lorentz showed that for randomly distributed nonpolar molecules, this third term is zero. This assumption is certainly not valid any more for polar molecules and has been amended, e.g., by Onsager through the introduction of a reaction field (comp. Eq. 26). The DRF method tries to compute the field more accurately, firstly by performing a molecular simulation of the molecules in the sphere and subsequently by computing the electric properties of the solute in the presence of the total field by a quantum-mechanical DFT computation. In the simulation step, a classical molecular dynamics simulation of the molecules in the sphere is performed, using the discrete reaction field polarizable force field and a soft wall force potential to prevent the molecules from leaving the sphere. A subset of the configurations is kept and subjected to a QM/MM calculation, where the solute is treated quantum-mechanically, with the remaining solvent molecules contributing an electrostatic term to the solute Hamiltonian via partial charges and a polarization term via distributed polarizabilities.

The force field applied in DRF consists of a standard 6-12 Lennard–Jones potential for the repulsion–dispersion interactions, partial charges for the electrostatic component, and distributed polarizabilities for the polarization interaction, taken from the Applequist model with Thole damping. The Applequist model considers the total molecular polarizabilities as emerging from a set of *interacting* atomic polarizabilities. The atomic multipoles create internal fields, which induce multipoles on all other atoms (*atomic self-polarization* is excluded). This model has been developed by Applequist et al. (1972), first for fields and dipole polarizabilities and later for higher-rank fields and polarizabilities. The DRF method uses partial charges and isotropic atomic dipole polarizabilities, although the model can be easily generalized to anisotropic polarizabilities. This model, called the dipole-interaction model (DIM), has also been extended to second hyperpolarizabilities by Jensen et al. (2003). In the DIM, the induced dipole of site k of a molecule subject to an (possibly inhomogeneous) electric field \underline{E} is a sum of two contributions:

$$\mu_{k,\alpha} = \alpha_k \left[E_{k,\alpha} + \sum_{k' \neq k} T_{\alpha k, \beta k'} \mu_{k',\beta} \right], \quad (55)$$

where the first term takes the external electric field into account and the second term considers the effect of the fields arising from induced dipoles from the other sites in

the molecule. Together with the equation $\mu_\alpha^{\text{mol}} = \sum_k \mu_{k,\alpha}$ and the definition of the molecular polarizability $\mu_\alpha^{\text{mol}} = \alpha_{\alpha\beta}^{\text{mol}} E_\beta$, a condition for the α_i can be deduced:

$$\alpha_{\alpha\beta}^{\text{mol}} = \sum_{ij} [\alpha_i^{-1} + T_{ij}]^{-1} \quad (56)$$

In combination with the assumption that the atomic polarizabilities are *transferable* between molecules, the condition of Eq. 56 is sufficient to derive atomic polarizabilities by fitting them to a set of known molecular polarizabilities, either taken from experiment or from high-level ab initio computations. A problem which remains in this description of interacting *polarizable points* is that the electrostatic interaction may grow without bounds with decreasing distance between the points. To circumvent this so-called *polarization catastrophe*, Thole (1981) proposed to replace the interaction between point charges by the interaction of spherically smeared-out charge densities.

One advantage of this interaction model is that it is able to describe the electrostatic interaction of flexible molecules, without a need to reparametrize the atomic polarizabilities. The distributed polarizability model of section “Distributed Multipole and Polarizability Analysis,” on the other hand, would require a new calculation of the distributed properties for each different geometry.

The partial charges and distributed polarizabilities are also used in the QM/MM calculations. The coupling Hamiltonian $\hat{H}_{QM/MM}$ of Eq. 47 is given by

$$\hat{H}_{QM/MM}(r) = \sum_s \left[\frac{q_s}{|r - \underline{R}_s|} + \mu_{s,\alpha}^{\text{ind}} \frac{(r_\alpha - R_{s,\alpha})}{|r - \underline{R}_s|^3} \right], \quad (57)$$

where the summation is over all atoms s of the surrounding solvent molecules. The induced dipoles μ_s^{ind} have to be computed self-consistently via an equation similar to Eq. 55, where the external field \underline{E}_k is replaced in the first step of the self-consistent procedure by the initial field $\underline{F}_s^{\text{init}}$:

$$\underline{F}_s^{\text{init}}(\omega) = \underline{F}_s^{\text{QM}}(\omega) + \delta_{0\omega} \underline{F}_s^{\text{MM}} + \underline{F}_s^{\text{mac}}(\omega) \quad (58)$$

where the first two terms are the fields due to the charge density of the solute and the solvent molecules, respectively, and the last term is the macroscopic field. Due to partial charges and the nonzero charge density of the solute, there will be a nonzero field even in the absence of the macroscopic field term, and in effect, in this case the *solute* properties as defined, e.g., in Eq. 15 will be obtained. If the macroscopic field is included, the DRF implementations will compute instead the *effective* properties of the solute. These are defined in the DRF model via an effective (optical) macroscopic field $\underline{F}^{\text{eff}} = \underline{F}^{\text{mac}} + \frac{4\pi}{3} \underline{P}$, i.e.,

$$\begin{aligned} \mu_\alpha^{\text{ind}}(\omega) &= \alpha_{\alpha\beta}^{\text{eff}}(\omega) F_{\beta,k}^{\text{eff}}(\omega) + K_{12}^{(2)} \frac{1}{2} \beta_{\alpha\beta\gamma}^{\text{eff}}(-\omega; \omega_1, \omega_2) F_\beta^{\text{eff}}(\omega_1) F_\gamma^{\text{eff}}(\omega_2) + \\ &K_{123}^{(3)} \frac{1}{6} \gamma_{\alpha\beta\gamma\delta}^{\text{eff}}(-\omega; \omega_1, \omega_2, \omega_3) F_\beta^{\text{eff}}(\omega_1) F_\gamma^{\text{eff}}(\omega_2) F_\delta^{\text{eff}}(\omega_3) \dots \end{aligned} \quad (59)$$

The macroscopic susceptibilities in the DRF method are finally obtained by multiplication of the effective properties with the cavity-field factors $L(\omega)$ for optical frequencies ω , e.g., for the linear and THG effects by

$$\chi_{ZZ}^{(1)}(\omega) = NL(\omega) \langle \alpha_{\alpha\beta}^{\text{eff}}(\omega) \rangle_{ZZ} \quad (60)$$

$$= N \frac{\langle \alpha_{\alpha\beta}^{\text{eff}}(\omega) \rangle_{ZZ}}{1 - \frac{4\pi}{3} \langle \alpha_{\alpha\beta}^{\text{eff}}(\omega) \rangle_{ZZ}} = N \frac{\bar{\alpha}^{\text{eff}}(\omega)}{1 - \frac{4\pi}{3} \bar{\alpha}^{\text{eff}}(\omega)} \quad (61)$$

$$\chi_{ZZZZ}^{(3)}(-3\omega; \omega, \omega, \omega) = \frac{N}{6} L(3\omega) L^3(\omega) \langle \gamma_{\alpha\beta\gamma\delta}^{\text{eff}}(-3\omega; \omega, \omega, \omega) \rangle_{ZZZZ} \quad (62)$$

$$= \frac{N}{6} L(3\omega) L^3(\omega) \bar{\gamma}_{||}^{\text{eff}}(-3\omega; \omega, \omega, \omega) \quad (63)$$

where $\bar{\alpha}^{\text{eff}} = \sum_{\alpha} \alpha_{\alpha\alpha}^{\text{eff}}$ and $\bar{\gamma}_{||}^{\text{eff}} = \frac{1}{15} \sum_{\alpha\beta} (\gamma_{\alpha\alpha\beta\beta}^{\text{eff}} + \gamma_{\alpha\beta\alpha\beta}^{\text{eff}} + \gamma_{\beta\beta\alpha\alpha}^{\text{eff}})$ are the orientationally averaged effective polarizability and second hyperpolarizability, respectively. Equation 60 defines the cavity field factor $L(\omega)$ in the model and shows how it may be computed in DRF. In the case of a *static* field component, however, the cavity field factors are computed via the Onsager equation:

$$L(0) = \frac{\epsilon(0)(2 + n^2(0))}{2\epsilon(0) + n^2(0)}, \quad (64)$$

where $\epsilon(0)$ and $n(0)$ are the linear dielectric constant and the refractive index, respectively, at zero frequency. The computation of $\epsilon(0)$ by simulation methods is quite cumbersome and is thus usually taken from experiment in the DRF model. On the other hand, $n(0)$ can be obtained via $n(0) = \sqrt{1 + 4\pi\chi_{ZZ}^{(1)}(\omega)}$.

In the same spirit, the EFISH effect $\chi_{ZZZZ}^{(3)}(-2\omega; \omega, \omega, 0)$, which contains a static field component, is not computed via a molecular simulation with applied fields. As mentioned above, an alternative is the assumption of a Boltzmann distribution in the field and an expansion of the exponential terms, assuming that the field effect is small compared to kT (Reis et al. 2006). Then,

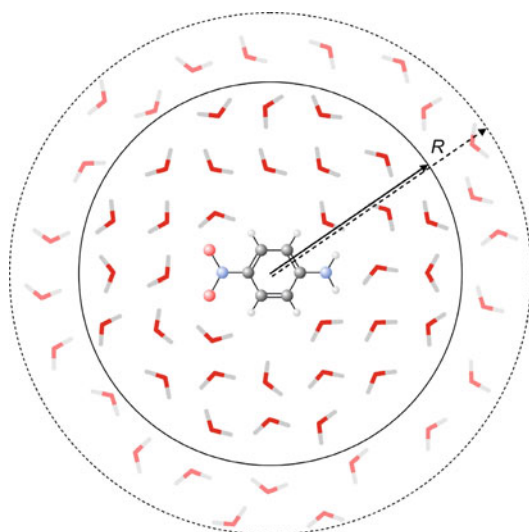
$$\chi_{ZZZZ}^{(3)}(-2\omega; \omega, \omega, 0) = \frac{N}{6} L(2\omega) L^2(\omega) L(0) \langle \gamma^{\text{eff}}(-2\omega; \omega, \omega, 0) \rangle_{ZZZZ} \quad (65)$$

$$= \frac{N}{6} L(2\omega) L^2(\omega) L(0) \left(\bar{\gamma}_{||}^{\text{eff}}(-2\omega; \omega, \omega, 0) + \frac{\underline{\mu}^{\text{sol}} \cdot \underline{\beta}^{\text{eff}}(-2\omega; \omega, \omega)}{3kT} \right) \quad (66)$$

where $\underline{\beta}^{\text{eff}}$ is the *vectorial part* of the averaged effective first hyperpolarizability, with components $\beta_i = \frac{1}{3} \sum_{j=1}^3 \beta_{ijj} + \beta_{jij} + \beta_{jji}$. Instead of $\underline{\beta}$ one often encounters the quantity $\beta_{||} = \frac{3}{5} \underline{\beta}$.

As mentioned in section “Definition of Macroscopic and Microscopic Quantities,” other authors define effective properties via the macroscopic field $\underline{F}^{\text{mac}}$, e.g., Wortmann and Bishop (1998) and Cammi et al. (2000), instead of the effective macroscopic field (which, incidentally, is just the local field of the Lorentz model).

Polarizable Embedding (PE)



The polarizable embedding (PE) (Olsen and Kongsted 2011) is a quite general method for the calculation of molecular properties in liquids. It is a QM/MM method with a separate simulation step, which creates a set of configurations. Molecular dynamics with classical, polarizable, or nonpolarizable force fields are usually applied for this step. Selected configurations are then subjected to the QM/MM calculation, using response theory to compute the property of interest of the solute molecule at the QM level self-consistently in the polarization field of the surrounding solvent molecules, described by MM. A unique feature of PE is that it allows to employ different levels of theory at the QM step, including Hartree–Fock (HF) and DFT (Olsen et al. 2010), MP2 and the second-order polarization propagator approximation (SOPPA) (Eriksen et al. 2012), multiconfigurational SCF (MCSCF) (Hedegård et al. 2013), and several coupled-cluster levels (Sneskov et al. 2011), although not all levels were already implemented for each property

at the time of writing. It also allows a large variety of molecular properties to be computed, e.g., nonlinear optical properties, one- and two-photon absorption spectra, NMR properties, etc. PE is implemented in the program packages Dalton (2015) and TURBOMOLE (<http://www.turbomole.com>). The description below is mainly based on the implementation of PE in Dalton (2015). The implementation in TURBOMOLE differs mainly by the possibility to use the resolution of identity (RI) in conjunction with the CC2 method for the description of ground and excited electric-state properties (PERI-CC2), unfortunately, however, not for (hyper)polarizabilities. In Dalton, singlet linear response is implemented for HF, DFT, SOPPA, MCSCF, CC2, and CCSD for closed-shell systems and for HF and DFT for open-shell systems. Singlet quadratic response is implemented for HF and DFT for closed-shell systems. Using finite field techniques, partially static (hyper)polarizabilities are also available.

As the molecular simulation and the QM/MM calculation are separated, any kind of simulations may be used, and we will only remark here on some points connected with the QM/MM step as implemented in the Dalton program package.

The total Hamiltonian is partitioned according to Eq. 47, with the interaction term $\hat{H}_{QM/MM}$ defined as

$$\hat{H}_{QM/MM} = \hat{H}^{\text{es}} + \hat{H}^{\text{pol}} + \hat{H}^{\text{vdW}} \quad (67)$$

where \hat{H}^{es} is the electrostatic interaction between the solute (treated by QM) and the multipole moments of the solvent molecules (the MM part). The latter may be, e.g., just point charges or permanent multipole moments, either molecular or distributed. The induction term \hat{H}^{pol} may be written as

$$\hat{H}^{\text{pol}} = -\frac{1}{2} \sum_a \underline{\mu}_a^{\text{ind}} \cdot \underline{F}^{QM}(\underline{r}_a) \quad (68)$$

where the summation runs over all MM molecules, $F^{QM}(\underline{r}_a)$ is the electric field at a MM molecule due to the solute molecule (electrons and nuclei), and $\underline{\mu}_a^{\text{ind}}$ is the induced dipole moment of the MM molecule a , given in linear approximation by

$$\underline{\mu}_a^{\text{ind}} = \underline{\alpha}_a \cdot \underline{E}_a^{\text{total}} \quad (69)$$

where $\underline{\alpha}_a$ is the dipole–dipole polarizability of the MM molecule a and $\underline{E}_a^{\text{total}}$ is the total electric field on a , due to the solute and the permanent and induced moments of the MM molecules. Thus, this field, as well as the field $\underline{F}^{QM}(\underline{r}_a)$ in Eq. 68, depends in turn on $\underline{\mu}_a^{\text{ind}}$. The set of coupled equations require thus either an iterative solution or an analytical inversion of the full matrix equation. This has to be repeated for each step of the wave function optimization.

Although only dipole–dipole polarizabilities are taken into account, they may be distributed over the molecules in the form of *localized* polarizabilities (i.e., without intermolecular polarizabilities (Misquitta 2015)).

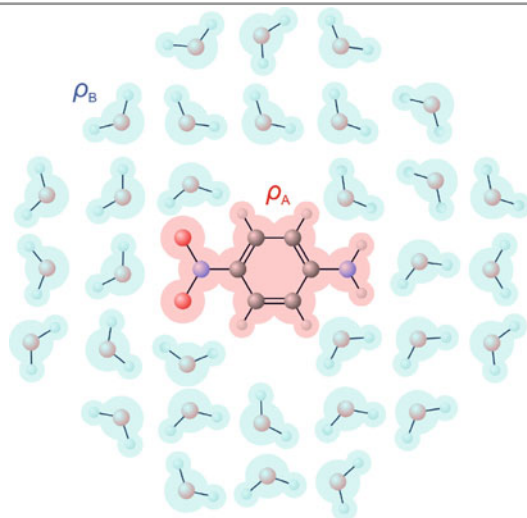
Finally, the term \hat{H}^{vdW} takes into account dispersion and short-range repulsion using a 6-12 Lennard–Jones potential, which depends only on the position of nuclei.

By combining Eqs. 47 and 67, 68, and 69, the equations to solve for the quantum-mechanical method considered, i.e., CC, DFT, etc., can be finally set up. It turns out that the equations are quite similar to the vacuum case, extended by some additional terms considering the new QM/MM interactions. A point to note is that the whole system of equations is solved to self-consistency. With the optimized wave function thus obtained, first-order properties of the solute molecule can then be computed. For higher-order properties, response theory is applied in conjunction with the QM method considered. This leads again to certain modifications of the corresponding vacuum equations. For details we refer to the literature in Nielsen et al. (2007), Osted et al. (2006), and references therein.

We note that the QM/MM step is usually performed on a cluster of the solute embedded in solvent molecules, which is extracted from the larger simulation box of the classical molecular dynamics or Monte Carlo calculation. The size of the cluster is varied until the properties of interest are converged with respect to cluster size, which may require different sizes depending on the property considered. With the converged cluster size determined, clusters of similar size are then extracted from a sufficiently large number of snapshots from the simulation trajectory and their properties computed with the QM/MM method and averaged to yield the final simulated value for the solution. The properties extracted in this way are therefore *solute* properties in the definition of Wortmann and Bishop (1998). Effective properties and thus also susceptibilities are computed by multiplying the solute properties with the corresponding cavity field factors obtained in the Lorentz model, usually computed using the polarizabilities computed within the PE model, thus only assuming that the form of the cavity field factors is appropriate.

Frozen-Density Embedding (FDE)

A weak point of the QM/MM models described so far is that the MM description of the solvent environment and solute–solvent interactions relies in a large extent on the predominantly electrostatic nature of these interactions. Although this assumption is valid in many cases, it can be expected that QM effects such as the exchange repulsion and dispersion interactions can play an important role especially in the inner solvent shell and these effects are either neglected or modeled only indirectly in an empirical way in the QM/MM approach. A careful parameterization of the force field (FF) used in the MM part can in some cases yield more accurate results than large-scale DFT calculations, but it is unclear if the given FF parameterization is generally applicable. Moreover, the role of particular interaction contributions in the description of the interaction energy and interaction-induced properties is not the same (see section “Interaction-Induced Properties”), and thus the good performance of a selected FF does not automatically guarantee its applicability to molecular property calculations.



As a compromise between the full QM treatment based on supermolecular calculations and combined QM/MM models, Wesolowski and Warshel (1993) proposed an alternative scheme within the DFT framework, in which the frozen electron density of the solvent is used to construct an embedding potential that enters the Hamiltonian in the calculation of the solute properties. In this so-called *frozen-density embedding* (FDE) scheme, both subsystems (solute and solvent) are treated at a QM level (not necessarily the same), but the electron density of the subsystems are determined separately, so the full calculation for the supersystem is avoided. This makes the QM calculations feasible for a solute molecule surrounded by physically reasonable number of solvent molecules.

The FDE variants can be classified according to (i) the degree of allowed relaxation of subsystem electron densities, (ii) the QM levels used to describe the solute and the solvent, and (iii) the approximation of nonadditive kinetic energy functional (see further). In the simplest model, the solvent electron density is kept frozen in the molecular (response) property calculation within a limited orbital space of the solute. The solvent embedding potential can be calculated at a low QM level or even constructed from the sum of the electron densities of isolated solvent molecules by using the sum-of-molecular-fragments approach (Neugebauer et al. 2005). This approximated electron density of the environment can be improved by including the relaxation effects through so-called *freeze-and-thaw* cycles (Wesolowski and Weber 1993), in which the roles of the frozen and nonfrozen systems are interchanged. Depending on the size and character of the system (e.g., the presence of the hydrogen bonds), one can decide for a full (i.e., reaching the self-consistency) or partial relaxation schemes. In the latter case, a typical choice is to update the solvent electron density in two freeze-and-thaw cycles. Obviously, the described approximations can be combined to achieve a reasonable compromise between the accuracy and computational demands. For example, it can appear useful to split the solvent environment into inner and outer

shells and to treat the former by the freeze-and-thaw strategy while for the latter one can use the sum-of-molecular-fragments density (Jacob et al. 2006). In addition, different DFT levels can be used for the solute molecule and inner and outer solvent shells, respectively. In the following we briefly summarize basic features of the FDE approach, and we put them in the context of molecular property calculations.

As already mentioned, in the FDE formalism, the total electron density (ρ_{tot}) of the cluster is split into two (or more) parts – the electron density of the solute (ρ_A) and that of the solvent (ρ_B):

$$\rho_{\text{tot}}(\underline{r}) = \rho_A(\underline{r}) + \rho_B(\underline{r}) \quad (70)$$

Under the assumption that the electron density ρ_B of the solvent is known and kept frozen, ρ_A can be obtained by minimizing the energy of the total system $E[\rho_{\text{tot}}]$. Following the idea of the Kohn–Sham (KS) method, ρ_A can be expressed in terms of KS orbitals (for a closed-shell solute molecule containing N_A electrons):

$$\rho_A(\underline{r}) = 2 \sum_{i=1}^{N_A/2} \phi_i^*(\underline{r})\phi_i(\underline{r}) \quad (71)$$

which can be determined through KS-like equations in which the effect of the frozen solvent electron density is included by an embedding term in the effective potential for ρ_A (in a.u.):

$$\left[-\frac{\nabla^2}{2} + V_{\text{eff}}^{\text{KS}}[\rho_A](\underline{r}) + V_{\text{eff}}^{\text{emb}}[\rho_A, \rho_B](\underline{r}) \right] \phi_i(\underline{r}) = \epsilon_i \phi_i(\underline{r}) \quad i = 1, 2, \dots, N_A/2 \quad (72)$$

where $V_{\text{eff}}^{\text{KS}}$ is the effective KS potential of the isolated solute. The embedding potential $V_{\text{eff}}^{\text{emb}}[\rho_A, \rho_B]$ can be written as

$$\begin{aligned} V_{\text{eff}}^{\text{emb}}[\rho_A, \rho_B](\underline{r}) &= v_B(\underline{r}) + \int \frac{\rho_B(\underline{r}')}{|\underline{r} - \underline{r}'|} d\underline{r}' + V_{\text{XC}}[\rho_A + \rho_B](\underline{r}) \\ &\quad - V_{\text{XC}}[\rho_A](\underline{r}) + \frac{\partial T_s^{\text{nadd}}[\rho_A, \rho_B]}{\partial \rho_A(\underline{r})} \end{aligned} \quad (73)$$

with $v_B(\underline{r})$ being the field of clamped nuclei in the subsystem B (solvent) and V_{xc} is the exchange-correlation potential. The last term in Eq. 73 is the functional derivative of the nonadditive kinetic energy functional, which is defined by the following decomposition of the total kinetic energy:

$$T_s^{\text{nadd}}[\rho_A, \rho_B] = T_s[\rho_A, \rho_B] - T_s[\rho_A] - T_s[\rho_B] \quad (74)$$

Although this scheme is exact in the exact functional limit, in practical calculations, an approximate density-dependent functional is used for the nonadditive kinetic energy functional and its functional derivative. It was shown by Casida and

Wesołowski (2004) that the FDE strategy can be combined with the time-dependent response theory, and thus it can also be applied in molecular property calculations within the TDDFT framework. If we assume that the response to an external electromagnetic field is localized to the embedded system (solute), i.e., that the response of the frozen environment (solvent) can be neglected, one can obtain solute (hyper)polarizabilities by solving the conventional TDDFT response equations with the kernel modified by adding an effective embedding contribution, which is evaluated for the ground-state density of the solvent. In consistency with conventional TDDFT, in which the adiabatic local density approximation (ALDA) is applied for the kernel, the embedding contribution containing the nonadditive kinetic energy functional derivative is also approximated by using the Thomas–Fermi density-dependent functional. Let us note that using the ALDA approximation in linear response property calculations can lead to slight discrepancies when compared to finite field property calculations.

Using the described TDDFT/FDE approach, Jacob et al. (2006) evaluated the mean polarizability and its anisotropy for a water molecule surrounded by 127 water molecules. They have shown that FDE predicts a decrease of $\bar{\alpha}$ in solution by about 7% compared to the isolated molecule, which was in contradiction with the DRF approach predicting even a slight increase of this quantity. This qualitative disagreement was explained in terms of different consideration of the short-range (mainly exchange repulsion) and polarization effects in the two approaches. In DRF, the direct solvent effect on the HOMO–LUMO gap (leading to the decrease of $\bar{\alpha}$) is overcompensated by the response of the solvent to the polarization of the solvated molecule. This response stabilizes the polarized molecule and therefore causes an increase in the solute polarizability. On the other hand, in the TDDFT/FDE calculations, the polarization of the solvent is neglected, while the hydrogen bonding and Pauli repulsion are accounted for, which naturally results in the decrease of $\bar{\alpha}$. In order to include the solvent polarization effects in the FDE calculations, Jacob et al. (2006) proposed to calculate the static polarizabilities by the ND-FF method from the change of the dipole moment due to an external electric field. Since the electron density is split in FDE, the ND-FF/FDE approach enables to control whether the external field directly affects only the solute molecule (solvent molecules are then affected only indirectly through their response to the polarization of the solute) or the field is imposed on the whole cluster. While the former method leads to the solute properties, the latter enables to determine effective properties.

Distributed Multipole Moments and (Hyper)polarizabilities

As mentioned in section “Definition of Macroscopic and Microscopic Quantities,” a converged description of the reaction field acting on a solute may require high-order multipoles as well as reaction field gradients in the one-point multipole description. This is especially problematic for large or nonspherical solutes as discussed by Misquitta (2015) in this handbook. A solution for this problem is to use several points for the multipole expansions, i.e., to distribute the multipole moments and

(hyper)polarizabilities in some way over the molecules. There have been several distribution schemes developed, all with their own advantages and disadvantages (see, e.g., Stone 2013, Chs. 7 and 9). Some of them have already been mentioned in previous subsections. Below we describe two approaches, which have been previously employed for the calculation of macroscopic susceptibilities, although not yet for liquids, and/or which are implemented in generally available program packages.

Distributed Multipole and Polarizability Analysis

The distributed multipole analysis (DMA) by Stone is available in several program packages, such as GDMA (<http://www-stone.ch.cam.ac.uk/pub/gdma/index.php>), and others. A fully distributed polarizability approach has been developed by Stone and Misquitta, and it is implemented in the program package CamCASP (Misquitta and Stone 2007). Both approaches are also described in Misquitta (2015). Unfortunately, this approach has not been extended to hyperpolarizabilities. Fully distributed (first) hyperpolarizabilities have been implemented using the Atoms-in-Molecules (AIM) approach by Bader (Reis et al. 2000), but this approach is not implemented in any publicly available program package. Distributed polarizabilities using the AIM approach have been computed and employed in interaction studies of dimers, but again no generally available implementation is known to us. For details of computations of distributed electric properties, see Stone (2013); here, we only point out the consequences of using a distributed scheme for the calculation of macroscopic susceptibilities.

In a fully distributed model, the charge density of a molecule is partitioned according to some scheme over several sites, which may be atoms, small groups of bonded atoms, or some other sites not coinciding with any nucleus, e.g., the middle of a bond or a nonnuclear attractor. This scheme may be based on real-space partitioning, as in the case of AIM, or based on properties of Gaussian functions used to describe the wave function in the case of DMA. The partitioning will assign to each site t a charge q_t , a dipole μ_t , and higher multipoles. In an external field, each site t of the molecule will develop on-site polarizabilities α_{tt}^{ab} , i.e., charge–charge ($a = b = 0$), charge–dipole ($a = 0, b = 1$ or $a = 1, b = 0$), and dipole–dipole ($a = b = 1$) polarizabilities, if we restrict ourselves to the dipole approximation. In addition there will be site–site polarizabilities α_{tu}^{ab} , which describe the change of multipole a at site t due to the b -th derivative of the potential at site u (or vice versa). As the potentials on each site do not need to be equal among the sites, this description allows the treatment of inhomogeneous fields in a natural way, even for extended molecules.

In order to describe the permanent local field in terms of distributed properties, we need to introduce the potential $\Phi_{tk}^{(0)}$ and the local field $F_{tk}^{(0)} = -\nabla\Phi_{tk}^{(0)}$ at a point in site t of a molecule k . Then, the set of equations determining the permanent local field in the distributed treatment which correspond to Eqs. 14 and 15 in the one-point expansion can be written as

$$-\Phi_{tk}^{(0)} = \sum_{l \neq k} \sum_{u \in l} \left[T_{tk;ul}^{00} q_{ul} + T_{tk;\beta ul}^{01} \mu_{\beta u,l} + \dots \right] \quad (75)$$

$$F_{\alpha tk}^{(0)} = \sum_{l \neq k} \sum_{u \in l} \left[T_{\alpha tk;ul}^{10} q_{ul} + T_{\alpha tk;\beta ul}^{11} \mu_{\beta u,l} + \dots \right] \quad (76)$$

and

$$q_{tk} = q_{tk}^{(0)} + \sum_{u \in k} \left[-\alpha_{tu,k}^{00} \Phi_{uk}^{(0)} + \alpha_{t\alpha u,k}^{01} F_{\alpha u,k}^{(0)} + \dots \right] \quad (77)$$

$$\mu_{\alpha t,k} = \mu_{\alpha t,k}^{(0)} + \sum_{u \in k} \left[-\alpha_{\alpha tu,k}^{10} \Phi_{uk}^{(0)} + \alpha_{\alpha t\beta u,k}^{11} F_{\beta u,k}^{(0)} + \dots \right] \quad (78)$$

The distribution model can be extended beyond polarizabilities, which leads first to distributed first hyperpolarizabilities β_{tuv}^{abc} , describing the change of multipole a at site t due to the simultaneous application of the b -th order derivative of the potential at site u and the c -th order derivative of the potential at site v . Up to $a = b = c = 1$, there are already eight tensors, although not all of them are different, because it is, e.g., $\beta_{tuv}^{100} = \beta_{utv}^{010} = \beta_{vut}^{001}$. A similar symmetry relation holds for α_{uv}^{ab} . Nevertheless, the number of nonzero values of fully distributed (hyper)polarizabilities can be quite large, and several localization methods have been devised to reduce the number of values, e.g., for interpretational purposes. For the computation of susceptibilities, such a localization does not seem to be necessary.

The molecular (hyper)polarizabilities can be reconstructed from the corresponding distributed properties, e.g., for the (dipole–dipole) polarizability α and (dipole–dipole–dipole) hyperpolarizability β :

$$\alpha_{\alpha\beta} = \sum_{tu} \left[r_{\alpha t} r_{\beta u} \alpha_{tu}^{00} + r_{\alpha t} \alpha_{t\beta u}^{01} + r_{\beta u} \alpha_{\alpha t u}^{10} + \alpha_{\alpha t \beta u}^{11} \right] \quad (79)$$

$$\beta_{\alpha\beta\gamma} = \sum_{tuv} \left[r_{\alpha t} r_{\beta u} r_{\gamma v} \beta_{tuv}^{000} + r_{\alpha t} r_{\beta u} \beta_{tu\gamma v}^{001} + r_{\alpha t} r_{\gamma v} \beta_{t\beta uv}^{010} + r_{\beta u} r_{\gamma v} \beta_{\alpha t uv}^{100} \right. \\ \left. + r_{\alpha t} \beta_{t\beta u\gamma v}^{011} + r_{\gamma v} \beta_{\alpha t \beta uv}^{110} + r_{\beta u} \beta_{\alpha t u\gamma v}^{101} + \beta_{\alpha t \gamma u\gamma v}^{111} \right] \quad (80)$$

where $r_{\alpha t}$ is the α component of the vector from the molecular origin to the atom t .

Distributed (hyper)polarizabilities have been used to compute the linear and first nonlinear susceptibility of the urea crystal (Reis et al. 2000), but not yet of liquids. For the use of distributed electric properties in dimers and other small clusters see, e.g., Misquitta (2015).

The Local Properties (LoProp) Approach

The LoProp approach was designed to yield a computationally simple method for the partitioning of multipolar moments and polarizabilities while still avoiding the problem of basis set sensitivity which plagues cheaper methods such as Mulliken

or Löwdin partitioning schemes (Gagliardi et al. 2004). It works essentially by first separating the atomic orbitals (the basis set) into subsets of occupied and unoccupied orbitals. This means that core and valence orbitals of the atoms in the system should be well described by the basis set applied, which is usually the case, if, e.g., atomic natural orbital basis sets (ANOs) are used. Other basis sets may need reconstructions to meet this requirement; see Gagliardi et al. (2004) for more details.

In the second step, the overlap matrix is diagonalized in a series of four transformations, while still maintaining the division into occupied and unoccupied subspaces. With the total transformation matrix determined, it is then possible to *localize* multipole moments to atomic centers and bonds between two centers. The molecular dipole moment $\underline{\mu}$ with respect to the point \underline{r}_0 can be expressed by

$$\underline{\mu}_{\underline{r}_0} = \sum_{ij} \underline{\mu}^{(ij)} + \sum_i q^{(i)}(\underline{r}^{(i)} - \underline{r}_0) \quad (81)$$

where i, j are the atomic centers localized at $\underline{r}_i, \underline{r}_j$, $q^{(i)}$ are the localized charges, and $\underline{\mu}^{(ii)}$ and $\underline{\mu}^{(ij)}$ are localized one-center and bond dipole moments, respectively. Due to construction, no bond charges $q^{(ij)}$ occur.

The computation of localized polarizabilities proceeds via the definition of the polarizability as change of the dipole moment in an applied field:

$$\alpha_{\alpha\beta} = \frac{\partial \mu_\alpha}{\partial F_\beta} = \frac{\partial \mu_\beta}{\partial F_\alpha} \quad (82)$$

In addition, *charge transfer* quantities Q^{ij} between two centers due to the applied field are introduced. The definition of Q^{ij} is not unique, and to determine them, they are *required* to be as small as possible as well as short ranged. Using a penalty function to satisfy these requirements, a Lagrangian is set up, which can easily be solved to yield the $Q^{(ij)}$. Finally, the localized polarizabilities are computed from a numerical finite field difference equation:

$$\alpha_{\alpha\beta}^{(ij)} = \frac{\mu_\alpha^{(ij)}(F_\beta) - \mu_\alpha^{(ij)}(-F_\beta)}{F_\beta} + \frac{[Q^{(ij)}(F_\beta) - Q^{(ij)}(-F_\beta)][r_\alpha^{(i)} - r_\alpha^{(j)}]}{2F_\beta} \quad (83)$$

where \underline{F} is the applied field and $P(\pm F_\alpha)$ is the property P in the field $\pm F_\alpha$. These localized polarizabilities are origin independent.

The LoProp scheme has been implemented in the MOLCAS 6.0 software (Molcas 6.0 2003). Unfortunately, it has not been extended to hyperpolarizabilities yet. Finally it is worth noting that LoProp polarizabilities are *localized* properties, not fully *distributed* ones in the sense used in section Distributed Multipole and Polarizability Analysis. One should also mention that LoProp polarizabilities have been shown to be transferable in a couple of series of homologues (Gagliardi et al.

2004), although more extensive tests are required to establish the validity and limits of this property more accurately.

Two Case Studies: Liquid Water and *p*-Nitroaniline in 1,4-Dioxane

In this final section we compare the results of several of the described methods for the linear and nonlinear susceptibilities of two specific systems, namely, liquid water (i.e., a “solution” of water in water) and *p*-nitroaniline in 1,4-dioxane. Especially the latter case is a warning example for many points which one needs to consider when theoretical NLO results are compared with experimental values. Both are systems with strong and specific intermolecular interactions, which are a challenge for a reliable treatment and can thus serve as a probe for the strengths and weaknesses of the specific method used. We emphasize, however, that the results collected here have not been computed under strictly comparable conditions, i.e., using the same basis sets, level of calculation, box sizes, etc. This will certainly lead to some differences for the results and should be kept in mind in the comparisons below.

Water

Liquid water is a difficult system for computational modeling in general and for the prediction of electric properties in particular. The presence of an extended and dynamic network of strong hydrogen bonds represents a serious difficulty for dielectric continuum models as well as for explicit models using molecular mechanics force fields for structure predictions. In particular, the reliable prediction of the molecular first hyperpolarizability in liquid water is a challenge for any model. Experimentally, it was found that $\beta(-2\omega, \omega, \omega)$ undergoes a change of $\sim 200\%$ from the gas phase to the liquid, involving a change of sign (Levine and Bethea 1976; Kaatz et al. 1998). Consequently, nearly all models tried their hand on this interesting system: the multipole expansion (MPE) method (Mikkelsen et al. 1995), DRF (Jensen et al. 2005), PE-CCSD (Osted et al. 2006), PE-DFT (Nielsen et al. 2007), and RLF (Reis et al. 2001). A selection of predicted values for some electric properties of isolated and solvated water is shown in Table 3, where we also include static values obtained with PCM using the CCSD/d-aug-cc-pVTZ level computed for this work.

Mikkelsen et al. (1995) compared results for $\beta(-\omega, \omega, 0)$ of the isolated molecule with those obtained for a cluster with five molecules (SM), the MPE model, and the combined SM+MPE model. They found that at the RHF level the MPE model does not reproduce the change of sign, the SM approach qualitatively reveals this change, and finally their combination yields a value at an order of magnitude comparable with experiment. Although the computed property $\beta(-\omega, \omega, 0)$ corresponds to the dc-Pockels effect, Ågren et al. (2004) subsequently confirmed their conclusions also for $\beta(-2\omega, \omega, \omega)$. Later work, using larger basis sets, more

Table 3 Experimental and ab initio values for selected properties of free and solvated water

Method	μ^{gas}	μ^{sol}	α^{sol}	β_z^{gas}	β_z^{sol}	Ref.
RPA/MPE				-4.4	-0.7	Ågren 2004
SM				-4.4	5.3	Ibid.
SM+MPE				-4.4	13.5	Ibid.
CCSD/PCM	0.73	0.85	10.81	-15.0	-6.4	This work
CCSD/MPE				-19.3	4.5	Kongsted et al. (2003)
CCSD/PE		1.05	10.31	-19.3	12.3	Osted et al. (2006)
DFT/PE		1.06	10.65	-	11.7	Nielsen et al. (2007)
DRF/DFT	0.71	1.01	9.95	-20.42	7.20 ^a	Jensen et al. (2005)
RLF/MP2	0.73	0.97	9.82	-18.6	25.6	Reis et al. (2001)
Experiment	0.73		10.45	-19.2/ - 22.0	19.2	See text

^aEffective property

advanced methods, and refined solvation approaches, e.g., using nonequilibrium solvation, did not change the main finding that the single-molecule continuum approach does not yield a change of sign for β in liquid water. However, Kongsted et al. (2003) showed that the MPE model, using CC2 and CCSD in combination with large diffuse basis sets, is able to predict the sign change for water, which they rationalize in terms of large charge penetration effects. In contrast, our values show that PCM is not able to reproduce the large changes of β upon solvation, even at the CCSD/d-aug-cc-pVTZ level.

Kongsted et al. (2003) also applied a QM/MM method, using an averaging procedure over the simulated trajectory to produce a mean configuration on which the QM calculation was performed. In this way, they were also able to predict the sign change. In addition, they showed that a qualitatively correct result can only be obtained if the MM part is *polarizable*. In a continuation of this work, Osted et al. (2006) applied the PE model at the CCSD/MM level and analyzed the results statistically. Although great effort was expended to obtain converged procedures, the agreement for the EFISH value with experiment was again only qualitative. Later work using the PE method with DFT instead of CCSD were in good agreement with the PE/CCSD/MM values (Nielsen et al. 2007). An application of the explicit DRF model to liquid water was reported in Jensen et al. (2005), where the sign change could also be predicted, although the final value was again only in qualitative agreement with experiment. Similarly, the RLF model is able to reproduce the large changes of β upon solvation, without, however, better agreement with experiment (Reis et al. 2001).

To summarize, although in general the explicit models yield better predictions compared with experiment than the continuum models, none of the methods is able to predict the experimental EFISH value quantitatively. One possible explanation is that all methods used the equation obtained by thermal averaging over the dipole orientations, i.e., Eq. 65, instead of modeling the orientational effect directly (Jensen

et al. 2005). Another possible explanation, of course, is that the single experimental value available may not be accurate.

It should be mentioned finally that nearly all methods are able to predict the linear optical effects, i.e., the refractive index $n(\omega)$ in good agreement with experiment.

EFISH of *p*-Nitroaniline in 1,4-Dioxane

Para-nitroaniline is a typical donor–acceptor compound with a quite large first hyperpolarizability and has been studied extensively both experimentally and theoretically. Many experimental studies were performed in the solvent 1,4-dioxane, which has long been known to show unusual dielectric properties: although the relative permittivity is with $\epsilon = 2.209$ that of a typical nonpolar solvent, it behaves like a medium polar solvent, e.g., in solvatochromic shift measurements. Several studies have shown that this is probably due to a large quadrupolar moment of the molecule (Geerlings et al. 2000; Cinacchi et al. 2006; Reis et al. 2006). Thus, this solvent could be a challenge for continuum models, which rely only on dipolar properties of the solvent, i.e., the relative permittivity ϵ and/or the refractive indices $n(\omega)$. Nevertheless, in several applications of continuum models on the EFISH signal of pNA in 1,4-dioxane, it was claimed that good agreement with experimental values has been found. As shown in Reis (2006), this is probably erroneous in most cases due to a mixture of several factors, among which we mention:

1. The correct conversion factor for many of the experimental values of β to the convention used here is 4. In a landmark paper (Willets et al. 1992), which was often used for the conversion of experimental to theoretical values, this factor was incorrectly given as 6. Several authors considered this factor correctly as too large and applied instead the (incorrect) factor of two (Jensen and van Duijnen 2005).
2. The experimental calibration factor for EFISH is based on the SHG value of quartz and has changed considerably over the years. The currently accepted value is about 60 % of the previously used value. Accordingly, this leads to some confusion in the choice of values with which to compare the theoretical results (Mikkelsen et al. 1994).
3. An experimental calibration factor independent of quartz was determined by Shelton, using the Hyper-Rayleigh scattering (HRS) signal of liquid CCl_4 . However, this calibration factor was later revised to a value which is a factor of 1.88 larger (Pyatt and Shelton 2001). The first factor applied to the experimental HRS signal of pNA/1,4-dioxane yields a value close to that of the EFISH-signal, if for the latter the (incorrect) conversion factor of 2 is applied (Jensen and van Duijnen 2005). The revised factor brings the HRS signal in good agreement with the EFISH values, using the correct factor of 4 (Reis 2006).

In Table 4, we compare some theoretical values reported in the literature with updated experimental values from different sources. As expected, the continuum

Table 4 Experimental and ab initio values for μ and $\beta_z(-2\omega; \omega, \omega)$ of free and solvated pNA

Method	μ^{gas}	μ^{sol}	β_z^{gas}	β_z^{sol}	Ref.
Experiment	(2.70) ^a	2.44–2.81	1787	4530–4694	
SCR/MCSCF	2.50	2.73	1527 ^b	2305 ^a	Mikkelsen et al. (1994)
SCR/MCSCF	–	–	1860	2670	Norman et al. (1997)
PCM/DFT	2.12	2.60	2743	3356	Benassi et al. (2015)
DRF/DFT	3.04	4.18	3545	7710, 4618 ^c	Jensen and van Duijnen (2005)
RLF/MP2	2.42	2.54	2092, 2542	4970–5130	Reis et al. (2006)

^a Ab initio value assuming a C_{2v} symmetry

^b $\beta_z(-\omega; \omega, 0)$

^c β_z^{eff} see Eq. 59

models in general yield values which are too low. Nevertheless, the depolarization ratio of certain components of β , which can be obtained from Hyper-Rayleigh scattering experiments and are independent of the absolute values, could be predicted quite well by the MPE approach (Norman et al. 1997). In addition, DFT does not appear to be a good choice for pNA at all, even if long-range corrected functionals like CAM-B3LYP are used, as not even the gas phase is predicted reasonable well (Benassi et al. 2015). A similar failure of CAM-B3LYP has also been found in a related study using the PE approach on the excited states of pNA dissolved in water (Eriksen et al. 2013).

The two explicit models DRF and RLF use instead a molecular charge distribution for 1,4-dioxane, either using point charges (DRF) or quadrupole and higher multipole moments (RLF). As a consequence, they are in much better agreement with experiment (note that for DRF, the effective hyperpolarizability should be compared with the experimental β_z^{sol}). Nevertheless, the gas-phase value for β_z used for DRF is much too high, which probably again is a DFT/TDDFT issue, and may throw some doubt on the good value obtained for the solution. The RLF approach yields values for β_z , which are in reasonable agreement both for the gas and solvated phases, although the quantum-mechanical approach used in that work (MP2 scaled with RPA or CC2 dispersion) is already a bit outdated. Unfortunately, no values using any of the more modern approaches PE or FDE are available.

Acknowledgements M.M., S.B., and H.R. acknowledge financial support by the project Mobilities – enhancing research, science, and education at the Matej Bel University (ITMS code: 26110230082, under the Operational Program Education financed by the European Social Fund).

Bibliography

Aidas, K., et al. (2014). The Dalton quantum chemistry program system. *WIREs Computational Molecular Science*, 4, 269–284 (Dalton, a Molecular Electronic Structure Program, Release DALTON2015.0, 2015). See <http://daltonprogram.org/>.

- Amovilli, C., & Mennucci, B. (1997). Self-consistent-field calculation of Pauli repulsion and dispersion contributions to the solvation free energy in the polarizable continuum model. *The Journal of Physical Chemistry B*, *101*, 1051–1057.
- Andersson, K., Barysz, M., & Berhardsson, A., et al. (2003). MOLCAS Version 6.0. Department of Theoretical Chemistry, Chemical Center, University of Lund, P.O.B. 124, S-22100 Lund
- Applequist, J., Carl, J. R., & Fung, K. K. (1972). Atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities. *Journal of the American Chemical Society*, *94*, 2952–2960.
- Avramopoulos, A., Papadopoulos, M. G., & Reis, H. (2007). Calculation of the microscopic and macroscopic linear and nonlinear optical properties of liquid acetonitrile. II Local fields and linear and nonlinear susceptibilities in quadrupolar approximation. *The Journal of Physical Chemistry B*, *111*, 2546–2553.
- Baranowska, A., Zawada, A., Fernández, B., Bartkowiak, W., Kędziera, D., & Kaczmarek-Kędziera, A. (2010). Interaction-induced electric properties and cooperative effects in model systems. *Physical Chemistry Chemical Physics*, *12*, 852–862.
- Baranowska, A., Fernández, A. B., & Sadlej, A. J. (2011). Importance of electron correlation effects and basis set superposition error in calculations of interaction energies and interaction-induced electric properties in hydrogen-bonded complexes: A model study. *Theoretical Chemistry Accounts*, *128*, 555–561.
- Benassi, E., Egidi, F., & Barone, V. (2015). General strategy for computing nonlinear optical properties of large neutral and cationic organic chromophores in solution. *The Journal of Physical Chemistry B*, *119*, 3155–3173.
- Boettcher, C. J. F., & Bordewijk, P. (1978). *Theory of electric polarization* (2nd ed., Vol. 1) Amsterdam: Elsevier.
- Bukowski, R., & Szalewicz, K. (2001). Complete ab initio three-body nonadditive potential in Monte Carlo simulations of vapor liquid equilibria and pure phases of argon. *The Journal of Chemical Physics*, *114*, 9518–9531.
- Cammi, R., Mennucci, B., & Tomasi, J. (1998). On the calculation of local field factors for microscopic static hyperpolarizabilities of molecules in solution with the aid of quantum-mechanical methods. *The Journal of Physical Chemistry A*, *102*, 870–875.
- Cammi, R., Mennucci, B., & Tomasi, J. (2000). An attempt to bridge the gap between computation and experiment for nonlinear optical properties: Macroscopic susceptibilities in solution. *The Journal of Physical Chemistry A*, *102*, 870–875.
- Coutinho, K., Canuto, S., & Mukherjee, P. K. (2005). The dipole polarizability of F⁻ in aqueous solution: A sequential quantum mechanics/molecular mechanics study. *Advances in Quantum Chemistry*, *48*, 141–150.
- Casida, M. E., & Wesolowski, T. A. (2004). Generalization of the Kohn Sham equations with constrained electron density formalism and its time-dependent response theory formulation. *International Journal of Quantum Chemistry*, *96*, 577–588.
- Cinacchi, G., Ingrosso, F., & Tani, A. (2006). Solvation dynamics by computer simulation: Coumarin C153 in 1,4-dioxane. *The Journal of Physical Chemistry B*, *110*, 13633–13641.
- Czyżnikowska, Ż., Góra, R. W., Zaleśny, R., Bartkowiak, W., Baranowska-Łączkowska, A., & Leszczynski, J. (2013). The effect of intermolecular interactions on the electric dipole polarizabilities of nucleic acid base complexes. *Chemical Physics Letters*, *555*, 230–234.
- Dyer, P. J., & Cummings, P. T. (2006). Hydrogen bonding and induced dipole moments in water: Predictions from the Gaussian charge polarizable model and Car-Parrinello molecular dynamics. *The Journal of Chemical Physics*, *125*, 144519.
- Eriksen, J. J., Sauer, S. P. A., Mikkelsen, K. V., Jensen, H. J. Aa., & Kongsted, J. (2012). On the importance of excited state dynamic response electron correlation in polarizable embedding methods. *Journal of Computational Chemistry*, *33*, 2012–2022.
- Eriksen, J. E., Sauer, S. P. A., Mikkelsen, K. V., Christiansen, O., Jensen, H. J. Aa., & Kongsted, J. (2013). Failures of TDDFT in describing the lowest intramolecular charge-transfer excitation in *para*-nitroaniline. *Molecular Physics*, *111*, 1235–1248.

- Ferrighi, L., Frediani, L., Cappelli, C., Salek, P., Ågren, H., Helgaker, T., & Ruud, K. (2006). Density-functional-theory study of the electric-field-induced second harmonic generation (EFISHG) of push-pull phenylpolyenes in solution. *Chemical Physics Letters*, *425*, 267–272.
- Fowler, P. W., & Madden, P. A. (1984). In-crystal polarizabilities of alkali and halide ions. *Physical Review B*, *29*, 1035–1042.
- Gagliardi, L., Lindh, R., & Karlström, G. (2004). Local properties of quantum chemical systems: The LoProp approach. *The Journal of Chemical Physics*, *121*, 4494–4500.
- Geerlings, J. D., Varma, C. A. G. O., & van Hemert M. C. (2000). Molecular dynamics studies of a dipole in liquid dioxanes. *The Journal of Physical Chemistry B*, *104*, 56–64.
- Góra, R. W., Zalesny, R., Zawada, A., Bartkowiak, W., Skwara, B., Papadopoulos, M. G., & Silva, D. L. (2011). Large Changes of static electric properties induced by hydrogen bonding: An ab initio study of linear HCN oligomers. *The Journal of Physical Chemistry A*, *115*, 4691–4700.
- Hammond, J. R., Govind, N., Kowalski, K., Autschbach, J., & Xantheas, S. S. (2009). Accurate dipole polarizabilities for water clusters $n=2-12$ at the coupled-cluster level of theory and benchmarking of various density functionals. *The Journal of Chemical Physics*, *131*, 214103.
- Hedegård, E. D., List, N. H., Jensen, H. J. Aa., & Kongsted, J. (2013). The multi-configuration self-consistent field method within a polarizable embedded framework. *The Journal of Chemical Physics*, *139*, 044101.
- Jacob, C. R., Neugebauer, J., Jensen, L., & Visscher, L. (2006). Comparison of frozen-density embedding and discrete reaction field solvent models for molecular properties. *Physical Chemistry Chemical Physics*, *8*, 2349–2359.
- Janssen, R. H. C., Bomont, J.-M., Theodorou, D. N., Raptis, S., & Papadopoulos, M. G. (1999). Computer simulation of the linear and nonlinear optical properties of liquid benzene: Its local fields, refractive index and second nonlinear susceptibility. *The Journal of Chemical Physics*, *110*, 6463–6474.
- Jaszuński, M., Rizzo, A., & Ruud, K. (2015). Molecular electric, magnetic, and optical properties. In J. Leszczynski (Ed.), *Handbook of computational chemistry* (Vol. 1). Springer.
- Jemmer, P., Fowler, P. W., Wilson, M., & Madden, P. A. (1997). Environmental effects on anion polarizability: Variation with lattice parameter and coordination number. *The Journal of Physical Chemistry A*, *102*, 8377–8385.
- Jensen, L., & van Duijnen, P. Th. (2005). The first hyperpolarizability of *p*-nitroaniline in 1,4-dioxane: A quantum mechanical/molecular mechanics study. *The Journal of Chemical Physics*, *123*, 074307.
- Jensen, L., Swart, M., van Duijnen, P. Th., & Snijders, J. G. (2002). Medium perturbations on the molecular polarizability calculated within a localized dipole interaction model. *The Journal of Chemical Physics*, *117*, 3316–3320.
- Jensen, L., Sylvester-Hvid, K. O., Åstrand, P.-O., & Mikkelsen, K. V. (2003). A dipole-interaction model for the molecular second hyperpolarizability. *The Journal of Physical Chemistry A*, *107*, 2270–2276.
- Jensen, L., Swart, M., & van Duijnen, P. Th. (2005). Microscopic and macroscopic polarization within a combined quantum mechanics and molecular mechanics model. *The Journal of Chemical Physics*, *122*, 034103.
- Kaatz, P., Donley, E. A., & Shelton, D. P. (1998). A comparison of molecular hyperpolarizabilities from gas and liquid phase measurements. *The Journal of Chemical Physics*, *108*, 849–856.
- Kanis, D. R., Ratner, M. A., & Marks, D. J. (1994). Design and construction of molecular assemblies with large second-order optical nonlinearities. Quantum chemical aspects. *Chemical Review*, *94*, 195–242.
- Kasprowicz-Kielich, B., & Kielich, S. (1975). A classical treatment of nonlinear processes of molecular relaxation in intense electric fields of high and low frequency. *Advances in Molecular Relaxation Processes*, *7*, 275–305.
- Klamt, A. (2011). The COSMO and COSMO-RS solvation models. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *1*(5), 699–709.

- Klamt, A., Mennucci, B., Tomasi, J., Barone, V., Curutchet, C., Orozco, M., & Luque, F. J. (2009). On the performance of continuum solvation methods. A comment on universal approaches to solvation modeling. *Accounts of Chemical Research*, *42*, 489–492.
- Kongsted, J., Osten, A., & Mikkelsen, K. V. (2003). Nonlinear optical response properties of molecules in condensed phases using the coupled cluster/dielectric continuum or molecular mechanics methods. *The Journal of Chemical Physics*, *119*, 10519–10535.
- Levine, B. F., & Bethea, C. G. (1976). Effects on hyperpolarizabilities of molecular interactions in associating liquid mixtures. *The Journal of Chemical Physics*, *65*, 2429–2438.
- Li, W., Zhou, X., Tian, W. Q., & Sun, X. (2013). A new scheme for significant enhancement of the second order nonlinear optical response from molecules to ordered aggregates *Physical Chemistry Chemical Physics*, *15*, 1810–1814.
- Lipiński, J. (1988). Modified all-valence INDO/spd method for ground and excited state properties of isolated molecules and molecular complexes. *International Journal of Quantum Chemistry*, *34*, 423–435.
- Lipiński, J., & Bartkowiak, W. (1997). Solvent effect on the electronic structure of molecules studied by the Langevin dipoles/Monte Carlo approach. *The Journal of Physical Chemistry A*, *101*, 2159–2165.
- Luo, Y., Ågren, H., & Mikkelsen, K. V. (1997). Unique determination of the cavity radius in Onsager reaction field theory. *Chemical Physics Letters*, *275*, 145–150.
- Marchal, R., Carbonniere, Ph., & Pouchan, C. (2009). A global search algorithm of minima exploration for the investigation of low lying isomers of clusters from density functional theory-based potential energy surfaces: The example of Si-n (n=3, 15) as a test case. *The Journal of Chemical Physics*, *131*, 114105.
- Marenich, A. V., Olson, R. M., Kelly, C. P., Cramer, C. J., & Truhlar, D. G. (2007). Self-consistent reaction field model for aqueous and nonaqueous solutions based on accurate polarized partial charges. *Journal of Chemical Theory and Computation*, *3*, 2011–2033.
- Mata, R. A., Costa Cabral, B. J., Millot, C., Coutinho, K., & Canuto, S. (2009). Dynamic polarizability, Cauchy moments, and the optical absorption spectrum of liquid water: A sequential molecular dynamics/quantum mechanical approach. *The Journal of Chemical Physics*, *130*, 014505.
- Medved', M., Budzák, Š., Laurent, A., & Jacquemin, D. (2015). Direct and indirect effects of dispersion interactions on NLO properties of weakly bound complexes. *Journal of Physical Chemistry A*, *119*, 3112–3124.
- Mennucci, B., Amovilli, C., & Tomasi, J. (1998). On the effect of Pauli repulsion and dispersion on static molecular polarizabilities and hyperpolarizabilities in solution. *Chemical Physics Letters*, *286*, 221–225.
- Mikkelsen, K. V., Luo, Y., Ågren, H., & Jørgensen, P. (1994). Solvent induced polarizabilities and hyperpolarizabilities of *para*-nitroaniline studied by reaction field linear response theory. *The Journal of Chemical Physics*, *100*, 8240–8250.
- Mikkelsen, K. V., Jørgensen, P., & Jensen, H. J. A. (1994). A multiconfiguration self-consistent reaction field response method. *The Journal of Chemical Physics*, *100*, 6597–6607.
- Mikkelsen, K. V., Luo, Y., Ågren, H., & Jørgensen, P. (1995). Sign change of hyperpolarizabilities of solvated water. *The Journal of Chemical Physics*, *102*, 9362–9367.
- Misquitta A. J. (2015). Intermolecular interactions. In J. Leszczynski (Ed.), *Handbook of computational chemistry* (Vol. 1). Springer.
- Misquitta, A. J., & Stone, A. J. (2007). CAMCASP: Cambridge package for Calculation of anisotropic site properties. University of Cambridge. <http://www-stone.ch.cam.ac.uk/programs/camcasp.html>.
- Moraldi, M., Celli, M., & Barocchi, F. (1989). Theory of virial expansion of correlation functions and spectra: Application to interaction-induced spectroscopy. *Physical Review A*, *40*, 1116–1126.
- Morita, A., & Kato, S. (1999). An ab initio analysis of medium perturbation on molecular polarizabilities. *The Journal of Chemical Physics*, *110*, 11987–11998.

- Munn, R. W. (1988). Electric dipole interactions in molecular crystals. *Molecular Physics*, *64*, 1–20.
- Munn, R. W., Luo, Y., Macak, P., & Ågren, H. (2005). Role of the cavity field in nonlinear optical response in the condensed phase. *The Journal of Chemical Physics*, *114*, 3105–3108.
- Neugebauer, J., Louwerse, M. J., Baerends, E. J., & Wesolowski, T. A. (2005). The merits of the frozen-density embedding scheme to model solvatochromic shifts. *The Journal of Chemical Physics*, *122*, 094115.
- Nielsen, C. B., Christiansen, O., Mikkelsen, K. V., & Kongsted, J. (2007) Density functional self-consistent quantum mechanics/molecular mechanics theory for linear and nonlinear molecular properties: Applications to solvated water and formaldehyde. *The Journal of Chemical Physics*, *126*, 154112.
- Norman, P., Luo, Yi, Jonsson, D., Ågren, H., Sylvester-Hvid, K. O., & Mikkelsen K. V. (1997). Hyperpolarizability depolarization ratios of nitroanilines. *The Journal of Chemical Physics*, *107*, 9063–9066.
- Olsen, J. M., & Kongsted, J. (2011). Molecular properties through polarizable embedding. *Advances in Quantum Chemistry*, *61*, 107–143.
- Olsen, J. M., Aidas, K., & Kongsted, J. (2010). Excited states in solution through polarizable embedding. *Journal of Chemical Theory and Computation*, *6*, 3721–3734.
- Osted, A., Kongsted, J., Mikkelsen, K. V., Åstrand, P.-O., & Christiansen, O. (2006). Statistically mechanically averaged coupled cluster/molecular dynamics method. *The Journal of Chemical Physics*, *124*, 124503.
- Pipolo, S., Corni, S., & Cammi, R. (2014). The cavity electromagnetic field within the polarizable continuum model of solvation. *The Journal of Chemical Physics*, *140*, 16.
- Pyatt, R. D., & Shelton, D. P. (2001). Hyper-Rayleigh scattering from CH₄, CD₄, CF₄, and CCl₄. *The Journal of Chemical Physics*, *114*, 9938–9946.
- Reis, H. (2006). Problems in the comparison of theoretical and experimental hyperpolarizabilities revisited. *The Journal of Chemical Physics*, *125*, 014506.
- Reis, H., Papadopoulos, M. G., Hättig, C., Ángyán, J. G., & Munn, R. W. (2000). Distributed first and second order hyperpolarizabilities: An improved calculation of nonlinear optical susceptibilities of molecular crystals. *The Journal of Chemical Physics*, *112*, 6161–6172.
- Reis, H., Papadopoulos, M. G., & Theodorou, D. N. (2001). Calculation of refractive indices and third-harmonic generation susceptibilities of liquid benzene and water: Comparison of continuum and discrete local field theories. *The Journal of Chemical Physics*, *114*, 876–881.
- Reis, H., Papadopoulos, M. G., & Grzybowski, A. (2006). Computer simulation of the linear and nonlinear optical susceptibilities of *p*-nitroaniline in cyclohexane, 1,4-dioxane, and tetrahydrofuran in quadrupolar approximation. II. Local field effects and optical susceptibilities. *The Journal of Physical Chemistry B*, *125*, 18537–18552.
- Rinaldi, D., & Rivail, J.-L. (1973). Polarisabilités moléculaires et effet diélectrique de milieu à l'état liquide. Étude théorique de la molécule d'eau et de ses dimères. *Theoretica Chimica Acta*, *32*, 57–70.
- Serr, A., & Netz, R. R. (2006). Polarizabilities of hydrated and free ions derived from DFT calculations. *International Journal of Quantum Chemistry*, *106*, 2960–2974.
- Skwara, B., Bartkowiak, W., Zawada, A., Góra, R. W., & Leszczynski, J. (2007). *Chemical Physics Letters*, *436*, 116–123.
- Skwara, B., Kaczmarek, A., Góra, R. W., & Bartkowiak, W. (2008). On decomposition of interaction-induced electric properties of HF dimer. *Chemical Physics Letters*, *461*, 203–206.
- Sneskov, K., Schwabe, T., Kongsted, J., & Christiansen, O. (2011). The polarizable embedding coupled cluster method. *The Journal of Chemical Physics*, *134*, 104108.
- Sokalski, W. A., & Poirier, R. A. (1983). Cumulative atomic multipole representation of the molecular charge distribution and its basis set dependence. *Chemical Physics Letters*, *98*, 86–92.
- Sokalski, W. A., Roszak, S., & Pecul, K. (1988). An efficient procedure for decomposition of the SCF interaction energy into components with reduced basis set dependence. *Chemical Physics Letters*, *153*, 153–159.

- Stone, A. J. (2013). *The theory of intermolecular forces* (2nd ed.) Oxford: Clarendon.
- Suponitsky, K. Y., & Masunov, A. E. (2013). Supramolecular step in design of nonlinear optical materials: Effect of $\pi \cdots \pi$ stacking aggregation on hyperpolarizability. *The Journal of Chemical Physics*, *139*, 094310.
- Sylvester-Hvid, K. O., Mikkelsen, K. V., Norman, P., Jonsson, D., & Ågren, H. (2004). Sign change of hyperpolarizabilities of solvated water, revised: Effects of equilibrium and nonequilibrium solvation. *The Journal of Physical Chemistry A*, *108*, 8961–8965.
- Szalewicz, K. (2012). Symmetry-adapted perturbation theory of intermolecular forces. *WIREs Computational Molecular Science*, *2*, 254–272.
- Thole, B. T. (1981). Molecular polarizabilities calculated with a modified dipole interaction. *Chemical Physics*, *59*, 341–350.
- Tomasi, J., Mennucci, B., & Cammi, R. (2005). Quantum mechanical continuum solvation models. *Chemical Review*, *105*, 2999–3094.
- Turbomole V6.2, 2010a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007. Available from <http://www.turbomole.com>.
- Wang, C. K., Wang, Y. H., Su, Y., & Luo, Y. (2003). Solvent dependence of solvatochromic shifts and the first hyperpolarizability of *para*-nitroaniline: A nonmonotonic behavior. *The Journal of Chemical Physics*, *119*, 4409–4412.
- Warshel, A., & Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology*, *103*, 227–249.
- Wesołowski, T. A., & Warshel, A. (1993). Frozen density functional approach for ab initio calculations of solvated molecules. *The Journal of Physical Chemistry*, *97*, 8050–8053.
- Wesołowski, T. A., & Weber, J. (1996). Kohn-Sham equations with constrained electron density: An iterative evaluation of the ground-state electron density of interacting molecules. *Chemical Physics Letters*, *248*, 71–76.
- Willetts, A., Rice, J. E., Burland, D. M., & Shelton, D. P. (1992). Problems in the comparison of theoretical and experimental hyperpolarizabilities. *The Journal of Chemical Physics*, *97*, 7590–7599.
- Wortmann, R., & Bishop, D. M., (1998). Effective polarizabilities and local field corrections for nonlinear optical experiments in condensed media. *The Journal of Chemical Physics*, *108*, 1001.
- Zalesny, R., Bartkowiak, W., Styrz, S., & Leszczynski, J. (2002). Solvent effects on conformationally induced enhancement of the two-photon absorption cross section of a pyridinium-N-phenolate betaine dye. A quantum chemical study. *The Journal of Physical Chemistry A*, *106*, 4032–4037.

Patrizia Calaminici, Aurelio Alvarez-Ibarra, Domingo Cruz-Olvera,
Victor-Daniel Domínguez-Soria, Roberto Flores-Moreno,
Gabriel U. Gamboa, Gerald Geudtner, Annick Goursot,
Daniel Mejía-Rodríguez, Dennis R. Salahub,
Bernardo Zuniga-Gutierrez, and Andreas M. Köster

Contents

Introduction	796
Theory	799
Kohn-Sham Density Functional Theory	799
The LCGTO Kohn-Sham Method	801
Auxiliary Density Functional Theory	806
Auxiliary Density Perturbation Theory	808
ERI Expansion in ADFT	810
Variational Fitting of Exact Exchange in ADFT	814
The ADFT-GIAO Methodology	820
Applications	827

P. Calaminici (✉), A. Alvarez-Ibarra, D. Cruz-Olvera, G.U. Gamboa, G. Geudtner,
D. Mejía-Rodríguez and A.M. Köster (✉)
Departamento de Química, CINVESTAV, Av. Instituto Politécnico Nacional, México, D.F.,
México
e-mail: pcalamin@cinvestav.mx; akoster@cinvestav.mx

V.-D. Domínguez-Soria
Departamento de Ciencias Básicas, UAM-A, Avenida San Pablo, México, D.F., México

R. Flores-Moreno
Departamento de Química, CUCEI, Universidad de Guadalajara, Guadalajara Jalisco, México

A. Goursot
Institut Charles Gerhardt, UMR 5253 CNRS, Ecole de Chimie de Montpellier, Montpellier,
Cédex 5, France

D.R. Salahub
Department of Chemistry, CMS Centre for Molecular Simulation, IQST Institute for Quantum
Science and Technology, University of Calgary, Calgary, AB, Canada

B. Zuniga-Gutierrez
Departamento de Ciencias Computacionales, Universidad de Guadalajara, Guadalajara Jalisco,
México

Dynamics of Sodium Clusters	827
Structure of Zeolites	833
Stability of Giant Fullerenes	842
Transition State Search in Metal Clusters	847
Conclusion	852
Bibliography	853

Abstract

The working equations of auxiliary density functional theory (ADFT) and auxiliary density perturbation theory (ADPT) are derived in the framework of the linear combination of Gaussian-type orbital expansion. The inclusion of hybrid functionals into ADFT is presented. Its extension for the calculation of magnetic properties is outlined. The ADFT and ADPT implementations in the density functional theory program deMon2k are discussed. Special attention is given to the efficient calculation of electron repulsion integrals in nanostructures. The use of ADFT and ADPT in first-principles Born-Oppenheimer molecular dynamics at the pico- to nanosecond time scale is reviewed. In particular, the long-standing mystery of the discrepancy between experiments and computations for the polarizability of small sodium clusters is resolved. Applications of the parallel deMon2k ADFT implementation to systems on the nanometer scale are reviewed. This includes Al-zeolites and giant fullerenes. It is shown that structures as large as C₅₄₀ can be fully optimized within a few days without any symmetry constraints in the ADFT framework employing all-electron basis sets. The successful application of a hierarchical transition state finder for the study of selected sodium cluster rearrangements is presented, too.

Introduction

Within the broad spectrum of computational methods included in this Handbook of Computational Chemistry, the present chapter will focus on auxiliary density functional theory (ADFT) (Köster et al. 2004b) as implemented in the software package deMon, the most recent version of which is known as deMon2k (Köster et al. 2011a; Geudtner et al. 2011).

Density functional theory (DFT) has shown solid progress since its formal beginnings in the theorems of Hohenberg and Kohn (1964) and Kohn and Sham (1965) and its even older roots dating back to the early work of Thomas and Fermi (Thomas 1927; Fermi 1927, 1928a,b), Slater (1951), and others (Gaspar 1954; Schwarz 1972). Our deMon program, too, has a significant track record of steady improvement over the years.

Looking back to the 1960s and early seventies, the most popular DFT method was the X α -scattered wave (SW) method of Slater and Johnson (Johnson 1966; Johnson and Smith 1970). This method, with its unsightly muffin-tin potential, was the trailblazer for DFT in molecular applications. It had good success for the

spectroscopy of inorganic complexes, including those containing transition metals (Johnson and Smith 1971, 1972; Johnson and Messmer 1974; Messmer et al. 1975, 1978; Weber et al. 1977; Salahub 1978). But the lack of a full potential prevented the SW technique from performing even such simple tasks as geometry optimizations which, at the time, were coming under control by the usual wavefunction-based methods of quantum chemistry, at least for simple molecules. Hence $X\alpha$ -SW remained a rather specialized fringe methodology, which was either ignored or scorned by mainstream quantum chemists.

The seventies and early eighties saw work that changed all of that. A few “ab initio quantum chemists” got interested in DFT and slowly but surely brought over some of the methodology and techniques that would allow a broader and broader base of functionality for DFT codes. This was the case for the earliest versions of deMon (St-Amant and Salahub 1990; Casida et al. 1996) as well as for several other codes, such as LCGTO- $X\alpha$ (Dunlap and Rösch 1990), DGAUSS (Andzelm and Wimmer 1992), ADF (Velde et al. 2001), and DMOL (Delley 1990). The deMon program, for example, was based on Gaussian functions, and much of the machinery comes over to the DFT world straightforwardly, although some aspects like numerical integration (Becke 1987) remain special. Now geometry optimizations held no mystery for the DFT codes, and one could start to think about real chemical applications. Then, in the late eighties, the holy grail of chemical accuracy was reached, at least for organic thermochemistry, with the development of hybrid functionals (Becke 1993a, b). This really got the chemists interested in DFT and that interest has persisted until today.

In parallel, a broad range of physical and chemical properties and processes came into range. The nineties saw the development of powerful methodology that included NMR chemical shifts and coupling constants (Malkin et al. 1993a, b, 1994; Valerio et al. 1998), core-electron spectroscopies (Triguero and Pettersson 1998; Triguero et al. 1998), optical properties through the time-dependent-DFT (TD-DFT) (Casida 1995; Jamorski et al. 1996) approach, photoelectron spectroscopy including the ZEKE (zero-electron kinetic energy) technique (Yang et al. 1996; Calaminici et al. 2001, 2003) molecular polarizabilities and hyperpolarizabilities (Sim et al. 1992; Guan et al. 1995; Calaminici et al. 1998, 1999, 2000), and many others. DFT and programs like deMon were now established as components of the quantum chemical tool kit. The first decade of the twenty-first century, like its predecessors, continued to show steady progress. Born-Oppenheimer molecular dynamics (BOMD) (Wei and Salahub 1994, 1997; Wei et al. 2000; Krishnamurty et al. 2003; Calaminici et al. 2007b; Gamboa et al. 2008; Krishnamurty et al. 2008a, b; Vásquez et al. 2009) have now become “routine” (though sometimes costly) for systems with a hundred or so atoms for simulation times of hundreds of picoseconds or even nanoseconds (Goursot et al. 2013). The recent calculations of melting temperatures and latent heats of free metal clusters have probed the quantitative predictive power of all-electron ADFT-BOMD simulations. The excellent agreement with experimental results (Vásquez et al. 2013, 2015) has further strengthened the confidence in these ab initio molecular dynamics simulations. As a result, ADFT-BOMD has also gained increasing popularity in

global optimizations (Köster et al. 2011b; Geudtner et al. 2013; Medel et al. 2014). Geometry optimizations and transition state searches (Campo and Köster 2008) can now be performed for systems containing several hundred atoms (Dominguez-Soria et al. 2007, 2008; Calaminici et al. 2009) as well as transition metal clusters (Cruz-Olvera et al. 2015); the thousand-atom barrier has been broken (Salahub et al. 2005; Dominguez-Soria et al. 2009), thanks mainly to the development of ADFT (auxiliary density functional theory) (Köster et al. 2004b). More recently, ADFT has also been extended to constrained density functional theory (de la Lande et al. 2013). Hybrid, QM/MM, and QM/QM' methodologies (Lev et al. 2010; Bertran et al. 2010) are now being applied with considerable confidence to systems with a level of complexity that would not have been contemplated just a few years ago. The implementation of force fields into deMon2k now permits multi-scale calculations within deMon2k itself (Salahub et al. 2015). These modules have been further extended by the implementation of asymptotic expansions for electrostatic interaction integrals (Alvarez-Ibarra et al. 2012; Alvarez-Ibarra and Köster 2013) and the efficient variational fitting of exact exchange (Mejia-Rodriguez and Köster 2014). So, at the present time, deMon2k is one of the most powerful LCGTO-DFT codes available. In particular, it includes hybrid functionals without the need for four-center electron repulsion integrals (Mejia-Rodriguez et al. 2015a).

In the next section, we will present the underlying theory, including a description of ADFT and auxiliary density perturbation theory (ADPT), a rather new non-iterative alternative to the coupled perturbed Kohn-Sham (CPKS) methodology. These general theory sections are followed by three more specific methodological sections addressing the computationally demanding three-center electron repulsion integral (ERI) calculations in ADFT, the variational fitting of exact exchange in ADFT, and the ADFT-GIAO (GIAO: Gauge Including Atomic Orbitals) method for the calculation of magnetic properties. Then, in the Applications section, we will describe four recent studies that represent the state of the art in “pure DFT” on large and/or difficult molecular and cluster problems, the largest of which are well into the nano- regime. We solve a long-standing mystery of the discrepancy between experiments and computations for the polarizabilities of sodium clusters (up to the nonamer) – finite temperature effects are the key. In the second application, large models for Na^+ and protonated Al-zeolites are constructed, and the acidic sites are analyzed, including important vibrational effects. The structures of giant fullerenes up to C_{540} have been optimized using all-electron basis sets. The results confirm that for all of the large fullerenes, a faceted shape is preferred over a spherical shape. The calculated binding energies are in the diamond range but considerably below the value for graphene. Finally, a hierarchical transition state algorithm is applied to locate transition states for small sodium cluster rearrangements with six to ten atoms. The minimum structures needed as input were obtained from Born-Oppenheimer molecular dynamics simulations. To connect the found transition states with the corresponding minimum structures, the intrinsic reaction coordinates were calculated. This work demonstrates how rearrangement mechanisms that are not intuitive can be studied in metal clusters.

Theory

Kohn-Sham Density Functional Theory

Density functional theory (DFT) is based on the Hohenberg-Kohn theorem (Hohenberg and Kohn 1964). According to this theorem, the ground-state energy, E , of a many-electron system with the external potential $v(\mathbf{r})$ can be expressed by the following density functional (Parr and Yang 1989; Dreizler and Gross 1990):

$$E[\rho] = T[\rho] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + V_{ee}[\rho]. \quad (1)$$

Here $\rho(\mathbf{r})$ is the electron density, $T[\rho]$ is the kinetic energy, and $V_{ee}[\rho]$ collects all electron-electron interaction energies. A major technical problem is the accurate description of the kinetic energy functional $T[\rho]$. In the Kohn-Sham approximation, this problem is avoided by introducing orbitals of a noninteracting reference system. Levy and Perdew (Levy 1979; Levy and Perdew 1985) have shown that these orbitals are delivered by the following minimization procedure:

$$T_{KS}[\rho] = \min_{\Psi \rightarrow \rho} \langle \Psi | \hat{\mathbf{T}} | \Psi \rangle. \quad (2)$$

Here Ψ is a Slater determinant composed of the orthonormal Kohn-Sham orbitals $\psi_i(\mathbf{r})$, and $T_{KS}[\rho]$ is the corresponding Kohn-Sham kinetic energy. This energy can be calculated from the occupied orbitals of the noninteracting system,

$$T_{KS}[\rho] = -\frac{1}{2} \sum_i^{\text{occ}} \langle \psi_i[\rho] | \nabla^2 | \psi_i[\rho] \rangle. \quad (3)$$

These Kohn-Sham orbitals are functionals of the density and can be derived from it as shown by Parr and coworkers (Zhao and Parr 1992, 1993; Zhao et al. 1994). Imposing the constraint that occupied orbitals of the noninteracting system reproduce the true ground-state density,

$$\rho(\mathbf{r}) = \sum_i^{\text{occ}} |\psi_i(\mathbf{r})|^2, \quad (4)$$

to the variation in Eq. (2) by using a local Lagrange multiplier $v_{KS}(\mathbf{r})$, the following set of equations is obtained:

$$\left(-\frac{1}{2} \nabla^2 + v_{KS}(\mathbf{r}) \right) \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}). \quad (5)$$

These are the Kohn-Sham equations and $v_{KS}(\mathbf{r})$ determines (within a trivial constant) the external potential of the Kohn-Sham reference system.

In order to find a more explicit representation of $v_{KS}(\mathbf{r})$, we now rewrite the energy functional Eq. (1) using the Kohn-Sham kinetic energy term Eq. (3) and the explicit expression for the classic electronic (Coulomb) interaction energy,

$$J[\rho] = \frac{1}{2} \iint \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \quad (6)$$

and it then follows

$$E[\rho] = T_{KS}[\rho] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + J[\rho] + E_{xc}[\rho]. \quad (7)$$

Here the newly introduced exchange-correlation energy functional is defined as

$$E_{xc}[\rho] \equiv T[\rho] - T_{KS}[\rho] + V_{ee}[\rho] - J[\rho]. \quad (8)$$

This quantity collects all nonclassical interactions between the electrons and the difference of the kinetic energies of the interacting and noninteracting system. The accuracy of the Kohn-Sham method is mainly determined by the quality of the approximation used for the calculation of $E_{xc}[\rho]$. The kinetic energy difference appearing in $E_{xc}[\rho]$ is on the order of magnitude of the correlation energy (Almbladh and Pedrosa 1984).

Based on the Hohenberg-Kohn theorem, the ground-state density minimizes the energy functional Eq. (7) and hence satisfies the Euler equation:

$$\mu = \frac{\delta E[\rho]}{\delta \rho(\mathbf{r})}, \quad (9)$$

where μ is the Lagrange multiplier associated with the normalization of the electronic density to the number of electrons n in the system:

$$\int \rho(\mathbf{r}) d\mathbf{r} = n. \quad (10)$$

The functional derivative of the energy functional Eq. (7) yields

$$\begin{aligned} \frac{\delta E[\rho]}{\delta \rho(\mathbf{r})} &= \frac{\delta T_{KS}[\rho]}{\delta \rho(\mathbf{r})} + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \\ &= \frac{\delta T_{KS}[\rho]}{\delta \rho(\mathbf{r})} + v_{KS}(\mathbf{r}). \end{aligned} \quad (11)$$

Thus, the Kohn-Sham potential has the following explicit form:

$$v_{KS}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}[\rho; \mathbf{r}], \quad (12)$$

where the exchange-correlation potential is defined as

$$v_{xc}[\rho; \mathbf{r}] \equiv \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}. \quad (13)$$

Inserting Eq. (12) in Eq. (5) yields the canonical Kohn-Sham orbital equations:

$$\left(-\frac{1}{2}\nabla^2 + v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}[\rho; \mathbf{r}] \right) \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}). \quad (14)$$

These equations have to be solved iteratively to reach self-consistency. They can be cast in matrix form yielding Roothaan-Hall-like equations (Roothaan 1951; Hall 1951).

The LCGTO Kohn-Sham Method

In the linear combination of Gaussian-type orbital (LCGTO) ansatz, the Kohn-Sham orbitals are expanded in atomic orbitals:

$$\psi_i(\mathbf{r}) = \sum_{\mu} c_{\mu i} \mu(\mathbf{r}). \quad (15)$$

In our notation, $\mu(\mathbf{r})$ represents an atomic orbital and $c_{\mu i}$ the corresponding molecular orbital coefficient. To avoid unnecessary complications in the presentation, we restrict ourselves to the closed-shell case. The extension to the open-shell formalism (Pople and Nesbet 1954; Roothaan 1960; Binkley et al. 1974) is straightforward. With the LCGTO expansion, we find for the electron density Eq. (4):

$$\rho(\mathbf{r}) = \sum_{\mu, \nu} P_{\mu\nu} \mu(\mathbf{r}) \nu(\mathbf{r}). \quad (16)$$

$P_{\mu\nu}$ represents an element of the closed-shell density matrix defined as

$$P_{\mu\nu} \equiv 2 \sum_i^{\text{occ}} c_{\mu i} c_{\nu i}. \quad (17)$$

Using the above expansions for the Kohn-Sham orbitals Eq. (15) and the density Eq. (16), the Kohn-Sham energy expression Eq. (7) can be rewritten in terms of atomic orbitals,

$$E = \sum_{\mu, \nu} P_{\mu\nu} H_{\mu\nu} + \frac{1}{2} \sum_{\mu, \nu} \sum_{\sigma, \tau} P_{\mu\nu} P_{\sigma\tau} \langle \mu\nu || \sigma\tau \rangle + E_{xc}[\rho], \quad (18)$$

with

$$H_{\mu\nu} = \langle \mu | -\frac{1}{2}\nabla^2 | \nu \rangle - \sum_C^{\text{Atoms}} \langle \mu | \frac{Z_C}{|\mathbf{r} - \mathbf{C}|} | \nu \rangle \quad (19)$$

and

$$\langle \mu\nu \| \sigma\tau \rangle \equiv \iint \frac{\mu(\mathbf{r})\nu(\mathbf{r})\sigma(\mathbf{r}')\tau(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \quad (20)$$

The SCF convergence is based on this energy expression. The $H_{\mu\nu}$ in Eq. (19) denotes an element of the one-electron Hamiltonian matrix. They are built from the kinetic and nuclear-attraction energy of the electrons and describe the distribution of an electron in the nuclear framework. The computation of this matrix has a formal quadratic scaling with the number of basis functions N of the system. The second term in Eq. (18) represents the Coulomb repulsion energy of the electrons. In contrast to Hartree-Fock theory, the calculations of the Coulomb and exchange energies are separated in DFT. The Coulomb term introduces a formal N^4 scaling. The exchange-correlation energy can be obtained by means of an integral of the following type:

$$E_{xc}[\rho] = \int \rho(\mathbf{r}) \epsilon_{xc}[\rho] d\mathbf{r}. \quad (21)$$

In general, the explicit form of the exchange-correlation energy density, $\epsilon_{xc}[\rho]$, will not permit an analytic solution of this integral. Therefore, a three-dimensional numerical integration has to be performed. It scales formally as $N^2 \times G$, where G denotes the number of grid points in the numerical integration.

From the above discussion, it follows that the calculation of the Coulomb repulsion energy represents the most demanding computational task in Eq. (18). The introduction of the variational approximation of the Coulomb potential (Dunlap et al. 1979; Mintmire and Dunlap 1982; Mintmire et al. 1982) reduces the formal scaling of this term to $N^2 \times M$, where M is the number of auxiliary functions which is usually three to five times N . This technique is nowadays used in most LCGTO-DFT programs. It is identical to the so-called resolution of the identity (RI) (Vahtras et al. 1993; Hamel et al. 2001; Flores-Moreno and Ortiz 2009) that cropped up in wave-function methods, too. The variational approximation of the Coulomb potential, as implemented in deMon2k, is based on the minimization of the following self-interaction term:

$$\mathcal{E}_2 = \frac{1}{2} \iint \frac{[\rho(\mathbf{r}) - \tilde{\rho}(\mathbf{r})][\rho(\mathbf{r}') - \tilde{\rho}(\mathbf{r}')]}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \quad (22)$$

The approximate density $\tilde{\rho}(\mathbf{r})$ is expanded in primitive Hermite Gaussians (Köster 2003) which are centered at the atoms:

$$\tilde{\rho}(\mathbf{r}) = \sum_{\bar{k}} x_{\bar{k}} \bar{k}(\mathbf{r}). \quad (23)$$

The primitive Hermite Gaussian auxiliary functions are indicated by a bar. An (unnormalized) auxiliary function $\bar{k}(\mathbf{r})$ centered at atom C with exponent ζ_k has the form

$$\bar{k}(\mathbf{r}) = \left(\frac{\partial}{\partial C_x} \right)^{\bar{k}_x} \left(\frac{\partial}{\partial C_y} \right)^{\bar{k}_y} \left(\frac{\partial}{\partial C_z} \right)^{\bar{k}_z} e^{-\zeta_k(\mathbf{r}-\mathbf{C})^2}. \quad (24)$$

In deMon2k, these auxiliary functions are normalized with respect to the Coulomb norm and grouped into *s*, *spd*, and *spdfg* sets. The exponents are shared within each of these sets (Andzelm et al. 1985, 1987). Based on the exponent range of the primary atomic basis set, an automatic generation of auxiliary function sets, indicated by the abbreviation GEN, is available in deMon2k. A detailed description is given in Calaminici et al. (2007a) for all-electron calculations and in Calaminici et al. (2005) for effective and model core potential calculations.

With the LCGTO expansion for $\rho(\mathbf{r})$ and $\tilde{\rho}(\mathbf{r})$, we obtain the following representation for \mathcal{E}_2 :

$$\begin{aligned} \mathcal{E}_2 = & \frac{1}{2} \sum_{\mu, \nu} \sum_{\sigma, \tau} P_{\mu\nu} P_{\sigma\tau} \langle \mu\nu \| \sigma\tau \rangle \\ & - \sum_{\mu, \nu} \sum_{\bar{k}} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} + \frac{1}{2} \sum_{\bar{k}, \bar{l}} x_{\bar{k}} x_{\bar{l}} \langle \bar{k} \| \bar{l} \rangle. \end{aligned} \quad (25)$$

The two- and three-center electron repulsion integrals are defined as

$$\langle \bar{k} \| \bar{l} \rangle \equiv \iint \frac{\bar{k}(\mathbf{r}) \bar{l}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \quad (26)$$

$$\langle \mu\nu \| \bar{k} \rangle \equiv \iint \frac{\mu(\mathbf{r}) \nu(\mathbf{r}) \bar{k}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'. \quad (27)$$

In deMon2k, optimized integral recurrence relations for the calculation of the three-center electron repulsion integrals (ERIs) in Eq. (27) are implemented (Köster 1996). The expansion coefficients $x_{\bar{k}}$ of the approximate density are determined by the minimization of \mathcal{E}_2 :

$$\frac{\partial \mathcal{E}_2}{\partial x_{\bar{m}}} = - \sum_{\mu, \nu} P_{\mu\nu} \langle \mu\nu \| \bar{m} \rangle + \sum_{\bar{k}} x_{\bar{k}} \langle \bar{k} \| \bar{m} \rangle = 0 \quad \forall \bar{m}. \quad (28)$$

At this point, it is useful to introduce the Coulomb matrix,

$$\mathbf{G} = \begin{pmatrix} \langle \bar{1} \| \bar{1} \rangle & \langle \bar{1} \| \bar{2} \rangle & \dots & \langle \bar{1} \| \bar{m} \rangle \\ \langle \bar{2} \| \bar{1} \rangle & \langle \bar{2} \| \bar{2} \rangle & \dots & \langle \bar{2} \| \bar{m} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \bar{m} \| \bar{1} \rangle & \langle \bar{m} \| \bar{2} \rangle & \dots & \langle \bar{m} \| \bar{m} \rangle \end{pmatrix}, \quad (29)$$

and the Coulomb vector,

$$\mathbf{J} = \begin{pmatrix} \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{1} \rangle \\ \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{2} \rangle \\ \vdots \\ \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{m} \rangle \end{pmatrix}. \quad (30)$$

With \mathbf{G} and \mathbf{J} , the following system of inhomogeneous equations for the determination of the fitting coefficients, collected in \mathbf{x} , can be formulated:

$$\mathbf{G} \mathbf{x} = \mathbf{J}. \quad (31)$$

A straightforward solution is obtained by the inversion of the Coulomb matrix \mathbf{G} :

$$\mathbf{x} = \mathbf{G}^{-1} \mathbf{J}. \quad (32)$$

Because \mathbf{G} is symmetric and positive definite, in principle, its inversion can be very efficiently performed via Cholesky decomposition (Press et al. 1992). However, if very large auxiliary function sets are used, the Coulomb matrix tends to become ill conditioned. As a consequence, the Cholesky decomposition might fail. For this reason, in deMon2k, we perform by default a singular value decomposition (SVD) of \mathbf{G} at the initial SCF and use in the following energy calculations (e.g., in a geometry optimization or molecular dynamic run) a numerical solver (Dominguez-Soria et al. 2009) that acts only on the nonredundant space of the SVD. This has proven to be a good compromise between accuracy and efficiency. The approach is suitable for systems with thousands of atoms and tight SCF convergence criteria (10^{-9} a.u.). The SVD threshold can be altered by the keyword MATINV (Köster et al. 2011c). Its default value is 10^{-6} .

After the description of the calculation of the fitting coefficients, we now turn to the energy and SCF calculation. Because \mathcal{E}_2 is positive definite, the following inequality holds:

$$\frac{1}{2} \sum_{\mu,\nu} \sum_{\sigma,\tau} P_{\mu\nu} P_{\sigma\tau} \langle \mu\nu \| \sigma\tau \rangle \geq \sum_{\bar{k}} \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} x_{\bar{l}} \langle \bar{k} \| \bar{l} \rangle.$$

With this inequality, an approximate SCF energy, which is based on Eq. (18), can be derived:

$$E = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\bar{k}} \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} x_{\bar{l}} \langle \bar{k} \| \bar{l} \rangle + E_{xc}[\rho]. \quad (33)$$

The variation of this energy expression with respect to the molecular orbital coefficients, constraining the Kohn-Sham orbitals to be orthonormal,

$$\sum_{\mu, \nu} c_{\mu i} S_{\mu \nu} c_{\nu j} = \delta_{ij} \quad \forall i, j, \quad (34)$$

yields

$$\begin{aligned} \frac{\partial E}{\partial c_{\mu i}} = & 4 \sum_{\nu} \left(H_{\mu \nu} + \sum_{\bar{k}} \langle \mu \nu \parallel \bar{k} \rangle x_{\bar{k}} + \langle \mu \mid v_{xc}[\rho] \mid \nu \rangle \right) c_{\nu i} \\ & - 4 \sum_{\nu} \sum_j S_{\mu \nu} c_{\nu j} \epsilon_{ji} \quad \forall \mu, i. \end{aligned} \quad (35)$$

Here, the derivative of the exchange-correlation energy, restricting ourselves to local functionals for clarity, was developed as (Gel'fand and Fomin 1963):

$$\frac{\partial E_{xc}[\rho]}{\partial c_{\mu i}} = \int \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \frac{\partial \rho(\mathbf{r})}{\partial c_{\mu i}} d\mathbf{r} = 4 \sum_{\nu} \langle \mu \mid v_{xc}[\rho] \mid \nu \rangle c_{\nu i}. \quad (36)$$

The ϵ_{ji} in Eq.(35) are the undetermined Lagrange multipliers. Because the electronic density is invariant to unitary transformations of the occupied molecular orbitals (MOs), it is possible (and convenient) to choose a set of MOs for which the off-diagonal multipliers are zero. These MOs are called canonical and are the solutions of the canonical Kohn-Sham equations,

$$\mathbf{K} \mathbf{c} = \mathbf{S} \mathbf{c} \boldsymbol{\epsilon}, \quad (37)$$

with the elements of the Kohn-Sham matrix \mathbf{K} defined as

$$K_{\mu \nu} \equiv H_{\mu \nu} + \sum_{\bar{k}} \langle \mu \nu \parallel \bar{k} \rangle x_{\bar{k}} + \langle \mu \mid v_{xc}[\rho] \mid \nu \rangle. \quad (38)$$

In Eq. (37), \mathbf{S} represents the overlap matrix, \mathbf{c} the molecular orbital coefficient matrix, and $\boldsymbol{\epsilon}$ the diagonal matrix of the Lagrange multipliers, i.e., the Kohn-Sham orbital energies.

As can be seen from Eq. (38), the Kohn-Sham matrix elements depend on the fitting coefficients $x_{\bar{k}}$ and the molecular orbital coefficients $c_{\mu i}$ via the dependence of the exchange-correlation potential on the orbital density $\rho(\mathbf{r})$. In deMon2k, this theoretical model is selected by the keyword specification VXCTYPE BASIS (Köster et al. 2011c) and is often referred to as the BASIS approach. It should be noted that the SCF convergence can be guided either by the molecular orbital coefficients or by the fitting coefficients that arise from the variational fitting of the Coulomb potential. In any case, the variational fitting of the Coulomb potential turns the original energy minimization into a MinMax variation (Köster et al. 2002). In deMon2k, the SCF convergence is guided by the fitting coefficients because they form a vector rather than a matrix. As a result, memory-efficient SCF convergence

acceleration methods suitable for very large systems are available (Köster et al. 2009).

Due to the use of efficient three-center electron repulsion integral algorithms (Köster 1996, 2003; Alvarez-Ibarra and Köster 2013), the numerical integration of the exchange-correlation energy and potential is the computationally most demanding task in this approach. Therefore, a more efficient approach for the calculation of the exchange-correlation energy and potential is desirable. In fact, the use of auxiliary functions for the calculation of the exchange-correlation energy and potential has a long history in DFT methods (Baerends et al. 1973; Sambe and Felton 1975). In programs like deMon-KS (Casida et al. 1996), DGAUSS (Andzelm and Wimmer 1992) or GTOFF (Trickey et al. 2004), the exchange-correlation potential is expanded in auxiliary functions as proposed by Sambe and Felton (1975). The expansion coefficients are obtained by a least squares fit on a grid. Because this fit and the corresponding energy expression are not variational, only approximate gradients and higher-energy derivatives are available (Verluis and Ziegler 1988; Gusarov et al. 2006).

Auxiliary Density Functional Theory

As an alternative to the fitting of the exchange-correlation potential by auxiliary functions, the direct use of the auxiliary function density from the variational fitting of the Coulomb potential for the calculation of the exchange-correlation potential has been investigated over the last years (Laikov 1997; Köster 1998; Köster et al. 2004b; Birkenheuer et al. 2005; Belpassi et al. 2006; Janetzko et al. 2008). The resulting energy expression, from now on named the auxiliary density functional theory (ADFT) energy, is variational and has the form (Köster et al. 2004b)

$$E = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\bar{k}} \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu \| \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k},\bar{l}} x_{\bar{k}} x_{\bar{l}} \langle \bar{k} \| \bar{l} \rangle + E_{xc}[\bar{\rho}]. \quad (39)$$

In deMon2k, this theoretical model is selected by the keyword specification VXCTYPE AUXIS (Köster et al. 2011c); that is the default setting and often referred to as the AUXIS approach. Because the approximate density is a linear combination of auxiliary functions, the density calculation at each grid point scales linearly. In contrast, with the orbital density, products of basis functions have to be evaluated. Obviously, this represents a considerable simplification of the grid work. In particular, using auxiliary function sets that share the same exponents reduces significantly the number of expensive exponential function evaluations at each grid point. Moreover, by using Hermite Gaussian auxiliary functions, the Hermite polynomial recurrence relations (Saunders 1983) can be used for the function calculations on the grid. For the numerical integration of the exchange-correlation energy and potential adaptive and fixed grids (Krack and Köster 1998; Köster et al.

2004a) of various qualities are available in deMon2k. The default setting guarantees energy accuracies in the range of 10^{-5} a.u. for most systems. Higher accuracies can be selected with the GRID keyword (Köster et al. 2011c).

The variation of the ADFT energy expression Eq. (39) with respect to the molecular orbital coefficients, again constraining the Kohn-Sham orbitals to orthonormality Eq. (34), yields

$$\begin{aligned} \frac{\partial E}{\partial c_{\mu i}} &= 4 \sum_v \left(H_{\mu v} + \sum_{\bar{k}} \langle \mu v \parallel \bar{k} \rangle (x_{\bar{k}} + z_{\bar{k}}) \right) c_{vi} \\ &- 4 \sum_v \sum_j S_{\mu v} c_{vj} \epsilon_{ji} \quad \forall \mu, i. \end{aligned} \quad (40)$$

The derivative of the exchange-correlation energy, again restricting ourselves to local functionals, now contains the fitted density:

$$\frac{\partial E_{xc}[\tilde{\rho}]}{\partial c_{\mu i}} = \int \frac{\delta E_{xc}[\tilde{\rho}]}{\delta \tilde{\rho}(\mathbf{r})} \frac{\partial \tilde{\rho}(\mathbf{r})}{\partial c_{\mu i}} d\mathbf{r}. \quad (41)$$

The resulting functional derivative defines the exchange-correlation potential calculated with the fitted density:

$$v_{xc}[\tilde{\rho}; \mathbf{r}] \equiv \frac{\delta E_{xc}[\tilde{\rho}]}{\delta \tilde{\rho}(\mathbf{r})}. \quad (42)$$

The derivative of the fitted density with respect to molecular orbital coefficients is given by

$$\frac{\partial \tilde{\rho}(\mathbf{r})}{\partial c_{\mu i}} = \sum_{\bar{l}} \frac{\partial x_{\bar{l}}}{\partial c_{\mu i}} \bar{l}(\mathbf{r}). \quad (43)$$

After differentiation of Eq. (32), it follows for the derivative of the fitted density:

$$\frac{\partial \tilde{\rho}(\mathbf{r})}{\partial c_{\mu i}} = 4 \sum_{\bar{k}, \bar{l}} \sum_v c_{vi} \langle \mu v \parallel \bar{k} \rangle G_{\bar{k}\bar{l}}^{-1} \bar{l}(\mathbf{r}). \quad (44)$$

Substituting this expression and the definition of the approximate exchange-correlation potential Eq. (42) into Eq. (41) yields

$$\frac{\partial E_{xc}[\tilde{\rho}]}{\partial c_{\mu i}} = 4 \sum_{\bar{k}} \sum_v c_{vi} \langle \mu v \parallel \bar{k} \rangle z_{\bar{k}}, \quad (45)$$

with

$$z_{\bar{k}} = \sum_{\bar{l}} G_{\bar{k}\bar{l}}^{-1} \langle \bar{l} | v_{xc}[\tilde{\rho}] \rangle. \quad (46)$$

In order to distinguish the $x_{\bar{k}}$ and $z_{\bar{k}}$ coefficients, we name them Coulomb and exchange-correlation coefficients, respectively. The corresponding Kohn-Sham matrix elements are defined as

$$K_{\mu\nu} = H_{\mu\nu} + \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle (x_{\bar{k}} + z_{\bar{k}}). \quad (47)$$

These matrix elements depend only on the fitting coefficients. In contrast to the traditional fitting of the exchange-correlation potential, ADFT employs the fitted density for the exchange-correlation potential calculation, and hence the numerical integration of v_{xc} is not avoided. However, the work on the grid is considerably reduced because only one-center terms are involved in the integrals in Eq. (46). Despite the fact that these terms are evaluated numerically, the calculation of the Kohn-Sham elements Eq. (47) scales linearly (Köster et al. 2003). Because the ADFT energy is variational, analytic gradients (Köster et al. 2004b) and higher-energy derivatives can be formulated.

Auxiliary Density Perturbation Theory

For second energy derivatives that appear in the calculation of polarizabilities, chemical hardness, van der Waals coefficients, vibrational frequencies, and other second-order properties, the perturbed density matrix is required. McWeeny's self-consistent perturbation (SCP) theory (McWeeny 1962, 2001; Diercksen and McWeeny 1966; McWeeny and Diercksen 1968; Dodds et al. 1977; Raynes et al. 1977; McWeeny et al. 1977) represents a direct approach for the calculation of this matrix. For clarity of presentation, we assume perturbation-independent basis and auxiliary functions and restrict ourselves to closed-shell systems. Under these conditions, the elements of the perturbed density matrix are given by the SCP formalism as (McWeeny et al. 1977)

$$P_{\mu\nu}^{(\lambda)} \equiv \frac{\partial P_{\mu\nu}}{\partial \lambda} = 2 \sum_i^{\text{occ}} \sum_a^{\text{uno}} \frac{\mathcal{K}_{ia}^{(\lambda)}}{\varepsilon_i - \varepsilon_a} (c_{\mu i} c_{\nu a} + c_{\mu a} c_{\nu i}). \quad (48)$$

Extension to the case of perturbation-dependent basis and auxiliary functions is straightforward but can lead to tedious expressions. Therefore, we will discuss here only Eq. (48). In this equation, λ denotes the perturbation parameter, e.g., an electric field component in the calculation of polarizabilities, ε_i and ε_a orbital energies of the i th occupied and a th unoccupied orbital, and $\mathcal{K}_{ia}^{(\lambda)}$ the perturbed Kohn-Sham matrix in the molecular orbital representation:

$$\mathcal{K}_{ia}^{(\lambda)} = \sum_{\mu,\nu} c_{\mu i} c_{\nu a} K_{\mu\nu}^{(\lambda)}, \quad (49)$$

with

$$K_{\mu\nu}^{(\lambda)} = H_{\mu\nu}^{(\lambda)} + \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle (x_{\bar{k}}^{(\lambda)} + z_{\bar{k}}^{(\lambda)}). \quad (50)$$

The perturbation of the exchange-correlation coefficients is given by

$$z_{\bar{k}}^{(\lambda)} = \sum_{\bar{l}} G_{\bar{k}\bar{l}}^{-1} \langle \bar{l} | v_{xc}^{(\lambda)}[\bar{\rho}] \rangle. \quad (51)$$

Since $v_{xc}[\bar{\rho}; \mathbf{r}]$ itself is a (local) functional of the density, it follows

$$\begin{aligned} \langle \bar{l} | v_{xc}^{(\lambda)}[\bar{\rho}] \rangle &= \iint \bar{l}(\mathbf{r}) \frac{\delta v_{xc}[\bar{\rho}; \mathbf{r}]}{\delta \bar{\rho}(\mathbf{r}')} \frac{\partial \bar{\rho}(\mathbf{r}')}{\partial \lambda} d\mathbf{r} d\mathbf{r}' \\ &= \sum_{\bar{m}} \langle \bar{l} | f_{xc}[\bar{\rho}] | \bar{m} \rangle x_{\bar{m}}^{(\lambda)}, \end{aligned} \quad (52)$$

with the exchange-correlation kernel defined as

$$f_{xc}(\mathbf{r}, \mathbf{r}') = \frac{\delta^2 E_{xc}[\bar{\rho}]}{\delta \bar{\rho}(\mathbf{r}') \delta \bar{\rho}(\mathbf{r})} = \frac{\delta v_{xc}[\bar{\rho}; \mathbf{r}]}{\delta \bar{\rho}(\mathbf{r}')}. \quad (53)$$

Compared to the standard LCGTO kernel $\langle \mu\nu | f_{xc} | \sigma\tau \rangle$, the scaling of the kernel calculation is reduced by almost two orders of magnitude in the ADPT approach. With this result, we now rewrite the perturbed Kohn-Sham matrix elements in molecular orbital representation as

$$\mathcal{K}_{ia}^{(\lambda)} = \mathcal{H}_{ia}^{(\lambda)} + \sum_{\bar{k}, \bar{l}} \langle ia | \bar{k} \rangle M_{\bar{k}\bar{l}} x_{\bar{l}}^{(\lambda)}, \quad (54)$$

with

$$M_{\bar{k}\bar{l}} = \delta_{\bar{k}\bar{l}} + \sum_{\bar{m}} G_{\bar{k}\bar{m}}^{-1} \langle \bar{m} | f_{xc}[\bar{\rho}] | \bar{l} \rangle. \quad (55)$$

On the other hand, we find as the derivative of the fitting equation system Eq. (31):

$$\sum_{\bar{k}} G_{\bar{m}\bar{k}} x_{\bar{k}}^{(\lambda)} = \sum_{\mu, \nu} P_{\mu\nu}^{(\lambda)} \langle \mu\nu | \bar{m} \rangle. \quad (56)$$

Substituting Eq. (54) and Eq. (56) into Eq. (48) yields

$$\begin{aligned} \sum_{\mu, \nu} P_{\mu\nu}^{(\lambda)} \langle \mu\nu | \bar{m} \rangle &= 4 \sum_i^{\text{occ}} \sum_a^{\text{uno}} \frac{\langle \bar{m} | ia \rangle \mathcal{H}_{ia}^{(\lambda)}}{\varepsilon_i - \varepsilon_a} \\ &\quad + 4 \sum_i^{\text{occ}} \sum_a^{\text{uno}} \sum_{\bar{k}, \bar{l}} \frac{\langle \bar{m} | ia \rangle \langle ia | \bar{k} \rangle}{\varepsilon_i - \varepsilon_a} M_{\bar{k}\bar{l}} x_{\bar{l}}^{(\lambda)} \\ &= \sum_{\bar{k}} G_{\bar{m}\bar{k}} x_{\bar{k}}^{(\lambda)}. \end{aligned} \quad (57)$$

We now define the elements of the Coulomb coupling matrix as

$$A_{\bar{k}\bar{l}} = \sum_i^{\text{occ}} \sum_a^{\text{uno}} \frac{\langle \bar{k} \| ia \rangle \langle ia \| \bar{l} \rangle}{\varepsilon_i - \varepsilon_a}. \quad (58)$$

Similarly, the elements of the perturbation vector are given by

$$b_{\bar{k}}^{(\lambda)} = \sum_i^{\text{occ}} \sum_a^{\text{uno}} \frac{\langle \bar{k} \| ia \rangle \mathcal{H}_{ia}^{(\lambda)}}{\varepsilon_i - \varepsilon_a}. \quad (59)$$

With these matrices and vectors, Eq. (57) can now be recast into

$$(\mathbf{G} - 4\mathbf{A}\mathbf{M}) \mathbf{x}^{(\lambda)} = 4\mathbf{b}^{(\lambda)} \iff \mathbf{x}^{(\lambda)} = \left(\frac{1}{4} \mathbf{G} - \mathbf{A}\mathbf{M} \right)^{-1} \mathbf{b}^{(\lambda)}. \quad (60)$$

In the case of perturbation-dependent basis and auxiliary functions, only $\mathbf{b}^{(\lambda)}$ has to be modified. The matrix that needs to be inverted is exactly the same as in Eq. (60). Thus, the calculation of the perturbed fitting coefficients is reduced to the solution of the above inhomogeneous equation system. In contrast to the traditional coupled-perturbed Kohn-Sham equation system (Fournier 1990; Komornicki and Fitzgerald 1993), the dimension of Eq. (60) is M^2 , and, therefore, the memory requirement for the solution of this equation system is similar to that of the fitting equation system Eq. (31) in the corresponding SCF calculation. With the perturbed fitting coefficients, the perturbed Kohn-Sham matrix can be constructed via Eq. (54), and the perturbed density matrix can then be calculated by Eq. (48).

So far Eq. (60) has been applied successfully for the calculation of static and dynamic molecular dipole-dipole and dipole-quadrupole polarizabilities (Flores-Moreno and Köster 2008; Shedge et al. 2010, 2012, 2011; Carmona-Espíndola et al. 2010; Calaminici et al. 2012; Mejia-Rodriguez et al. 2015b) and static and dynamic molecular hyperpolarizabilities (Carmona-Espíndola et al. 2012; Karne et al. 2015) as well as for the calculation of Fukui functions (Flores-Moreno et al. 2008; Flores-Moreno 2010). Also a time-dependent ADFT variant was developed (Ipatov et al. 2006; Carmona and Köster 2013). Calculation of second derivatives with respect to nuclear displacements is currently under development.

ERI Expansion in ADFT

Due to the efficient calculation of the exchange-correlation potential in ADFT, the calculation of three-center ERIs, Eq. (27), arises as a potential bottleneck in the ADFT SCF procedure. The reason is twofold. First, the number of ERIs increases formally as $N^2 \times M$ (N basis functions and M auxiliary functions) with system size. Second, the increasing popularity of BOMD simulations raises demand for fast SCF solutions for small- to medium-sized systems. Two major developments have been

recently implemented in deMon2k to improve the computational performance of the ERI calculation. These are the double asymptotic expansion of ERIs (Alvarez-Ibarra and Köster 2013) and the mixed SCF scheme (Alvarez-Ibarra and Köster 2015). A brief discussion of these implementations is presented in the following subsections.

Double Asymptotic Expansion of ERIs

In large molecular systems, a large portion of ERIs involves atoms that are far apart from each other. Due to the form of the Coulomb operator, ERIs cannot be discarded loosely. However, the calculation of so-called far-field ERIs (because they occur in the far field of the electron density of the involved atoms) can be reduced by performing multipole expansions of the Coulomb operator. Multipole expansions are only feasible if the electron density centers are widely separated from each other, limiting their use to large molecular systems. The fast multipole method (FMM), originally developed by Greengard and Rokhlin (1987), has been adapted to electronic structure methods by several researchers (White and Head-Gordon 1994; Petersen et al. 1994; Kudin and Scuseria 1998; White et al. 1995; Strain et al. 1996). However, the double asymptotic ERI expansion used in deMon2k differs from the FMM in two ways (Alvarez-Ibarra and Köster 2013). First, the asymptotic expansion is performed in the final form of the ERI thus allowing the expansion to be applied independently from other integrals (e.g., nuclear-attraction integrals and other ERIs). Second, the space division procedure is avoided in exchange for an ERI classification based on the distance between the atomic centers and the extension of the orbitals and the Coulomb potential generated by the auxiliary function (information that is known from the input definition). The double asymptotic expansion of far-field ERIs has the following form:

$$\langle \mu v || \bar{k} \rangle^A \sim \left(\frac{\pi}{\zeta_{\bar{k}}} \right)^{3/2} \sum_{m_x, m_y, m_z} \frac{(-1)^M}{m_x! m_y! m_z!} \left(\frac{\pi}{\zeta_{\bar{k}}} \right)^{3/2} T_{AC}(m + \bar{k}) S_{\mu v}^A(m), \quad (61)$$

where

$$T_{AC}(m + \bar{k}) = \left(\frac{\partial}{\partial C_x} \right)^{m_x + \bar{k}_x} \left(\frac{\partial}{\partial C_y} \right)^{m_y + \bar{k}_y} \left(\frac{\partial}{\partial C_z} \right)^{m_z + \bar{k}_z} \frac{1}{|\vec{A} - \vec{C}|}, \quad (62)$$

$$S_{\mu v}^A(m) = \langle \mu + m | v \rangle, \quad (63)$$

and m_x , m_y and m_z are the expansion indices that sum up to M , i.e., $M = m_x + m_y + m_z$. Orbital μ is centered on A , whereas the Coulomb potential generated by the auxiliary function \bar{k} is centered on C . The superscript in the expanded integral Eq. (61) indicates the expansion center. There is an equivalent expansion for orbital v (Alvarez-Ibarra and Köster 2013). The double asymptotic expansion transforms the two-electron three-center integral, Eq. (27), into a series of modified

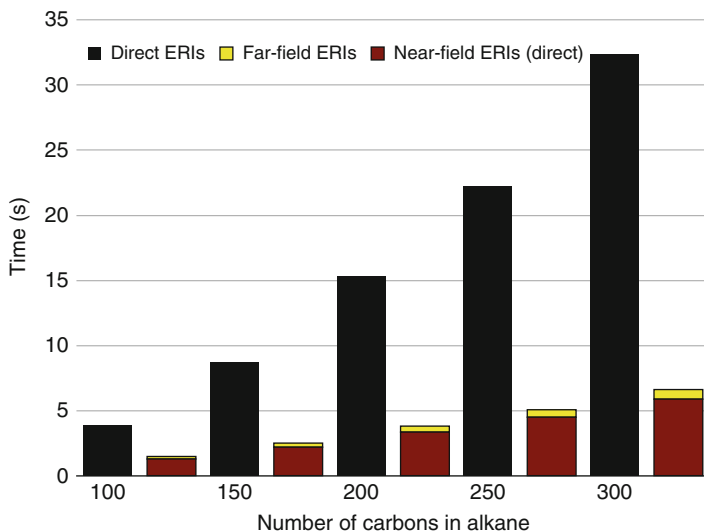


Fig. 1 Calculation time per SCF cycle of ERIs for selected n-alkane systems. *Black bars* indicate times using the direct SCF (for all integrals). *Red and yellow bars* indicate near- and far-field ERI times, respectively. *Red and yellow bars* add to the total time using the near- and far-field separation as described in the text

two-center overlap integrals with prefactors that depend only on atomic coordinates. When the number of far-field ERIs dominates significantly over the number of near-field ERIs (ERIs for which the asymptotic expansion is not applicable), the factorization of auxiliary function information to the atomic level, Eq. (62), and the efficient algorithms for the calculation of overlap integrals (Obara and Saika 1986) reduce enormously the calculation time required for the two-electron ERIs.

Figure 1 shows the calculation time for ERIs per SCF cycle in a number of selected n-alkanes using sixty 2.53 GHz Xeon E5649 processors with 2 GB RAM per CPU. The DZVP basis set and the GEN-A2 auxiliary function set were used for the calculation, along with the VWN functional for the exchange-correlation potential. Black bars indicate the time for ERI calculations with the optimized recurrence relations, i.e., treating all ERIs as if they were in the near field, whereas red and yellow bars indicate the time used calculating near-field ERIs via recurrence relations and far-field ERIs via the double asymptotic expansion, respectively. Since near-field ERIs are calculated using the same formulas as in the direct SCF, the difference between bars corresponds to the time savings obtained by the double asymptotic expansion of far-field ERIs. For the smallest n-alkane system ($C_{100}H_{202}$), the ERI calculation time is already reduced by half employing the double asymptotic expansion. Note that this plot shows a linear scaling for the new ERI calculation scheme with respect to system size. This is due to the large number of far-field

ERIs in the alkane chains with respect to the number of near-field ERIs. Note also that for systems with few or non-existing far-field ERIs, the new method will not significantly impact CPU timings. In order to address this situation, a new SCF scheme is proposed, which is presented in the following subsection.

The Mixed SCF Scheme

Regardless of the molecular system size, the ERI calculation may become a bottleneck in the direct SCF procedure if many iteration cycles are required for convergence. This is particularly true if the variational fitting of the Coulomb potential is used since ERIs are needed twice per SCF cycle, namely, for the calculation of the Coulomb vector, Eq. (30), and the calculation of the Kohn-Sham matrix, Eq. (38). In this situation, even the calculation of near-field ERIs alone can become a bottleneck since, according to Fig. 1, these calculations dominate the total ERI calculation time. Since multipole or asymptotic expansions are not applicable to these ERIs, the solution implemented in deMon2k is the so-called mixed SCF scheme (Alvarez-Ibarra and Köster 2015), named after the combination of the conventional (Clementi 1963) and direct (Almlöf 1982) SCF schemes. In the mixed SCF, the near-field ERIs are calculated once and stored in memory (as in the conventional SCF), whereas the far-field ERIs are calculated in a recurrent fashion (as in the direct SCF) using the efficient double asymptotic expansion. The difference between the mixed SCF and other modified SCF schemes is that the determination of the ERIs to be stored is based on the near- and far-field ERI separation, which is readily available from input information. In order to avoid I/O operations, the mixed SCF scheme stores near-field ERIs only in RAM, never on hard disk. If no far-field ERIs are in the system, the mixed SCF reduces to a conventional SCF that uses only RAM for ERI storage.

Initially the mixed SCF algorithm determines the amount of RAM necessary for the allocation of the working fields for the SCF procedure. Then, the amount of free RAM is used to store near-field ERIs. When the SCF procedure is about to start, near-field ERIs are calculated and stored in RAM. Then, in the SCF cycles, they are used in the two corresponding modules, Coulomb vector and Kohn-Sham matrix calculation. In the parallel version of the mixed SCF, only RAM from the slave processors is used to store ERIs, leaving the RAM of the master processor free to store SCF fields. Far-field ERIs are calculated on every processor since their memory demands are much smaller. This algorithm produces an outstanding performance, which allows the reduction of the ERI calculation time to a practically insignificant fraction of the total electronic structure calculation time. Since the mixed SCF scheme uses all the free RAM available in the computational architecture, calculation of larger systems is possible simply by requesting more processors for the job with no significant overhead due to the ERI calculation.

Figure 2 depicts the calculation time for ERIs per SCF cycle in the same *n*-alkanes from Fig. 1. Also the computational details are the same as for Fig. 1. In this case, the comparison is between the direct SCF with the double asymptotic expansion algorithm for the far-field ERIs and the mixed SCF scheme. Red and

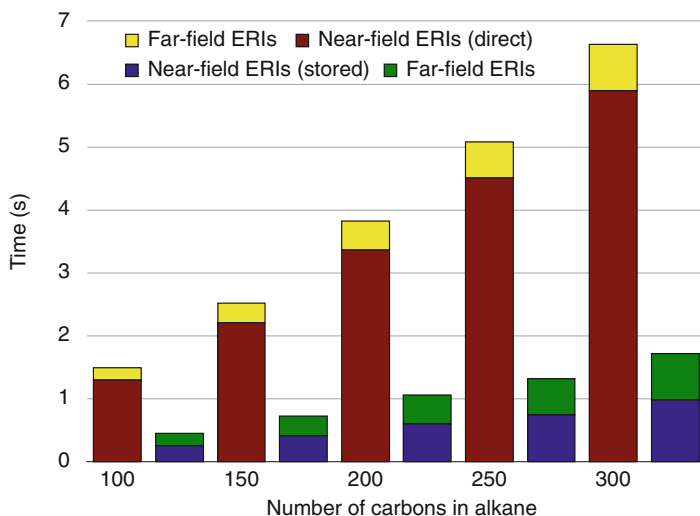


Fig. 2 Calculation time per SCF cycle of ERIs for selected n-alkane systems. *Red* and *yellow bars* indicate direct SCF near- and double asymptotically expanded far-field ERI times, respectively. *Blue* and *green bars* indicate mixed SCF near- and double asymptotically expanded far-field ERI times

yellow bars show the calculation time for near-field and far-field ERIs in the direct SCF, whereas blue and green bars indicate the calculation time for near-field and far-field ERIs in the mixed SCF. Note that far-field ERI times are the same for both SCF schemes, since they are calculated in the same way. For the near-field ERIs, in-RAM storage achieves speedups that range from three and a half to six in this example. If we compare times from the original direct SCF scheme, i.e., without the double asymptotic ERI expansion, the ERI calculation time per SCF cycle for $C_{100}H_{202}$ is reduced from 4s to 0.5s, and for $C_{300}H_{602}$, it is reduced from 32s to less than 2s. As a conclusion, it is reasonable to state that the mixed SCF scheme eliminates the ERI computational bottleneck in ADFT when a computational architecture with enough RAM is used for the calculation.

Variational Fitting of Exact Exchange in ADFT

Hybrid functionals, where a certain portion of exact exchange is admixed with an exchange-correlation functional, are now a standard tool in quantum chemical calculations (Becke 2014). The functional form of exact exchange in DFT is equivalent to the Fock exchange in Hartree-Fock theory. Therefore, computing contributions from exact exchange in Kohn-Sham methods leads to a formal N^4 scaling. In order to avoid this unfavorable scaling, we implemented a variational fitting exact exchange approach in deMon2k which allows the computation of hybrid functionals with both the AUXIS and BASIS approaches.

In closed-shell systems, exact exchange is given by

$$E_{\text{EXX}} = -\frac{1}{4} \sum_{\mu, \nu} \sum_{\sigma, \tau} P_{\mu\sigma} P_{\nu\tau} \langle \mu\nu \| \sigma\tau \rangle \quad (64)$$

$$= -\sum_{i,j}^{\text{occ}} \langle ij \| ij \rangle, \quad (65)$$

where we used Eq. (17) for transforming Eq. (64) into Eq. (65). In the variational fitting exact exchange approach, the orbital products

$$\rho_{ij}(\mathbf{r}) \equiv \psi_i(\mathbf{r})\psi_j(\mathbf{r}) \quad (66)$$

are expanded in linear combinations of the same atom-centered primitive Hermite Gaussian auxiliary functions used to approximate the density:

$$\tilde{\rho}_{ij}(\mathbf{r}) = \sum_{\bar{k}} x_{\bar{k}ij} \bar{k}(\mathbf{r}). \quad (67)$$

For the variational fitting, the self-interaction term,

$$\varepsilon_2^{\text{EXX}} = -\sum_{i,j}^{\text{occ}} \iint \frac{[\rho_{ij}(\mathbf{r}) - \tilde{\rho}_{ij}(\mathbf{r})][\rho_{ij}(\mathbf{r}') - \tilde{\rho}_{ij}(\mathbf{r}')]}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}', \quad (68)$$

which is negative definite, must be maximized. Thus, the variations of $\varepsilon_2^{\text{EXX}}$ with respect to each fitting coefficient $x_{\bar{k}ij}$ must vanish:

$$\frac{\partial \varepsilon_2^{\text{EXX}}}{\partial x_{\bar{k}ij}} = \sum_{i,j}^{\text{occ}} \langle ij \| \bar{k} \rangle - \sum_{i,j}^{\text{occ}} \sum_{\bar{l}} \langle \bar{k} \| \bar{l} \rangle x_{\bar{l}ij} \equiv 0 \quad \forall \bar{k}, i, j. \quad (69)$$

Similar to the variational Coulomb fitting, one can cast a set of inhomogeneous equation systems, one for each orbital product

$$\mathbf{G}\mathbf{x}_{ij} = \mathbf{J}_{ij} \quad \forall i, j, \quad (70)$$

where \mathbf{x}_{ij} collects the fitting coefficients for each orbital product and \mathbf{J}_{ij} is defined as

$$\mathbf{J}_{ij} = \begin{pmatrix} \langle ij \| \bar{1} \rangle \\ \langle ij \| \bar{2} \rangle \\ \vdots \\ \langle ij \| \bar{m} \rangle \end{pmatrix}. \quad (71)$$

It is important to note that $\langle ij || \bar{k} \rangle = \langle ji || \bar{k} \rangle$ by the permutational symmetry properties of the three-center electron repulsion integrals. Therefore, only $N_{\text{occ}} \times (N_{\text{occ}} + 1)/2$ inhomogeneous equation systems are needed. After the set of equation systems, defined in Eq. (70), has been solved, the fitted exact exchange energy can be computed as

$$\tilde{E}_{\text{EXX}} = - \sum_{i,j}^{\text{occ}} \sum_{\bar{k}, \bar{l}} x_{\bar{k}ij} \langle \bar{k} || \bar{l} \rangle x_{\bar{l}ij} \quad (72)$$

$$= - \sum_{i,j}^{\text{occ}} \sum_{\bar{k}, \bar{l}} \langle ij || \bar{k} \rangle G_{\bar{k}\bar{l}}^{-1} \langle \bar{l} || ij \rangle. \quad (73)$$

The total hybrid ADFT energy thus becomes

$$E = \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu} + \sum_{\bar{k}} \sum_{\mu,\nu} P_{\mu\nu} \langle \mu\nu || \bar{k} \rangle x_{\bar{k}} - \frac{1}{2} \sum_{\bar{k}, \bar{l}} x_{\bar{k}} x_{\bar{l}} \langle \bar{k} || \bar{l} \rangle + E_c[\tilde{\rho}] + (1 - \alpha) E_x[\tilde{\rho}] - \alpha \sum_{i,j}^{\text{occ}} \sum_{\bar{k}, \bar{l}} \langle ij || \bar{k} \rangle G_{\bar{k}\bar{l}}^{-1} \langle \bar{l} || ij \rangle, \quad (74)$$

where the exchange-correlation energy has been explicitly separated. Eq. (74) is only an illustrative case of how to build a hybrid functional, and, in practice, more than one mixing parameter can be found in a hybrid definition. Variation of Eq. (74) with respect to the molecular orbital coefficients under the orthonormality constraint, Eq. (34), yields

$$\begin{aligned} \frac{\partial E}{\partial c_{\mu i}} &= 4 \sum_{\nu} \left(H_{\mu\nu} + \sum_{\bar{k}} \langle \mu\nu || \bar{k} \rangle (x_{\bar{k}} + z_{\bar{k}}) - \alpha \sum_j^{\text{occ}} \sum_{\bar{k}, \bar{l}} \langle \mu j || \bar{k} \rangle G_{\bar{k}\bar{l}}^{-1} \langle \bar{l} || \nu j \rangle \right) c_{\nu i} \\ &\quad - 4 \sum_{\nu} \sum_j S_{\mu\nu} c_{\nu j} \epsilon_{ji} \end{aligned} \quad (75)$$

with

$$z_{\bar{k}} = \sum_{\bar{l}} G_{\bar{k}\bar{l}}^{-1} \langle \bar{l} || (1 - \alpha) v_x[\tilde{\rho}] + v_c[\tilde{\rho}] \rangle. \quad (76)$$

The corresponding Kohn-Sham matrix elements are defined as

$$K_{\mu\nu} = H_{\mu\nu} + \sum_{\bar{k}} \langle \mu\nu || \bar{k} \rangle (x_{\bar{k}} + z_{\bar{k}}) - \alpha \sum_j^{\text{occ}} \sum_{\bar{k}, \bar{l}} \langle \mu j || \bar{k} \rangle G_{\bar{k}\bar{l}}^{-1} \langle \bar{l} || \nu j \rangle. \quad (77)$$

Straightforward implementation of Eqs. (73) and (77) leads to an algorithm that scales as $N_{\text{occ}} \times N \times M^2$. It has been noted that this quartic scaling algorithm is only

advantageous to its four-center counterpart, Eq. (65), when $N_{\text{occ}} \ll N$, i.e., using basis sets of at least triple-zeta quality.

Many groups have proposed modifications to the variational exact exchange fitting method giving rise to algorithms like the atomic resolution of the identity (Sodt and Head-Gordon 2008), local density fitting (LDF) (Polly et al. 2004; Mejia-Rodriguez and Köster 2014), concentric atomic density fitting (Hollman et al. 2014), pair atomic resolution of the identity (Watson et al. 2003; Merlot et al. 2013; Manzer et al. 2015), as well as atom-centered Cholesky decomposition approaches (Aquilante et al. 2007). We have recently developed an LDF approach to exact exchange (LDF-EXX) which has proven to be robust and efficient (Mejia-Rodriguez and Köster 2014). LDF-EXX is based on the fact that the exact exchange contribution to the Kohn-Sham matrix,

$$-\alpha \sum_j^{\text{occ}} \sum_{\bar{k}, \bar{l}} \langle \mu_j \| \bar{k} \rangle G_{\bar{k}\bar{l}}^{-1} \langle \bar{l} \| \nu_j \rangle, \quad (78)$$

is invariant under unitary transformations of the molecular orbitals. Thus, any set of molecular orbitals generated through a standard localization procedure, like Edmiston-Ruedenberg (Edmiston and Ruedenberg 1963), Foster-Boys (Boys 1960; Foster and Boys 1960), or Pipek-Mezey (Pipek and Mezey 1989), can be used to calculate the exact exchange contribution. By using localized molecular orbitals, the three-center ERIs appearing in Eq. 78 have non-negligible values only when atomic orbitals $\mu(\mathbf{r})$ and $\nu(\mathbf{r})$ are close in space to the localized molecular orbital $\psi_j(\mathbf{r})$. This allows the definition of local sets of basis functions around each localized molecular orbital and the screening of a large number of integrals. Furthermore, since all orbital products $\psi_j(\mathbf{r})\mu(\mathbf{r})$ are also localized in space, the auxiliary functions needed to accurately fit these products can also be restricted to a region close to $\psi_j(\mathbf{r})$. In this way, local auxiliary function sets can also be defined for each localized molecular orbital. For sufficiently large systems, the sizes of the local sets remain constant, reducing the scaling to N_{occ} with an $N_{\text{local}} \times M_{\text{local}}^2$ constant prefactor.

A very important ingredient to ensure that N_{local} and M_{local} are kept at their smallest possible size is the localization procedure. Unfortunately, orbital localization algorithms based on Jacobi sweeps scale as N^3 and can become very unstable when large basis sets are employed. In order to partially overcome this situation, a combination of the pivoted Cholesky decomposition of the density matrix (Aquilante et al. 2006) followed by the Foster-Boys algorithm is employed in our LDF-EXX. This combination yields tightly localized molecular orbitals without jeopardizing the potential linear scaling of the LDF-EXX algorithm for systems composed of a few hundred atoms (Mejia-Rodriguez and Köster 2014), despite the fact that both steps (Cholesky decomposition and Foster-Boys localization) scale cubically. Thus, the LDF-EXX algorithm can be summarized as follows:

1. Pivoted Cholesky decomposition of $\mathbf{P} \mapsto \mathbf{C}$.
2. Foster-Boys localization of the resulting \mathbf{C} matrix.

3. Selection of the local basis and auxiliary function sets for each MO.
4. Calculation of the ERI matrix, with elements $E_{\bar{k}\mu,i} = \langle \bar{k} || \mu i \rangle$, for each MO.
5. Calculation, Cholesky decomposition, and inversion of $\mathbf{G}_i = \mathbf{L}_i \mathbf{L}_i^T$ for each MO.
6. Compute exact exchange contribution to the Kohn-Sham matrix: $-\alpha \sum_i^{\text{occ}} \mathbf{E}_i^T \mathbf{G}_i^{-1} \mathbf{E}_i$.

The computational efficiency of the LDF-EXX algorithm depends crucially on the selection of the local sets in Step 3. Our selection is based on atomic Löwdin populations for each MO. All auxiliary and basis functions centered on atoms with significant Löwdin populations for a given MO will form the local auxiliary and basis sets, respectively, of that particular localized orbital. In addition, the local basis set is extended with basis functions centered on atoms with non-negligible overlaps to any function in the previously selected local basis set. In deMon2k, significant atomic Löwdin populations are defined as the n largest atomic populations whose sum does not exceed a predefined threshold set by default to 0.99. Also in this context, non-negligible overlaps are those whose absolute values are greater than 10^{-6} a.u. (Mejia-Rodriguez and Köster 2014).

The extension of the preceding hybrid ADFT energy expression, Eq. (74), to the BASIS approach is trivial and has also been implemented in deMon2k. While using the GEN-A2* auxiliary function set, the above algorithm yields standard enthalpies of formation within 0.5 kcal/mol with respect to a four-center reference for the entire G3/99 molecular test set (Curtiss et al. 1997, 1998, 2001) and the PBE0 (Perdew et al. 1996b; Adamo and Barone 1999), B3LYP (Stephens et al. 1994) or M06-2X (Zhao and Truhlar 2006) hybrid functionals (Mejia-Rodriguez et al. 2015a). It is important to note that the meta-GGA is only available through the BASIS approach.

DNA Double Strands

In order to demonstrate the advantages of the LDF-EXX algorithm implemented in deMon2k, we show Hartree-Fock calculations on DNA double strands containing up to 8 adenine-thymine pairs. The structures were generated with the Accelrys Discovery Studio Client in the DNA B-strand conformation with both ends capped with hydroxyl groups. The phosphate groups were protonated in order to obtain neutral systems.

The largest DNA strand, depicted in Fig. 3, has the molecular formula $\text{C}_{160}\text{H}_{202}\text{O}_{92}\text{N}_{56}\text{P}_{14}$ with a total of 524 atoms. All computations were performed using the 6-31G** basis set (Hehre et al. 1972; Francl et al. 1982) and the GEN-A2/GEN-A2* approach described in Mejia-Rodriguez and Köster (2014). The largest system contains about 6,000 basis functions and more than 15,000 auxiliary functions.

Figure 4 shows the time needed for single point Hartree-Fock calculations using the LDF-EXX algorithm in an Intel Xeon X5675 @ 3.07 GHz with 4 GB of RAM memory. Convergence was achieved after 10 SCF iterations in all cases. Figure 4 also shows the scaling between two consecutive points. It can be seen that the

Fig. 3 Ball and stick representation of the DNA double strand in the B-strand conformation with 8 adenine-thymine base pairs

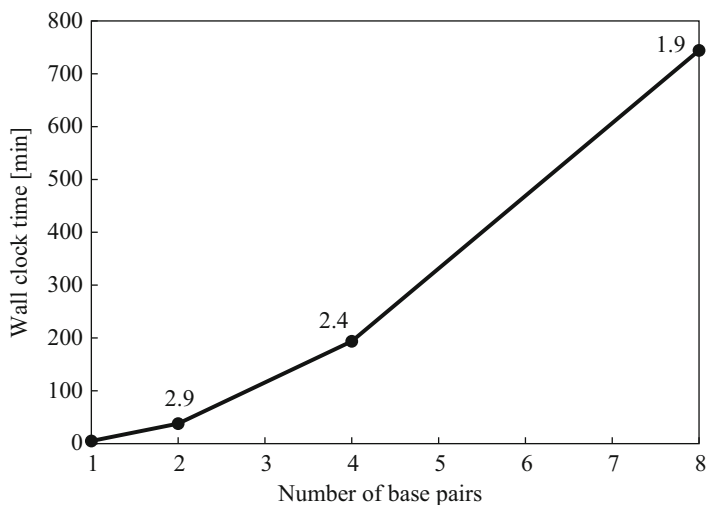
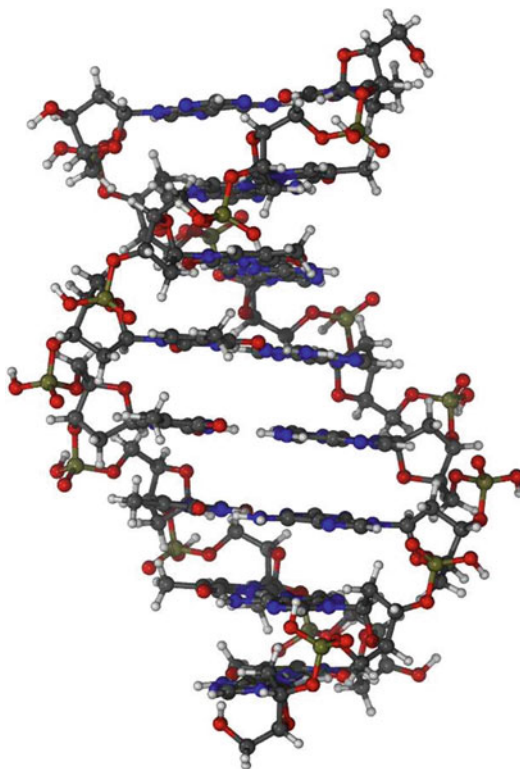


Fig. 4 Wall clock times, in minutes, for serial Hartree-Fock calculations of DNA double strands. All calculations were performed employing the LDF-EXX algorithm. The number above each point shows the scaling calculated with respect to the immediately anterior point

calculated scaling is diminishing and becomes subquadratic for the DNA₄ to DNA₈ step, showing the potential linear scaling that can be achieved by the algorithm.

It is important to note that the N^3 molecular orbital localization took less than 5% of the total time in all cases. Furthermore, since well-localized molecular orbitals were obtained, all \mathbf{E}_i matrices could be calculated in one batch up to DNA₄ and two batches for DNA₈. Despite the difference in the number of batches, the drop in the scaling remained constant for the steps DNA₂ to DNA₄ and DNA₄ to DNA₈.

Summary

In this section, a novel implementation of exact exchange has been reviewed. The algorithm has a potential linear scaling behavior for sufficiently large (extended) systems. The algorithm has been called local density-fitting exact exchange (LDF-EXX).

Furthermore, test calculations on DNA double strands containing more than 500 atoms and described with a double-zeta plus polarization basis set are shown. These results demonstrate the low-scaling behavior of the algorithm.

With the new LDF-EXX implementation, hybrid functional calculations within the ADFT framework are now possible in deMon2k. Further improvements are directed toward an efficient parallelization of the LDF-EXX algorithm so it can be employed in large-scale computations including BOMD simulations.

The ADFT-GIAO Methodology

The computation of magnetic properties from first principles demands a reformulation of the corresponding Hamilton operator. In order to include the magnetic field in the nonrelativistic Hamiltonian, the momentum operator \hat{p} is replaced by the conjugate momentum operator $\hat{\pi}$:

$$\hat{p} \longrightarrow \hat{\pi} = -i\vec{\nabla} + \vec{A}(\vec{r}). \quad (79)$$

The quantity $\vec{A}(\vec{r})$ denotes the vector potential of the total magnetic field, which is composed of two contributions. The first contribution, $\vec{A}^{\text{loc}}(\vec{r})$, originates from the intrinsic magnetic moment, $\vec{\mu}_C$, located at the position \vec{C} , of all the nuclei,

$$\vec{A}^{\text{loc}}(\vec{r}) = \frac{1}{c^2} \sum_C^{\text{Atoms}} \vec{\mu}_C \times \frac{\vec{r} - \vec{C}}{|\vec{r} - \vec{C}|^3}, \quad (80)$$

while the applied external magnetic field, \vec{B} , is responsible for the second one, $\vec{A}^{\text{ext}}(\vec{r})$,

$$\vec{A}^{\text{ext}}(\vec{r}) = \frac{1}{2} \vec{B} \times (\vec{r} - \vec{G}). \quad (81)$$

Here \vec{G} denotes the arbitrary origin of the vector potential, i.e., the gauge. The basic physical laws require that the observables are invariant with respect to the

chosen gauge origin. However, this gauge invariance is not necessarily ensured for approximate solutions of the Schrödinger equation. The origin of the problem has been traced back to the finite-basis set representation used for the molecular orbital expansion. Therefore, the enlargement of the basis set can resolve the problem, but this approach is only feasible for very small systems. A more practical solution is provided by so-called local or distributed gauge origin methods. The idea is to consider a different gauge origin for each atomic or molecular orbital. The developed schemes following this idea permit the computation of gauge-invariant results independently of the size of the basis set. Nowadays, the gauge-including atomic orbital (GIAO) (Ditchfield 1972) scheme has been established as a standard for magnetic property calculations. In GIAO, the atomic orbitals depend on the electronic coordinates and on the perturbation, i.e., the applied magnetic field, \vec{B} . The perturbation-dependent atomic orbitals are built by attaching an exponential factor, known as London factor (London 1937), to the unperturbed atomic orbital,

$$\phi_{\mu}(\vec{r}, \vec{B}) = \mu(\vec{r}) e^{-\frac{i}{2} \vec{B} \cdot (\vec{A} - \vec{G}) \times \vec{r}}. \quad (82)$$

As a consequence, the ADFT energy, Eq. (39), is now expressed in terms of these GIAOs. This combination is named the ADFT-GIAO methodology. Further, the elements of the one-electron Hamiltonian matrix, $H_{\mu\nu}$, also change due to the introduction of the conjugated momentum operator, given in Eq. (79), for the evaluation of the kinetic energy. Introducing Eqs. (79) and (82) into Eq. (19), we find, after some lengthy algebra,

$$H_{\mu\nu} = \left\langle \phi_{\mu} \left| \hat{H} \right| \phi_{\nu} \right\rangle = \left\langle \mu \left| \mathcal{L}_{AB} \hat{H}_g \right| \nu \right\rangle \quad (83)$$

with

$$\mathcal{L}_{AB} = e^{\frac{i}{2} \vec{B} \cdot (\vec{A} - \vec{B}) \times \vec{r}}, \quad (84)$$

being the product of the GIAO pair London factors where the vector \vec{B} denotes the center of the basis function $\nu(\vec{r})$ in Eq. (83). The corresponding gauge-invariant one-electron Hamilton operator has the form

$$\hat{H}_g = \hat{H}_0 + \vec{B} \cdot \hat{\xi}_g^p + \sum_C^{\text{Atoms}} \vec{\mu}_C \cdot \hat{\sigma}_g^p(\vec{C}) + \vec{B} \cdot \hat{\xi}_g^d \cdot \vec{B} + \sum_C^{\text{Atoms}} \vec{B} \cdot \hat{\sigma}_g^d(\vec{C}) \cdot \vec{\mu}_C. \quad (85)$$

Here \hat{H}_0 is the sum of the usual kinetic and nuclear-attraction operators of the electrons, as given in Eq. (19). The second and third terms correspond to the gauge-invariant paramagnetic vector operators for the magnetizability,

$$\hat{\xi}_g^p(\vec{B}) = -\frac{i}{2} (\vec{r} - \vec{B}) \times \vec{\nabla}, \quad (86)$$

and for the magnetic shielding,

$$\hat{\sigma}_g^p(\vec{C}) = -\frac{i}{c^2} \frac{\vec{r} - \vec{C}}{|\vec{r} - \vec{C}|^3} \times \vec{\nabla}. \quad (87)$$

The fourth and fifth terms are the gauge-invariant diamagnetic tensor operators for the magnetizability,

$$\hat{\xi}_g^d(\vec{B}) = \frac{1}{8} \left[|\vec{r} - \vec{B}|^2 \mathbf{E} - (\vec{r} - \vec{B}) \otimes (\vec{r} - \vec{B}) \right] \quad (88)$$

and the magnetic shielding,

$$\hat{\sigma}_g^d(\vec{C}) = \frac{1}{2c^2} \left[\frac{(\vec{r} - \vec{B}) \cdot (\vec{r} - \vec{C})}{|\vec{r} - \vec{C}|^3} \mathbf{E} - \frac{(\vec{r} - \vec{C}) \otimes (\vec{r} - \vec{B})}{|\vec{r} - \vec{C}|^3} \right]. \quad (89)$$

As the applied magnetic field and the intrinsic magnetic moments are treated as perturbations, the corresponding properties are computed as second-order energy derivatives. An element of the magnetic shielding tensor (Zuniga-Gutierrez et al. 2011) is given by the second-order energy derivative with respect to a Cartesian component of the external magnetic field and of the intrinsic magnetic moment of the nucleus C , i.e.,

$$\sigma_{\eta\lambda}(\vec{C}) = \lim_{\substack{\vec{B} \rightarrow 0 \\ \mu_C \rightarrow 0}} \frac{\partial^2 E}{\partial \mathcal{B}_\lambda \partial \mu_{C,\eta}} = \sum_{\mu,\nu} P_{\mu\nu}^{(\lambda)} H_{\mu\nu}^{(\eta)}(\vec{C}) + \sum_{\mu,\nu} P_{\mu\nu} H_{\mu\nu}^{(\lambda\eta)}(\vec{C}). \quad (90)$$

Here λ and η denote Cartesian components (x , y or z). As enclosed superscripts, (λ) and (η) , of perturbed matrix elements, they denote differentiation with respect to \mathcal{B}_λ and $\mu_{C,\eta}$, respectively. Thus, the expression for the derivatives of $H_{\mu\nu}$ appearing in Eq. (90) are

$$H_{\mu\nu}^{(\eta)}(\vec{C}) = \left\langle \mu \left| \hat{\sigma}_{g,\eta}^p \right| \nu \right\rangle \quad (91)$$

and

$$H_{\mu\nu}^{(\eta\lambda)}(\vec{C}) = \left\langle \mu \left| \mathcal{L}_{AB}^{(\lambda)} \hat{\sigma}_{g,\eta}^p + \hat{\sigma}_{g,\lambda\eta}^d \right| \nu \right\rangle \quad (92)$$

with

$$\mathcal{L}_{AB}^{(\lambda)} = \frac{i}{2} \left[(\vec{A} - \vec{B}) \times \vec{r} \right]_\lambda. \quad (93)$$

Because of the use of GIAOs, the basis functions are perturbation dependent. Thus the perturbed density matrix is given by (McWeeny 2001)

$$P_{\mu\nu}^{(\lambda)} = 2 \sum_i^{\text{occ}} \sum_a^{\text{uno}} \frac{K_{ia}^{(\lambda)} - \varepsilon_i S_{ia}^{(\lambda)}}{\varepsilon_i - \varepsilon_a} (c_{\mu i} c_{\nu a} - c_{\mu a} c_{\nu i}) - \frac{1}{2} \sum_{\sigma, \kappa} P_{\mu\sigma} S_{\sigma\kappa}^{(\lambda)} P_{\kappa\nu}. \quad (94)$$

Here $S_{\mu\nu}^{(\lambda)}$ is the first derivative of an element of the overlap matrix,

$$S_{\mu\nu}^{(\lambda)} = \left\langle \mu \left| \mathcal{L}_{AB}^{(\lambda)} \right| \nu \right\rangle, \quad (95)$$

and its form is independent of the employed methodology, BASIS or AUXIS. Note that for perturbation-independent basis functions, the terms including derivatives of elements of the overlap matrix vanish and, consequently, Eq. (94) reduces to Eq. (48), except for the sign in the parentheses. This difference originates from the pure imaginary paramagnetic shielding operator, Eq. (87), that causes the response of the electron density. A closer examination of Eq. (94) reveals that the response of the molecular orbitals must also be purely imaginary. According to the definition of the electron density in terms of molecular orbitals, we find for the response of the electron density:

$$\begin{aligned} \rho^{(\lambda)}(\vec{r}, \vec{\mathcal{B}}) &= \lim_{\vec{\mathcal{B}} \rightarrow 0} \frac{\partial}{\partial \mathcal{B}_\lambda} \sum_i^{\text{occ}} [\psi_i(\vec{r})]^* \psi_i(\vec{r}) \\ &= \sum_i^{\text{occ}} [\psi_i^{(\lambda)}(\vec{r})]^* \psi_i(\vec{r}) + [\psi_i(\vec{r})]^* \psi_i^{(\lambda)}(\vec{r}) \\ &= 0. \end{aligned} \quad (96)$$

This result must be emulated by the auxiliary density in order to compute magnetic properties within the ADFT framework. To do so, we have chosen the simplest ansatz by using perturbation-independent auxiliary functions (Zuniga-Gutierrez et al. 2011). With this choice, the linear response for a magnetic perturbation of the auxiliary density vanishes:

$$\tilde{\rho}^{(\lambda)}(\vec{r}) = \frac{\partial \tilde{\rho}(\vec{r})}{\partial \mathcal{B}_\lambda} = 0. \quad (97)$$

In Eq. (94), $K_{\mu\nu}^{(\lambda)}$ is the magnetic perturbed Kohn-Sham matrix whose form depends on whether the Kohn-Sham (BASIS) or the auxiliary density (AUXIS) is used for the evaluation of the exchange-correlation contributions. For the first choice, (conventional DFT) $K_{\mu\nu}^{(\lambda)}$ is found to be

$$K_{\mu\nu}^{(\lambda)} = H_{\mu\nu}^{(\lambda)} + \sum_{\vec{k}} \left\langle \mathcal{L}_{AB}^{(\lambda)} \mu\nu \middle| \vec{k} \right\rangle x_{\vec{k}} + \lim_{\vec{\mathcal{B}} \rightarrow 0} \frac{\partial^2 E_{xc}[\rho]}{\partial \mathcal{B}_\lambda \partial P_{\mu\nu}} \quad (98)$$

while for the latter (ADFT) $K_{\mu\nu}^{(\lambda)}$ is

$$K_{\mu\nu}^{(\lambda)} = H_{\mu\nu}^{(\lambda)} + \sum_{\bar{k}} \left\langle \mathcal{L}_{AB}^{(\lambda)} \mu\nu \middle| \bar{k} \right\rangle (x_{\bar{k}} + z_{\bar{k}}). \quad (99)$$

Here $H_{\mu\nu}^{(\lambda)}$ denotes the first derivative of the one-electron Hamiltonian contributions, Eq. (83), whose form is

$$H_{\mu\nu}^{(\lambda)} = \left\langle \mu \middle| \mathcal{L}_{AB}^{(\lambda)} \hat{H}_0 \middle| \nu \right\rangle + \left\langle \mu \middle| \hat{\xi}_{g,\lambda}^p \middle| \nu \right\rangle. \quad (100)$$

The second term in Eq. (100) corresponds to the molecular integral of the λ -Cartesian component of the paramagnetic magnetizability vector operator, Eq. (86). The $H_{\mu\nu}^{(\lambda)}$ and electronic Coulomb repulsion energy, second term in Eqs. (98) and (99), have the same form for both methodologies. The difference between Eqs. (98) and (99) lies in the exchange-correlation contributions. Within the conventional DFT framework, this contribution is evaluated with the magnetic field-dependent electron density that generates new expressions which have to be evaluated by numerical integration. This is an expensive computational task. The number of additional integrals depends on the approximation of the exchange-correlation functional: Three additional integrals for LDA and six for GGA. On the other hand, the exchange-correlation contributions are uniquely accounted for by the exchange-correlation coefficients $\{z_{\bar{k}}\}$, Eq. (46), in the ADFT expression of $K_{\mu\nu}^{(\lambda)}$, Eq. (99). Note that these coefficients have already been evaluated in the preceding SCF calculation. Therefore, they are read from memory. As a result, no extra numerical integration is required within ADFT-GIAO.

The ADFT-GIAO methodology has also been used for the computation of the magnetizability tensor (Zuniga-Gutierrez et al. 2012). It is computed as the second energy derivative with respect to Cartesian components of the applied magnetic field,

$$\begin{aligned} \xi_{\tau\lambda} = \lim_{\vec{B} \rightarrow 0} \frac{\partial^2 E}{\partial B_\tau \partial B_\lambda} &= \sum_{\mu,\nu} P_{\mu\nu}^{(\lambda)} K_{\mu\nu}^{(\tau)} + \sum_{\mu,\nu} P_{\mu\nu} \left[H_{\mu\nu}^{(\tau\lambda)} + \left\langle \mathcal{L}_{AB}^{(\tau\lambda)} \mu\nu \middle| \bar{k} \right\rangle (x_{\bar{k}} + z_{\bar{k}}) \right] \\ &\quad - \sum_{\mu,\nu} S_{\mu\nu}^{(\tau\lambda)} W_{\mu\nu} - \sum_{\mu,\nu} S_{\mu\nu}^{(\lambda)} W_{\mu\nu}^{(\tau)}. \end{aligned} \quad (101)$$

Here $W_{\mu\nu}$ and $W_{\mu\nu}^{(\tau)}$ denote elements of the weighted density matrix and its corresponding derivative. Moreover, second derivatives of elements of the one-electron Hamiltonian matrix,

$$\begin{aligned} H_{\mu\nu}^{(\tau\lambda)} &= \left\langle \mu \middle| \mathcal{L}_{AB}^{(\tau)} \mathcal{L}_{AB}^{(\lambda)} \hat{H}_0 \middle| \nu \right\rangle + \left\langle \mu \middle| \mathcal{L}_{AB}^{(\lambda)} \hat{\xi}_{g,\tau}^p \middle| \nu \right\rangle \\ &\quad + \left\langle \mu \middle| \mathcal{L}_{AB}^{(\tau)} \hat{\xi}_{g,\lambda}^p \middle| \nu \right\rangle + \left\langle \mu \middle| \hat{\xi}_{g,\lambda\tau}^d \middle| \nu \right\rangle, \end{aligned} \quad (102)$$

and of the overlap matrix,

$$S_{\mu\nu}^{(\tau\lambda)} = \left\langle \mu \left| \mathcal{L}_{AB}^{(\tau\lambda)} \right| \nu \right\rangle, \quad (103)$$

with

$$\mathcal{L}_{AB}^{(\tau\lambda)} = -\frac{1}{4} \left[\left(\vec{A} - \vec{B} \right) \times \vec{r} \right]_{\tau} \left[\left(\vec{A} - \vec{B} \right) \times \vec{r} \right]_{\lambda}, \quad (104)$$

are needed for the calculation of magnetizability tensor elements. These matrix elements can be straightforwardly calculated by extending molecular integral recurrence relations.

Magnetic properties related to the nuclear motion have also been computed with the ADFT-GIAO methodology, i.e., the rotational g-tensor (Zuniga-Gutierrez et al. 2015) and the nuclear spin-rotation constants (NSRC) (Zuniga-Gutierrez et al. 2015), that are experimentally accessible by molecular beam or microwave spectroscopy, respectively. Theoretically, these spectroscopic parameters are calculated as second energy derivatives with respect to a Cartesian element of the rotational angular momentum, \vec{J} , and either an element of the applied magnetic field for the rotational g-tensor or an element of the nuclear intrinsic magnetic moment for the NSRC. It is worth to mention that the rotational g-tensors and the NSRC can also be expressed in terms of the magnetizability and magnetic shielding tensors (Sauer 2011). They represent the most complex quantities to solve for the corresponding equations. Furthermore, due to the inherent efficiency of the ADFT-GIAO methodology, the 3D plotting of the induced magnetic field has become feasible for large systems, like doped fullerenes. Temperature and dynamic effects can also be studied with ADFT-GIAO by analyzing the average or instantaneous values of any of the magnetic properties along a molecular dynamics (MD) trajectory (Goursot et al. 2013; Zuniga-Gutierrez et al. 2015). For these kinds of calculations, the magnetic property is sometimes calculated for hundreds of thousands of different structures. All these types of calculations are available in the deMon2k code.

For all magnetic properties studied so far, i.e., magnetic shielding tensors, magnetizability tensors, rotational g-tensors, and nuclear spin-rotation tensors, the ADFT-GIAO methodology was thoroughly validated against conventional DFT-GIAO approaches. In all cases, ADFT-GIAO yields practically indistinguishable results, with much less computational effort. As a benchmark example, the magnetic shielding tensors for all atoms in triphenylene ($C_{18}H_{12}$) and the corresponding magnetizability were calculated employing the PW91 exchange-correlation functional in combination with the aug-cc-pVXZ (X=D,T,Q) basis sets and the GEN-A2 auxiliary function set. Figure 5 shows the timings for the calculations performed with the conventional DFT-GIAO (black bars) and with the ADFT-GIAO (gray bars) methodologies. Clearly, the ADFT-GIAO methodology outperforms the conventional DFT-GIAO method. Note that the difference in computational performance increases rapidly with the number of basis functions. For the largest

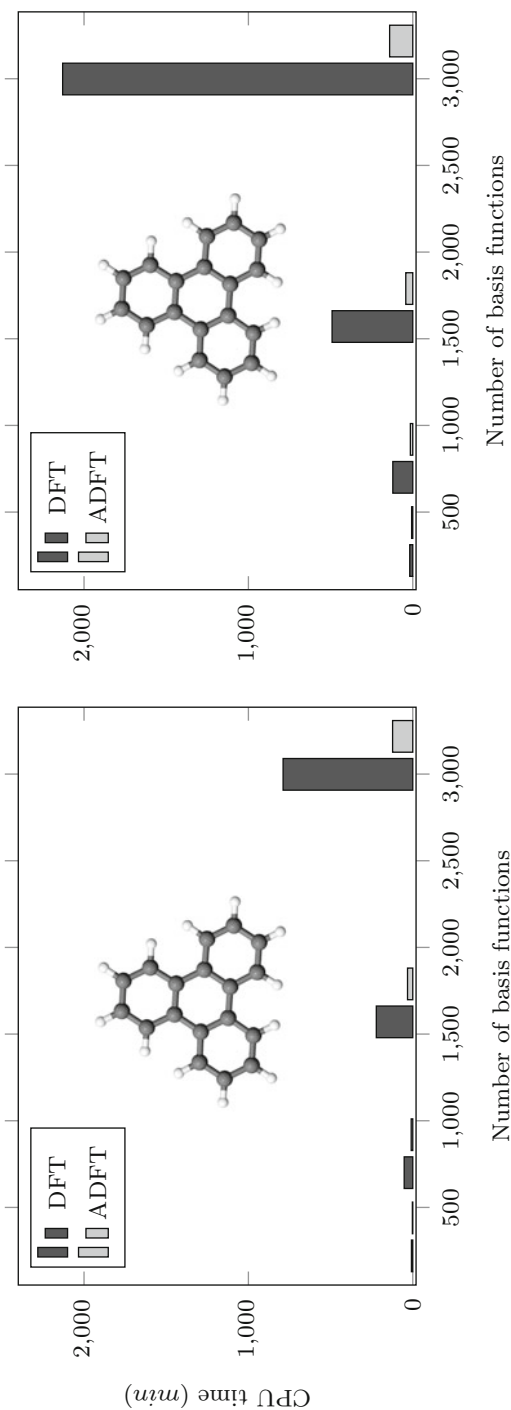


Fig. 5 Comparison of CPU timings between conventional DFT- and ADFT-GIAO methodologies for the calculation of the magnetic shielding (*left*) and magnetizability (*right*) tensors of the triphenylene

basis set, with 3108 basis functions, the ADFT-GIAO magnetic shielding and magnetizability calculations were 6 and 15 times faster than the corresponding DFT-GIAO calculations. Therefore, magnetic property calculations of nanosystems with hundreds or even thousands of atoms become feasible with the ADFT-GIAO methodology.

Figure 6 shows computational timings for magnetic properties in large systems: All magnetic shielding tensors for models of ZSM5 zeolites (Zuniga-Gutierrez et al. 2011) with one ($\text{Si}_{96}\text{O}_{200}\text{H}_{80}$), two ($\text{Si}_{192}\text{O}_{400}\text{H}_{128}$), and three ($\text{Si}_{288}\text{O}_{600}\text{H}_{176}$) cell units and the magnetizability tensor for models of a single-walled carbon nanotube ($\text{C}_{600}\text{H}_{40}$), a peapod ($\text{C}_{780}\text{H}_{40}$), and a double-walled carbon nanotube ($\text{C}_{952}\text{H}_{72}$) (Zuniga-Gutierrez et al. 2012). The calculations were performed with the LDA and PBE exchange-correlation functionals for magnetic shielding and magnetizability, respectively. In both cases, the DZVP basis set combined with the GEN-A2 auxiliary functions were used. The timings in Fig. 6 refer to parallel calculations with 8 Intel XeonTM cores with 2.4 GHz. The total time for the calculation of all magnetic shielding tensors in the largest ZSM5 model takes approximately 6 days. Note that the property calculation (gray bars) is computationally much less demanding than the SCF procedure (black bars).

Applications

Dynamics of Sodium Clusters

In this section, we review the results obtained from a systematic study of the temperature dependency of the polarizabilities of small sodium clusters at a reliable first-principles all-electron level of theory (Calaminici et al. 2007b; Gamboa et al. 2008).

Static polarizability measurements have been used for a long time to gain insight into the electronic structure of small clusters (Bonin and Kresin 1997; Knickelbein 2001, 2003, 2004; Bowlan et al. 2011). Because the polarizability is very sensitive to the distribution of the valence electron density, it can also be used as an indicator for chemical reactivity (Pearson 1973; Politzer 1987). Therefore, the study of the size dependency of the polarizability of simple clusters yields results of fundamental importance to chemistry and physics. A particular example resides in the polarizability studies on sodium clusters which were of paramount importance for the derivation of the jellium model (Knight et al. 1984). Today, this model is used with great success in cluster science (Bergeron et al. 2004, 2005). Several series of experimental data on sodium cluster polarizabilities are available in the literature (Molof et al. 1974a, b; Knight et al. 1985; Rayane et al. 1999; Tikhonov et al. 2001; Bowlan et al. 2011). In the pioneering work of Knight et al. (1984), the static polarizabilities of sodium clusters in a size range from 2 to 40 sodium atoms were studied. Later on, Rayane et al. (1999) repeated these measurements for a smaller size range from 2 to 22 sodium atoms. More recently, Tikhonov et al. (2001) measured the polarizability of selected sodium clusters up to 93 atoms. Just a few years ago, Bowlan et al. performed high precision measurements of

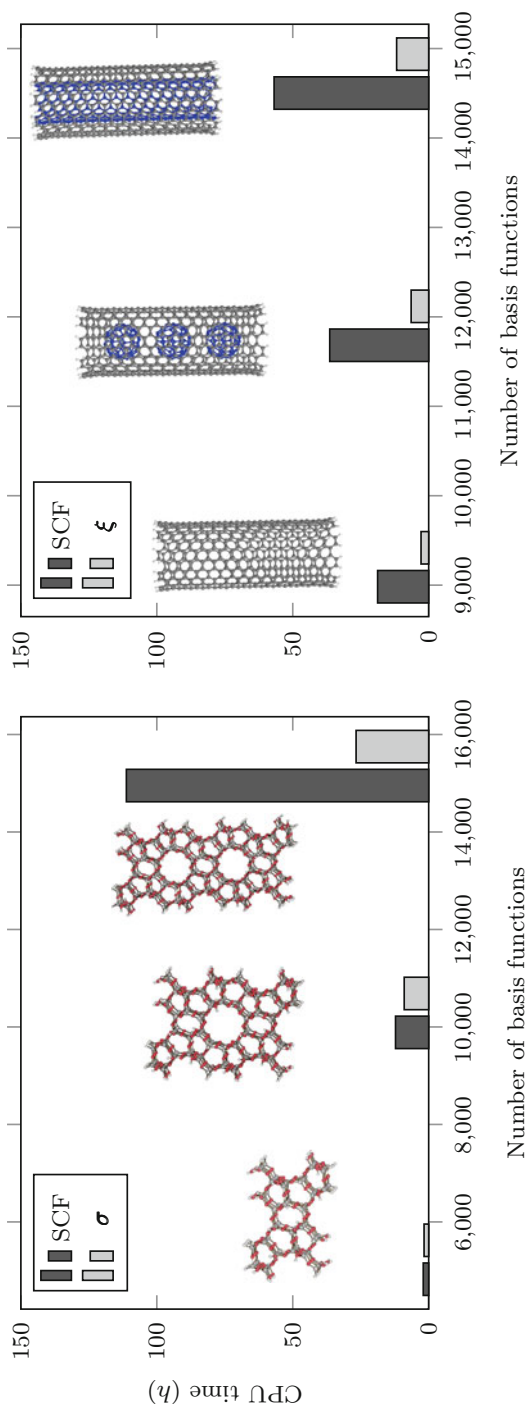


Fig. 6 Computational times for the calculation of magnetic shieldings (*left plot*) for ZSM5 models and the magnetizabilities (*right plot*) for models of carbon nanotubes. See text for details

cluster polarizabilities and dipole moments for sodium clusters containing less than 200 atoms. The overall agreement between these different experiments is quite satisfying. Nevertheless, discrepancies exist.

In Fig. 7, the experimental data for a series of sodium cluster polarizabilities are plotted together with our theoretical values. In particular, the pronounced oscillating behavior observed by Knight et al. up to the hexamer (Fig. 7a, dots) was not confirmed by the more recent study of Rayane and coworkers (Fig. 7a, squares). To emphasize the spread between these two experimental data sets, we have connected the data points in Fig. 7 by vertical lines. This figure also shows that for the larger sodium clusters with 7, 8, and 9 atoms excellent agreement between the reported data sets exist. The heptamer and octamer polarizabilities were also measured by Tikhonov et al. (2001) and are in good agreement with the depicted experimental data in Fig. 7, too.

In order to resolve the discrepancy in the measured polarizabilities of the smaller sodium clusters, many theoretical studies have been performed over the last two decades (see, e.g., Politzer 1987; Chandrakumar et al. 2004; Kümmel et al. 2000; Blundell et al. 2000; Kronik et al. 2000 and references therein). Most calculations employed density functional theory (Hohenberg and Kohn 1964; Kohn and Sham 1965), but wave-function-based studies are also available (Chandrakumar et al. 2004). A comparison of these theoretical studies reveals that the static polarizabilities of sodium clusters are severely underestimated at all reliable levels of theory. Over the last two decades, different corrections have been proposed to resolve this long-standing discrepancy between theory and experiment. More recently, it has been speculated that the mismatch between calculated and measured sodium cluster polarizabilities is due to finite temperature effects (Kümmel et al. 2000; Blundell et al. 2000; Kronik et al. 2000). In fact, this idea was already mentioned in the original experimental work (Knight et al. 1985). However, a systematic study of the temperature dependence of sodium cluster polarizabilities at a reliable first-principles all-electron level of theory was still absent until very recently. We have finally closed this gap by performing a temperature-dependent polarizability study of small sodium clusters by employing Born-Oppenheimer molecular dynamics (BOMD) simulations (Calaminici et al. 2007b; Gamboa et al. 2008).

The calculations were performed with the linear combination of Gaussian-type orbital density functional theory (LCGTO-DFT) deMon2k (Köster et al. 2011a) code. In Fig. 7, the crosses refer to all-electron polarizabilities calculated with the local density approximation (LDA) employing the exchange functional from Dirac (1930) in combination with the correlation functional proposed by Vosko, Wilk, and Nusair (VWN) (Vosko et al. 1980). The asterisks denote polarizabilities obtained with the gradient-corrected exchange-correlation functional proposed by Perdew, Burke, and Ernzerhof (PBE) (Perdew et al. 1996a).

The cluster structures were optimized at the corresponding level of theory employing a double-zeta valence polarization (DZVP) basis set (Godbout et al. 1992). For the polarizability calculations, a triple-zeta valence polarization (TZVP) basis set augmented with field-induced polarization (FIP) function was used (Calaminici et al. 1999). All calculations were performed in the framework of

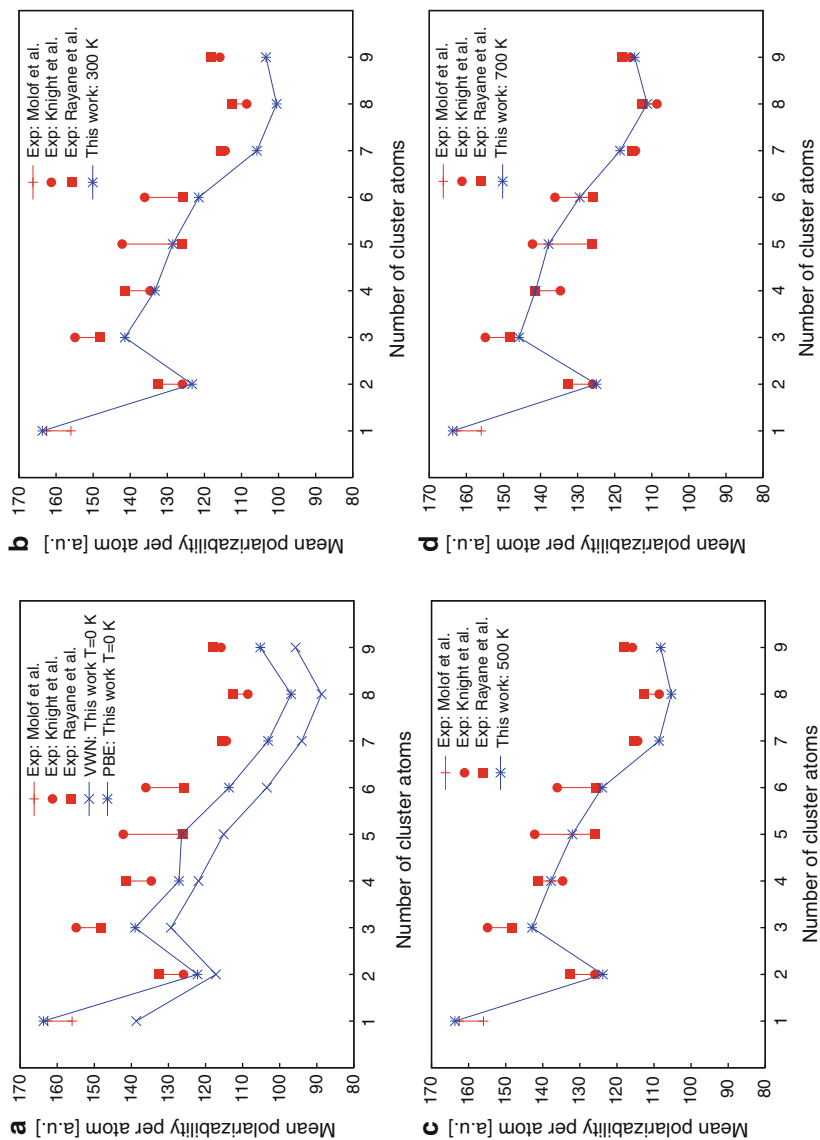


Fig. 7 Experimental and theoretical mean polarizabilities per atom [a.u.] of Na_n ($n=2-9$) clusters. The theoretical values are calculated at the VWN/TZVP-FIP/GEN-A2^{8*} (crosses) and PBE/TZVP-FIP/GEN-A2^{8*} (asterisks) level of theory. These values are connected in order to guide the eye. The individual graphs display calculated polarizabilities at 0 (a), 300 K (b), 500 K (c), and 700 K (d)

auxiliary density functional theory (ADFT) (Köster et al. 2004b) with A2 or GEN-A2* auxiliary function sets (Calaminici et al. 2007a). The latter was used in the analytical calculation of the cluster polarizabilities (Flores-Moreno and Köster 2008).

It is well established in the literature that calculated DFT polarizabilities at this level of theory differ by no more than 5 % from experiment (Calaminici et al. 1998). Figure 7a, however, shows that the calculated polarizabilities of the sodium clusters are not only considerably too low, but even at the level of the gradient-corrected PBE functional, differences of more than 10 % relative to the experimental values can occur. By and large, these results are confirmed by many other theoretical calculations.

In order to study the dynamics of small sodium clusters at finite temperatures, Born-Oppenheimer molecular dynamics (BOMD) calculations were performed at the above described PBE/DZVP/A2 level of theory. For each cluster, from the dimer to the nonamer, 18 trajectories were recorded in a temperature range from 50 to 900 K with intervals of 50 K. Each trajectory has a length of 220 ps and was recorded with a time step of 2 fs. Similar statistics have already been successfully applied to determine the melting temperatures of sodium clusters with LDA pseudo-potential DFT molecular dynamics (Chacko et al. 2005).

The temperature in the canonical BOMD simulation was controlled by a Nosé-Hoover chain thermostat (Nosé 1984; Hoover 1985; Martyna et al. 1992). In order to study the temperature dependency of the sodium cluster polarizabilities, the polarizability tensor α was calculated along the recorded trajectories. For this purpose, the first 20 ps of each trajectory were discarded, and α was then calculated in 100 fs time steps along the remaining 200 ps. Due to the computational demand of the analytical polarizability calculation along the BOMD trajectories, we employed the LDA kernel. Thus, the computational level for the calculation of the temperature-dependent part of the cluster polarizabilities was VWN/TZVP-FIP/GEN-A2*. The temperature-dependent mean sodium cluster polarizability was then calculated as

$$\bar{\alpha}(T) = \bar{\alpha}^{PBE}(0) + \delta\bar{\alpha}^{VWN}(T) \quad (105)$$

with

$$\bar{\alpha}(T) = \frac{1}{3} [\bar{\alpha}_{xx}(T) + \bar{\alpha}_{yy}(T) + \bar{\alpha}_{zz}(T)]. \quad (106)$$

This approximation assumes that the temperature dependency of $\bar{\alpha}(T)$, namely, $\delta\bar{\alpha}(T)$, is the same at PBE and VWN levels of theory. It should be remembered that the geometries are of course always determined from PBE BOMD calculations.

In Fig. 7, the calculated cluster polarizabilities at 0 (a), 300 (b), 500 (c), and 700 K (d) are depicted. As this figure shows, the individual cluster polarizabilities increase with temperature. Somewhere between 500 and 700 K, the calculated $\bar{\alpha}(T)$ per atom matches into the experimental data sets. In particular, the comparison of the calculated and experimental $\bar{\alpha}(T)$ per atom at 700 K for Na₇, Na₈, and Na₉, for which excellent agreement between the different experimental data sets exist,

is very satisfying. Moreover, it is interesting to note that the oscillating behavior of the $\bar{\alpha}$ per atom for the smaller clusters, which was observed in the original measurement (Knight et al. 1985) and also appears, less pronounced, in the $T=0$ K PBE calculations (Fig. 7a, asterisks), disappears at higher temperatures. Instead, the $\bar{\alpha}$ per atom decreases monotonically from Na_3 to Na_8 . Therefore, we can conclude that the finite temperature polarizabilities of small sodium clusters do not reflect individual molecular structures. Instead, they only reflect the shell closing at the dimer and octamer consistent with the jellium model. Our BOMD calculations also show that cluster fragmentations are not important for the cluster polarizabilities. Such fragmentations occur in our simulations above 800 K.

The change in the trend of the $\bar{\alpha}$ per atom for the small sodium clusters with increasing temperature is due to the different temperature dependencies of the individual cluster polarizabilities. In Fig. 8, the temperature dependencies of the individual cluster polarizabilities are presented. As this figure shows, for the smaller clusters up to the pentamer (Fig. 8a), the behavior of Na_4 is particular. Over a large temperature range up to 600 K the $\delta\bar{\alpha}(T)$ value for this cluster changes almost ideally linearly with the temperature. A closer analysis reveals that in this temperature range the D_{2h} rhombic structure of Na_4 rearranges only in the molecular plane. At 600 K and above, three-dimensional rearrangements occur. In this case, the temperature dependency of the polarizability reflects directly the dynamics of the cluster rearrangement. Figure 8a also shows that $\delta\bar{\alpha}(T)$ increases considerably faster with temperature for Na_4 than for Na_5 . As a result, the bump in the static $T=0$ K PBE polarizabilities at the pentamer (Fig. 7a) disappears in the finite temperature polarizabilities. For the larger clusters, we found very similar temperature dependencies of $\delta\bar{\alpha}(T)$ for Na_8 and Na_9 (Fig. 8b). The shape of our Na_8 temperature-dependent polarizability resembles previous studies with the extended Thomas-Fermi approximation (Blundell et al. 2000). The enlarged sodium hexamer polarizabilities in the temperature range between 200 and 600 K are due to the coexistence of two isomers, planar and pentagonal-pyramidal, as was already previously speculated in (Calaminici et al. 1999).

Summary

In this section, the results obtained by the first systematic study of the temperature dependency of the sodium cluster polarizabilities at a reliable first-principles all-electron level of theory were reviewed. The main results of this study are summarized as follows:

1. The calculated $\bar{\alpha}(T)$ per atom matches well the available experimental data sets at around 700 K;
2. The long-standing discrepancy between theory and experiment is resolved by inclusion of finite temperature effects in the electronic structure calculation;
3. The calculated finite temperature sodium cluster polarizabilities show characteristic minima at the dimer and octamer as expected from the jellium model;
4. Individual molecular structures besides these two are not resolved in the calculated finite temperature sodium cluster polarizabilities.

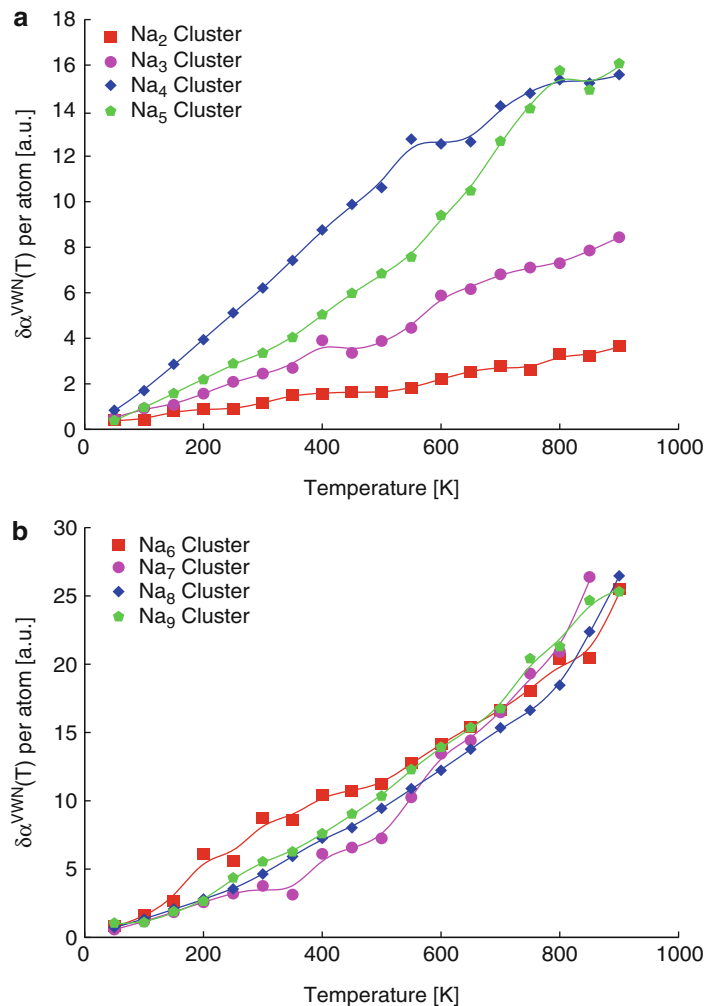


Fig. 8 Change of the mean polarizability per atom with temperature for Na_n clusters with $n = 2-5$ (a) and $n = 6-9$ (b). The calculations were performed at the VWN/TZVP-FIP/GEN-A2* level of theory

Structure of Zeolites

In this section, the calculated structural parameters and energetic properties of Na- and H-mordenites (MOR), using cluster models with more than 400 atoms, will be reviewed. These calculations (Dominguez-Soria et al. 2007, 2008) were performed in the framework of ADFT, using both the local (LDA) and the generalized gradient approximation (GGA), employing all-electron basis sets. For this study, the most populated T3, T4, and T1 Al sites have been investigated, using two different MOR

models, each containing two isolated Al sites. A detailed analysis of the structures, 3D contours of the molecular electrostatic potential (MEP), binding energies of Na cations and protons, and Brønsted O-H harmonic frequencies is discussed. If possible, comparison with available experimental results is made. The structural changes among Si/Al substitution as well as Na/H exchange are discussed as well.

Diffusion, adsorption, and reactivity of molecules within micro- or mesoporous materials are specifically related with the physicochemical properties of the material structures. It was originally underlined by Derouane et al. (1988) that the interactions of the host molecules with the material surfaces depend on the volume, shape, and topology of the cavities, which generate particular organizations of these molecules. The interrelationship between the porous materials and the host molecules has been referred to as “confinement” and attributed a large role in the selectivity and catalytic activity of zeolite materials, in particular, in acid-catalyzed reactions (Thibault et al. 1998; Anquetil et al. 1999; Smirnov et al. 1999).

Computer modeling based on ab initio techniques have been used in recent years to provide a better understanding of the physicochemical processes involved in the protonation of reactants. They lead to somewhat different conclusions, according to the methods applied and the reactions studied:

- (i) A confinement effect should originate from van der Waals interactions between the reactants and the zeolite framework (Rungsisirakun et al. 2005);
- (ii) Protonated products are stabilized by the long-range electrostatic effects of the framework (Vos et al. 2001; Demuth et al. 2001; Rozanska et al. 2003, 2005);
- (iii) Confinement acts through the stabilizing interaction of the framework oxygens on protonated transition states, which are also favored when matched with the zeolite cavities, van der Waals interactions being dominant when neutral species are considered (Vos et al. 2001; Demuth et al. 2001; Rozanska et al. 2003, 2005).

These results obtained with modern computational methods demonstrate that the concept of confinement is now widely invoked but not yet clearly quantified, differing with the level of accuracy and approximation. The relative influence of the local structure on the Brønsted acid strengths (large versus small cavities), of the solid-reactant dispersion interactions, and of electrostatic polarization effects on the reactants has not really been quantified yet. We believe that such a study necessitates first to set up an accurate and also efficient methodology in order to answer the following fundamental questions:

- (i) Is the Brønsted acid strength different in small and large cavities and what is the role of the local site geometry on this property?
- (ii) How dependent on the cavity size is the stabilization of the guest molecule through long-range dispersion and electrostatic polarization due to the solid framework?

We have focused our attention on the first question (Dominguez-Soria et al. 2007, 2008), studying the structure and the intrinsic properties of mordenite (MOR)

cationic sites, in particular sodium binding energies and acid strengths. For these studies, we proposed a new methodological approach based on very large model clusters, which can eventually be embedded in a classical environment, which we believe is a promising alternative to the idealized periodic representation of a zeolitic solid. This is particularly important for zeolites where the experimental Si/Al ratio is not very large, i.e., when the solid contains an Al distribution which can hardly be represented as periodic. In the next sections, we will compare our results with those reported experimentally for natural and synthetic zeolites. In section “Models and Methodology,” we discuss the zeolite models that we have chosen and the applied methodology. In section “Calculated Geometries and Properties,” we report the calculated geometries and properties and discuss them with respect to experiment and other theoretical results. The results obtained so far are summarized in section “Summary.”

Models and Methodology

Mordenites are natural and synthetic zeolites with Si/Al ratios of 4.3 to 6.0 in the former case and 5.0 to 12.0 in the latter (Passaglia 1975; Jacobs and Martens 1987). Synthetic mordenites are used for acidic catalysis. MOR catalysts are synthesized in the Na form followed by a mild treatment with NH_4Cl which leads to H-exchanged forms. The MOR structure can be described as composed of edge-sharing five-membered (5-m) rings of tetrahedra forming chains along the *c* crystallographic axis (Meier 1961). Their architectures comprise large mono-directional accessible 12-m ring channels of TO_4 tetrahedra where T stands for either Si or Al and small 8-m ring channels, which are interconnected through 8-m ring tubes. The topological symmetry of MORs is orthorhombic with space group *Cmcm* having in the unit cell four symmetrically independent tetrahedral sites, usually called T1, T2, T3, and T4. T1 and T2 sites connect four different rings while the T3 and T4 sites constitute the 4-m rings of the zeolite framework (Meier 1961). Since the first structural study of natural mordenite by Meier (1961), much work has been performed to solve several problems about the real symmetry of the solid due to the presence of Al tetrahedra and extra-framework cations, framework defects, and other structural distortions (Alberti et al. 1986; Alberti 1997). Having different Si/Al ratios, natural and synthetic zeolites also have different Al distribution patterns and slightly different T-O bond lengths (Schlenker et al. 1979).

Our strategy in this study for the determination of the structure and intrinsic properties of the catalyst has been to use clusters containing 120 tetrahedra, sufficiently large to enclose the main 12m-rings and the side pocket 8m-rings. These models include 2 unit cells along *a* and *b*, whereas the *2c* dimension has been cut at the middle of the second 12m-ring channel (see Fig. 9). The substitution of Si by Al has previously been studied extensively for natural and synthetic mordenites (Alberti et al. 1986; Alberti 1997; Schlenker et al. 1979; Ito and Saioto 1985; Simoncic and Armbruster 2004). Following these results based on crystal structure refinement (Alberti et al. 1986; Alberti 1997) and IR intensities of the OH stretching bands (Yang et al. 2001), it has been proposed that the Brønsted sites of MOR are related to an Al substitution with the probability $\text{T3} > \text{T4} > \text{T1} > \text{T2}$, associated with

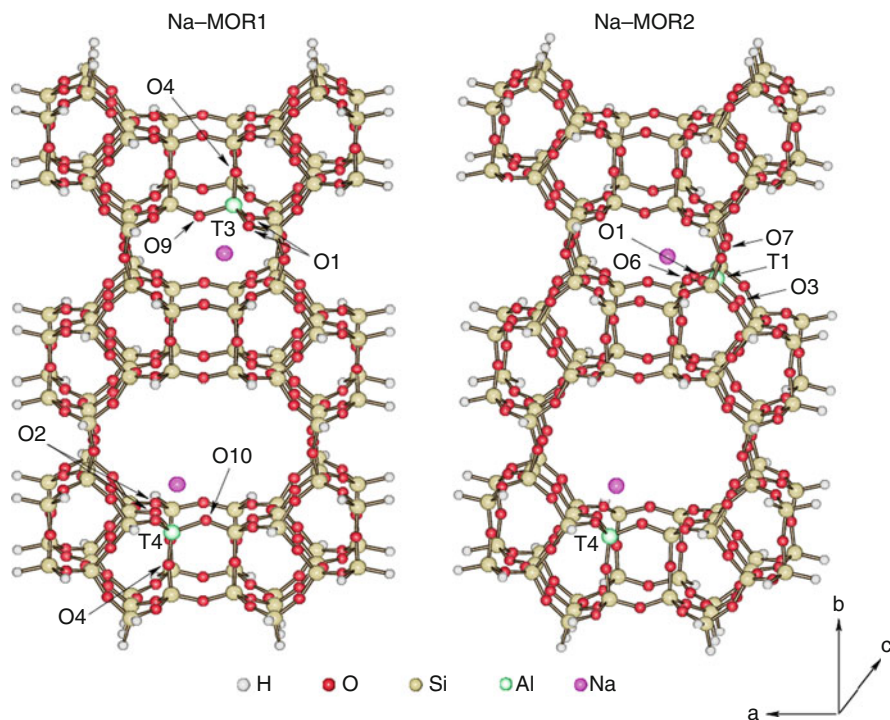


Fig. 9 Structures, T and O atoms legend of the studied mordenite models. The orientation of the three crystallographic axis a , b , and c is also given

a relative population of the OH sites which is probably larger in the 12-ring channel for natural zeolites (Alberti et al. 1986). The T4 site has also been proposed as the most probable candidate for Al substitution by Demuth et al. based on the relative stability of Al embedded clusters with Al substituted in turn at the 4 possible sites (Demuth et al. 2001). Due to the low probability of Al in T2 and the high probability of more Al sites in the main channel (Alberti 1997), we have selected two models with the distributions T4, T3 and T4, T1, associated with both Na cations (Na-MOR1, Na-MOR2) and protons (H-MOR1 and H-MOR2). The original models have been cut from a solid with the appropriate Al distribution, generated using the Cerius2 program (Cerius 2005) and terminated with hydrogens. In order to constrain the structure of the cluster to that of the solid, the coordinates of the terminal hydrogen atoms, positioned along the Si-O bonds, have been fixed during the subsequent optimizations, all other atomic coordinates being relaxed. Figure 9 illustrates the two models Na-MOR1 and Na-MOR2. In order to test the influence of the Al content on the mordenite geometrical parameters, we have also studied the same model without Al atoms, as will be discussed later.

For these studies, we have used an ADFT-based approach, using a linear combination of atomic orbitals, as implemented in the deMon2k program (Köster et al. 2011a). Geometry optimizations and molecular electrostatic potential (MEP)

calculations were performed at the local level using the correlation functional proposed by Vosko et al. (1980), whereas energetic properties (binding energies and proton affinities) have also been evaluated using the exchange-correlation functional of Perdew, Burke, and Ernzerhof (Perdew et al. 1996a). DFT-optimized double-zeta plus valence polarization (DZVP) basis sets were employed for all atoms (Godbout et al. 1992). For the fitting of the density, the A2 auxiliary function set was used (Godbout et al. 1992). The exchange-correlation potential was numerically integrated on an adaptive grid (Köster et al. 2004a). The grid accuracy was set to 10^{-5} a.u. in all calculations. The Coulomb energy was calculated by the variational fitting procedure proposed by Dunlap, Connolly, and Sabin (Dunlap et al. 1979; Mintmire and Dunlap 1982). A quasi-Newton method in internal redundant coordinates with analytical energy gradients was used for the structure optimization (Reveles and Köster 2004). The convergence was based on the Cartesian gradient and displacement vectors with a threshold of 10^{-4} and 10^{-3} a.u., respectively. For the geometry optimization of the mordenite structures, the parallel version of deMon2k (Calaminici et al. 2006; Geudtner et al. 2006) was used. Each mordenite model possesses more than 400 atoms with around 6000 orbital basis functions and 11,000 auxiliary functions. The calculations were performed on 6 or 8 2.4 GHz Intel Xeon CPUs. These nodes were connected with a Myrinet switch. The geometry optimization of such models takes about 13 days on 8 CPUs. In this time, roughly 25 optimization steps are performed. The cation binding energies (BE) to the framework have been calculated using the formula: $BE = E(\text{neutral model}) - E(\text{cation}) - E(\text{anionic model})$. The O-H stretching frequencies of the H-MOR1 model have been calculated within the harmonic approximation. This normal mode being uncoupled with all the other vibrational modes of the model, its frequency has been calculated using a partial Hessian analysis, as proposed by Li and Jensen (2002).

Calculated Geometries and Properties

General Structural Parameters of the Mordenite Models

Our purpose here was to simulate averaged bond length and bond angle data that can be compared with those provided by X-ray experiments performed on natural and synthetic mordenites with various Si/Al ratios. Comparison with available theoretical data was also presented (Dominguez-Soria et al. 2007). The main geometrical parameters (T-O bonds and T-O-T bond angles) of the fully siliceous and sodium mordenite models have been compared in Table 1 of Dominguez-Soria et al. (2007) with experimental data. The calculated values have been averaged over Na-MOR1 and Na-MOR2, namely, considering different Al distributions. This procedure mimics the experimental averaging over the non-distinguishable Si and Al positions. Among several experimental structures of natural mordenites (Alberti et al. 1986; Alberti 1997; Schlenker et al. 1979; Ito and Saioto 1985; Simoncic and Armbruster 2004), the structure obtained by Alberti et al. (1986) has been chosen for comparison together with the experimental structure of a natural and a synthetic Na-mordenite (Simoncic and Armbruster 2004).

As we have shown in Table 1 of Dominguez-Soria et al. (2007), the calculated bond lengths are very comparable with the experimental results, within 0.02 Å. We have noticed that this uncertainty is of the same order of magnitude as that obtained in previous theoretical work using the LDA level of theory and periodic boundary conditions (Demuth et al. 2001). The systematic lengthening of the T-O bonds in our calculations with respect to standard experimental bond distances could be interpreted in terms of a correction for thermal motions as given in Simoncic and Armbruster (2004). The average values of the T-O-T bond angles, presented in Table 2 of Dominguez-Soria et al. (2007), are in good agreement with experiment, within 4 degrees, which is also the average difference of our calculated bond angles with respect to those reported previously by Demuth et al. (2001). It is worth noting that the isotropic temperature factors provided by X-ray refinements for the zeolite oxygens are all very large (2.4 to 4.6 Å²) (Alberti et al. 1986; Simoncic and Armbruster 2004), leading to a substantial uncertainty on the oxygen positions. It was also shown that the averaged bond length and bond angle are rather insensitive to the Al content of the model. This allowed us to compare our models containing two Al sites with experimental mordenites which contain about 4 to 6 Al per unit cell.

Al sites of Na-MOR

Previous Data for Extra-Framework Cations

Experimentally, the positions of the extra-framework cations are not easily determined. Moreover, there is a clear difference between natural and synthetic mordenites, due to the different natures of their extra-framework cations, as it was demonstrated in the literature (Simoncic and Armbruster 2004). Whereas Ca and K cations predominate in natural mordenites, synthetic zeolites are grown using Na cations, leading, presumably, to different locations of the (Al, cation) pairs; although T3 remains slightly more populated than the other sites, the distribution of Al among the possible T sites is much more balanced according to the pattern reported by Simoncic and Armbruster (2004) (for an Al content of 6, one can expect 1.3 Al at T1 and T4, 0.8 at T2 and 2 at T3). Consequently, the Na cations compensating the Al charges are found distributed within two regions, i.e., the side pocket and the main channel. Due to the presence of water, about 60 % of the cations cannot be precisely determined, presumably being located mainly in the main channel (Simoncic and Armbruster 2004). The remaining cations are found either bonded to three oxygens around T3 in the side pocket or pushed downward into the 8m-ring tube.

Since our study concerns Na-MOR models, we focus on comparing our results with the experimental data provided for synthetic Na-mordenites. Previous theoretical work on Na-mordenite used periodic models involving 8 (Al, Na) pairs per unit cell, and results were compared with the natural zeolite data available (Demuth et al. 2001). The choice of occupying only T3 and T4 sites with equal populations was made, leading then to a strong concentration of cations in each cavity, related to the high symmetry of the solid. Therefore, it is hardly possible to make a direct comparison with this study, concerning the Na positions with respect to the framework oxygens.

Analysis of the Properties of the Cationic Sites in Na-MOR1 and Na-MOR2

The substitution of one Si by Al in the framework generates one negative charge, which has to be compensated by the presence of a counterion. The framework negative charge is distributed among the framework oxygens, more particularly to the four oxygens adjacent to Al. The location of the cation thus depends on the local geometry around the T site for large Si/Al ratios, i.e., when the Al sites can be considered isolated. In these conditions, it is well established that smaller Al-O-Si angles are generally associated with larger electrostatic potentials, i.e., larger attractive interactions with an incoming positive charge (Goursot et al. 1998). In order to locate the most probable positions of the counterions, we have used the well-known strategy of approximating the cation position by searching the largest electrostatic interaction of the system density with a positive point charge. To do so, the molecular electrostatic potential (MEP) was calculated around the three AlO_4 T4, T3, and T1 tetrahedra. The contour plots of these MEPs are illustrated in Figure 2a, 2b, and 2c of Dominguez-Soria et al. (2007), respectively. These MEP values have been calculated for the optimized structures of the corresponding T anions, at the T4 and T3 sites in Na-MOR1 and at the T4 and T1 sites in Na-MOR2, respectively. As expected, the MEP wells in Na-MOR1 surround the two O2 and O10 in the main channel at T4 and the two O1 and O9 at T3 in the side pocket. The lowest MEP minimum at T4 is located between the two O2 and O10 (-4.13 eV). The lowest minimum at T3 is located close to O9 (-4.22 eV). In Na-MOR2, the MEPs at T4 are identical to the ones calculated for T4 in Na-MOR1, showing that the two Al sites are independent. Interestingly, the T1 site of Na-MOR2 belongs to both the main channel and the side pocket. The O1 and O6 adjacent oxygens are in the side pocket, whereas O3 and O7 are located in the main channel (Fig. 9). The MEP well at T1 is essentially located within the side pocket, with the lowest minimum close to O6 (-3.94 eV). It is worthwhile to underline that the MEP values at O7 and O3, i.e., in the main channel, are equal to -3.60 and -3.10 eV, respectively, showing that a compensating Na^+ or H^+ at those oxygens is unfavored with respect to O6. This result indicates that the counterion at T1 will be more stable within the side pocket than in the main channel, at least after dehydration. This conclusion is also valid for protons (Dominguez-Soria et al. 2007).

Si substitution by Al leads, as expected, to lengthening the T-O bonds and decreasing the T-O-Si angles. The average elongation of the T-O bonds reaches 0.115 for T1, 0.124 for T3, and 0.123 Å for T4. The corresponding T-O-Si averaged decrease is found to be 3.3 degrees, 7.2 degrees, and 8.3 degrees for T1, T3, and T4, respectively. These changes are in a reasonable range, with a lower variation for the T1 site, due to its more rigid environment. The optimized structures of the Na^+ sites were reported in Table 3 of Dominguez-Soria et al. (2007), as well as the Na binding energies to the framework. From the reported Na-O distances, the cation can be considered as bi-coordinated at T4 and T1 and mono-coordinated at T3. This result correlates with the calculated positions of the MEP minima, as shown in Figure 2 of Dominguez-Soria et al. (2007). Interestingly we noticed that the shortest NaO distances correspond to the smallest Al-O-Si angles (smaller Al-O-Si angles correlate with more electron density available for cations in the O lone pair due to better hybridization conditions) (Goursot et al. 1998). This result can

be considered as the physical reason why the cation at T1 is located inside the side pocket. Moreover, this finding allows us to anticipate the preferred oxygen for protonation, i.e., O10 at T4, O9 at T3, and O6 at T1.

The calculated Na binding energies follow the ordering $T3 > T4 \approx T1$, which is very comparable with that indicated by the MEPs. This result indicates that electrostatics governs the cation – framework binding. Moreover, we noticed that the use of the PBE GGA exchange-correlation functionals does not modify the trend established on the basis of LDA calculations. For sites T1, T3, and T4, it can be pointed out that the ordering of the Na binding energies reproduces the (Na, Al) population pattern, reported for synthetic Na-MOR. This result supports the suggestion of Simoncic et al. that there is a synergistic effect between cations and (Al, Si) order during the growth of the solid (Simoncic and Armbruster 2004). Finally, we remark that the binding properties found at T1 and T3 are different although both sites belong to the side pocket. This is related to a sum of different effects, the predominant one being most probably the global electrostatic potential felt by the cation. For the same reason, T1 and T4 present comparable binding properties, despite their different locations in the zeolite framework. It is important here to stress that these conclusions are made on the basis of the existence of two independent Al sites. If several (Al, Na) pairs are present within the side pockets, the distribution of the cations would most probably be governed by their mutual repulsion (Campana et al. 1997).

For isolated sites, we have found that both T1 and T3 compensating cations are located in the side pocket area. In synthetic mordenites, at least two sites in side pockets are substituted by Al (≈ 2.5). Since the binding of Na^+ at T3 is favored with respect to that at T1, the three short Na-O bonds indicated by experiment (to O1, O1' and O9) can then be related to our calculated values at T3. In the presence of a second Na^+ to compensate Al in T1, one can expect that the mutual repulsion of the two cations will modify their position. Taking into account their relative stability, the experimental evidence that the other cation is pushed downward to the main cavity (Simoncic and Armbruster 2004) is supported by our results. Moreover we could argue further about Na-MOR with 6 Al per unit cell: this Al content would imply the presence of three (Na, Al) pairs in the side pockets (presumably two at T3 and one at T1). In this situation, one may expect that only two cations will remain inside the pocket, whereas the third one will be pushed outside due to cation-cation repulsion (this argument is verified by classical molecular mechanics simulations using Cerius2 associated with the cvff-aug-ionic force field (Demuth et al. 2001)).

Al Sites of H-MOR

As for the Na-MOR models, the protonated structures have been optimized, the original proton positions being those of the respective lowest MEP minima. At T4, the MEP well is situated at equivalent distances from the two O2 and O10. Whereas the Na^+ cation at T4 was found at equal distances from O10 and O2, the optimized position of the proton corresponds to a bonding to O2. In a previous theoretical study of the relative energies of various H-MOR models, the O2 protonation was also found to be more favorable than the O10 protonation (Li and Jensen 2002).

Consistency with the mentioned relationship between Al-O-Si bond angle values and cation binding energies would lead one to expect comparable proton affinities at O2 and O10. At T1 and T3, the proton attaches to O6 and O9, respectively. Moreover, the O2-H and O6-H bonds are pointing toward the center of the 8m-ring which contains the T4 and T1 site, respectively, these bonds lying in the 8m-ring planes. The proton at T4 can thus be reached by a molecule being in the main channel as well as coming from the side pocket. In contrast, the O9-H bond at T3 is pointing toward the center of the side pocket.

We could see that the Al-O bonds which do not connect to a proton are only little changed with respect to those of the Na-MOR models, with a maximum elongation of about 0.05 Å. As expected, the Al-O(H) bonds are elongated by about 0.12–0.16 Å with respect to the same bonds in the Na-MOR models. This effect is due to the covalent character of the OH bond, which induces the weakening of the Al-O bonds, whereas the Na⁺ – oxygen interactions are weaker and essentially electrostatic. The presence of the sodium cation, bonded to two O (T4, T1) or one O (T3), had thus much less impact on the Si-O-Al framework than that of the proton.

In contrast with the Na binding energies, we found that the proton affinity values are more sensitive to the inclusion of the GGA correction, leading to increased values with respect to LDA and to nonsignificant differences among the three acid sites (less than 0.08 eV) if one takes into account all methodological errors. From these values, we conclude that the sites belonging to the main channel and to the side pocket display similar proton affinities, showing thus that the strength of the OH bond is not dependent on the local structure of the studied sites. It is worth noting that proton affinities are not related to relative energies of models with different H sites, which are reported in previous theoretical work (Demuth et al. 2001), relative total energies being essentially related with relative Al site stabilities.

In order to complete the description of the OH bond strength, the vibrational harmonic frequencies of O2-H and O9-H in the H-MOR1 model have been calculated. The values are 3667 cm⁻¹ for O9-H at T3 and 3659 cm⁻¹ for O2-H at T4, with IR intensities of 174 and 182 km/mol, respectively. These frequencies, not corrected by any scaling factor, can be considered similar within the theoretical error bar. This result is consistent with the corresponding proton affinity found for the proton at O9 (T3) with respect to O2 (T4). They are shifted toward higher energies by about 50 cm⁻¹ with respect to the reported infrared (IR) experimental asymmetric band at about 3610 cm⁻¹. The comparison of our results with X-ray and IR experimental data allows us to suggest that the O9H stretching vibration (T3 site) is not responsible for the observed low-energy component at 3585 cm⁻¹.

Due to the similarity of the model and method used for both vibrations, it is most probable that the lower energy vibration is related to another Brønsted hydroxyl originating from:

- (i) Proton migration, after dehydration, to another oxygen, especially if both T1 and T3 sites are aluminated, leading to different local interactions;

- (ii) The presence of two T3 and one T1 aluminated sites in the side pocket (Simoncic and Armbruster 2004) will induce different MEP surfaces and different electrostatics;
- (iii) The presence of a (100) defect layer (Simoncic and Armbruster 2004) that induces geometry changes in the side pockets which are not reproduced in our models.

We can thus argue that our models with isolated Al sites ($\text{Si}/\text{Al}=23$) show that main channel and side pocket hydroxyls are energetically not different because of their framework environment. Their different energetical behaviors are most probably generated by the different content of framework Al at the 12-membered ring channels and in the side pockets obtained after crystallization.

Summary

In this work, Na and protonated models of a mordenite zeolite, including Al in T1, T3, and T4 crystallographic sites, were studied using all-electron DFT calculations. From the analysis of the obtained results, the following conclusions can be drawn:

- (i) A good agreement with experimental bond length and bond angle data reported for synthetic Na-mordenites has been obtained;
- (ii) The binding energies of the Na cation follow the same ordering as the populations of Al T sites, as derived from X-ray measurements on synthetic Na-mordenites; this result shows the existence of a synergistic effect between cations and (Al, Si) order during the growth of the solid, as suggested experimentally;
- (iii) The compensating cation at the T1 site is found to be more stable in the side pocket than in the main channel, which questions the assignments of the T1-associated proton to a main channel location;
- (iv) The calculated proton affinities at T1, T3, and T4 sites are equivalent, indicating that these Brønsted sites have similar acid strengths. One can thus infer that the different acidic behaviors at the different sites (Marie et al. 2000, 2004) do not originate from different acid strengths of these OH sites. This leads to the suggestion that the MOR acidic properties are more related to the electronic structure of the base than to effects of the solid framework.

Thus, for the evaluation of differences in local acidity, the presence of the associated base (CO, CH₃CN, NH₃, etc. . .) has to be taken into account. Further studies in this direction have been published by our laboratories (Dominguez-Soria et al. 2008, 2011).

Stability of Giant Fullerenes

Fullerenes are carbon nanostructures formed by the closing of a graphitic sheet with the needed curvature supplied by intersecting, among a given number of

graphitic hexagons, of 12 pentagons (Kadish and Ruoff 2007; Andreoni 2007). These carbon aggregates have been experimentally known for more than 20 years (Kroto et al. 1985) and, consequently, a large number of works, experimental as well as theoretical, focused on this subject (see, e.g., Seifert et al. 1996; Cioslowski 1995; Boltalina et al. 2000; Bühl and Hirsch 2001 and references therein).

One main reason for the great interest in the study of fullerenes is certainly to be found in their particularly appealing geometrical form. The best-known fullerene is the so-called buckminsterfullerene that contains 60 carbon atoms (C_{60}) and is composed of 12 pentagonal carbon rings located around the vertexes of an icosahedron and 20 hexagonal carbon rings at the centers of icosahedral faces (Kroto et al. 1985).

Larger fullerenes that have an icosahedral symmetry can be constructed (Kroto and McKay 1988; Itoh et al. 1996) as well. These clusters, known as *giant fullerenes*, can be thought of as cutout pieces of graphene that are folded into their final shapes (icosahedrons). This kind of procedure generates 12 pentagonal carbon rings situated around vertices of an icosahedron, while all other carbon rings are hexagonal. Giant or large fullerenes have been the subject of different theoretical studies in the last years. We direct the interested reader to Dunlap et al. (1991), York et al. (1994), Bakowies et al. (1995), Scuseria (1995), Xu and Scuseria (1996), Scuseria (1996), Haddon et al. (1997), Bates and Scuseria (1998), Heggie et al. (1998), Geudtner et al. (2006), Dunlap and Zope (2006), Shao et al. (2006, 2007), Zope et al. (2008) and references therein.

Most of these studies were focused either on understanding if the shape of these clusters is spherical or faceted (York et al. 1994; Bakowies et al. 1995; Scuseria 1995; Xu and Scuseria 1996; Scuseria 1996; Bates and Scuseria 1998; Geudtner et al. 2006), on calculating their response properties (Zope et al. 2008) or testing new algorithms developed for the investigation of large systems (Dunlap and Zope 2006; Geudtner et al. 2006). Most previous first-principles theoretical studies of large fullerenes have been performed either at the Hartree-Fock level of theory using symmetry restrictions and relatively small basis sets or employing analytic density functional theory (York et al. 1994; Bakowies et al. 1995; Scuseria 1995; Xu and Scuseria 1996; Scuseria 1996; Haddon et al. 1997; Bates and Scuseria 1998; Heggie et al. 1998; Dunlap and Zope 2006; Shao et al. 2006, 2007; Zope et al. 2008).

In a recent study, we have performed state-of-the-art calculations on the large C_{180} , C_{240} , C_{320} and C_{540} fullerenes by employing the linear combination of Gaussian-type orbital density functional theory (LCGTO-DFT) approach (Calaminici et al. 2009). The structures of these clusters were fully optimized without any symmetry constraints. This work represents the first systematic study on large fullerenes based on non-symmetry adapted first-principles calculations, and it demonstrates the capability of ADFT for energy calculations and structure optimizations of large-scale structures without any symmetry constraint.

In the next two sections, the computational details will be presented, and the most important results we have obtained in terms of structural changes, of the evolution of the bond lengths, and of the calculated binding energies will be reviewed.

Computational Details

All calculations were performed using the DFT program deMon2k (Köster et al. 2011a). The exchange-correlation potential was numerically integrated on an adaptive grid (Köster et al. 2004a). The grid accuracy was set to 10^{-5} a.u. in all calculations. The Coulomb energy was calculated by the variational fitting procedure proposed by Dunlap, Connolly, and Sabin (Dunlap et al. 1979; Mintmire and Dunlap 1982). The calculation of the exchange-correlation energy was performed with the auxiliary function density (Köster et al. 2004b), i.e., auxiliary density functional theory was used. The structure optimizations were performed with the local density approximation (LDA) employing the Dirac exchange functional (Dirac 1930) in combination with the correlation functional from Vosko, Wilk, and Nusair (VWN) (Vosko et al. 1980). DFT-optimized double-zeta plus valence polarization (DZVP) all-electron basis sets optimized for local functionals (Godbout et al. 1992) were employed. For the structure optimization, a quasi-Newton method in internal redundant coordinates with analytic energy gradients was used (Reveles and Köster 2004). The geometry optimizations were performed using the parallel version of the deMon2k code (Calaminici et al. 2006; Geudtner et al. 2006). The convergence was based on the Cartesian gradient and displacement vectors with a threshold of 10^{-4} and 10^{-3} a.u., respectively.

The diamond and graphene calculations were performed in the same theoretical framework using the cyclic cluster model (CCM) (Janetzko et al. 2008). The obtained energies using the CCM are in the range of other calculated cohesive energies for graphene (Trickey et al. 1992; Dunlap and Boettger 1996). Because the fullerene and graphene calculations are performed within the same theoretical framework, the relative energy differences found here are reliable.

Results and Discussion

The DFT-optimized singlet structures of C_{180} , C_{240} , C_{320} , and C_{540} are depicted in Fig. 10. These structures have been fully optimized at the all-electron level using DZVP basis sets in combination with the VWN functional. A long-standing discussion in the literature addresses the question of whether giant fullerenes prefer a faceted or a spherical shape. This question was raised considering pictures obtained by transmission electron microscopy (TEM) that have shown evidence of possible spheroidal structures in concentric carbon particles (Iijima 1980; Ugarte 1992, 1995). Using a divide-and-conquer method for density functional calculations, the structure and stability of C_{240} were studied, and the most stable structure was claimed to be highly spherical (York et al. 1994). However, this result was not confirmed by any successive theoretical work that have clearly shown evidence of a faceted shape for this fullerene (Dunlap et al. 1991; Bakowies et al. 1995; Scuseria 1995, 1996; Xu and Scuseria 1996; Haddon et al. 1997; Bates and Scuseria 1998). Depending on the viewing axis, simulated TEM of icosahedral fullerenes can provide either images with spherical or with faceted shapes (Scuseria 1995). In addition, an explanation of why experimental results showed rounder shapes for large fullerenes was also given (Scuseria 1995).

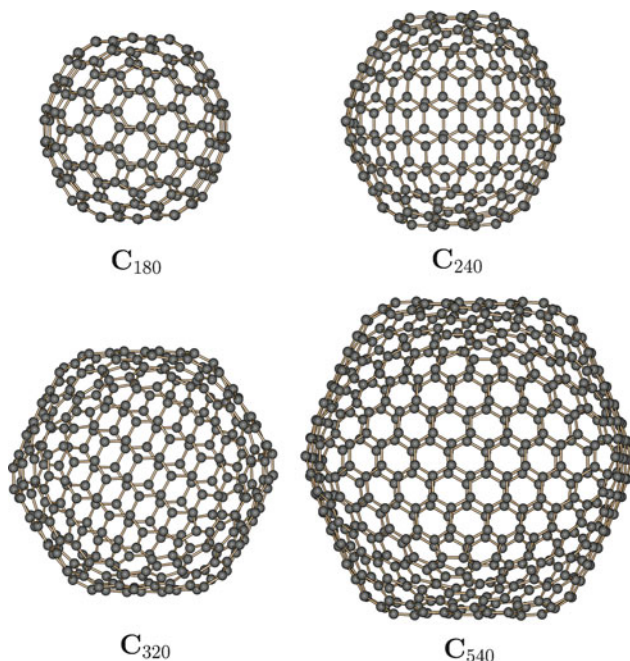


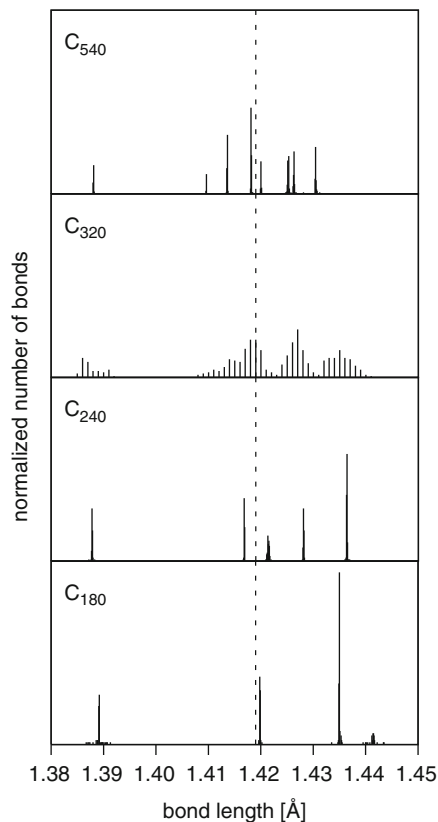
Fig. 10 Optimized structure of C_{180} , C_{240} , C_{320} , and C_{540} fullerenes. The calculations have been performed with the VWN functional in combination with DZVP basis sets

As Fig. 10 shows, our first-principles-based structure optimizations predict that larger fullerenes, C_{240} , C_{320} , and C_{540} , prefer a faceted shape. Moreover, even for the smallest fullerene here studied, C_{180} , there is clear evidence that the faceted shape is preferred over a spherical shape if first-principles all-electron optimizations without any symmetry restriction are performed (see Fig. 10). Details about the timing of these calculations are reported in Geudtner et al. (2006) where it is shown that the optimization of such large fullerenes is feasible within a few days on a parallel architecture with 32 to 64 cores.

We noticed that the obtained results are in agreement with most of the previous reported theoretical studies (Dunlap et al. 1991; Bakowies et al. 1995; Scuseria 1995, 1996; Xu and Scuseria 1996; Haddon et al. 1997; Bates and Scuseria 1998).

In order to gain more insight into the structural changes of these systems as the number of carbon atoms increases, a detailed analysis of the bond length evolution was performed. In Fig. 11, the normalized number of bonds for C_{180} , C_{240} , C_{320} , and C_{540} are plotted versus the bond length. The dashed line at 1.419 Å represents the graphene bond length obtained from the periodic (CCM) deMon2k calculation. Most obvious from this figure is the difference of C_{320} to all other fullerenes. In fact, whereas usually a discrete distribution of bond lengths is found, in C_{320} , a wide, in some ranges almost continuous, bond length distribution is observed. This clearly indicates a break in the expected high symmetry of the system. Our studies

Fig. 11 Normalized number of bonds for C_{180} , C_{240} , C_{320} , and C_{540} versus the bond lengths (in Å)



show that the C_{320} fullerene possesses a ground-state potential energy surface (PES) of higher multiplicity, most likely either triplet or quintet. Further test calculations indicated that for this fullerene also in the cases of triplet and quintet PESs the continuous bond length distribution observed for the singlet PES persists. Of course, only non-symmetry adapted optimizations can lead to such a result. To the best of our knowledge, this symmetry breaking in larger fullerenes due to their electronic structure has never been observed in previous calculations. As Fig. 10 shows, the observed symmetry breaking does not alter the global shape of the giant fullerene.

For the other systems, C_{180} , C_{240} , and C_{540} , the expected discrete bond length distribution is obtained, indicating that the symmetry of the electronic structures matches with the expected geometrical symmetry. In these systems, the number of different bond lengths increases with system size, and an accumulation of bond lengths around the graphene bond length is observed (Fig. 11).

More surprising is the trend that the longest bond length in the cluster shortens with increasing cluster size. This indicates that delocalization increases with cluster size despite the global building pattern, i.e., the appearance of 12 pentagons.

With the aim of guiding future desirable experiments on large fullerenes and to gain more information about their stability, we have also explored the behavior of the binding energy of the studied fullerenes with increasing fullerene size. The results of the uncorrected binding energy (in eV) per carbon atom obtained with the VWN functional have been illustrated in Figure 3 of Calaminici et al. (2009) showing that the binding energy increases monotonically with the increase in the number of carbon atoms. This indicates that the large fullerenes become more and more stable with increasing size. However, the increase turned out to be very moderate. We have also included the basis set superposition error (BSSE) in the calculation of the binding energies. The inclusion of the BSSE decreases the calculated binding energies of each studied fullerene by 0.02 eV but without altering the trend (Calaminici et al. 2009).

From the comparison of the CCM calculations, we found that the calculated binding energy of C_{540} of 8.75 eV is very close to the cohesive energy of diamond (8.78 eV). This result indicates that C_{540} has a similar binding energy to diamond which fuels the hope that such giant fullerenes could indeed be prepared. However, the binding energy of even the largest studied fullerene, C_{540} , is still far away from the corresponding value in graphene which was calculated to be 8.91 eV (Janetzko et al. 2008).

Summary

In this section, the results obtained from state-of-the-art density functional theory calculations performed on large fullerenes, such as C_{180} , C_{240} , C_{320} , and C_{540} , have been reviewed. The study was carried out with all- electron basis sets, and all structures were fully optimized without any symmetry restriction. The results obtained can be summarized as follows:

- (i) This work confirms that for all large fullerenes studied here, a faceted shape is preferred over the spherical shape;
- (ii) The analysis of the bond length evolution reveals for C_{320} a qualitatively different pattern than for the other fullerenes. The most likely explanation for this difference is a symmetry breaking in the electronic structure of this large fullerene;
- (iii) The shortening of the longest bond length with increasing cluster size indicates that delocalization increases with cluster size;
- (iv) The calculated binding energies are in the range of diamond but considerably below the graphene value. Thus, even giant fullerenes as those reviewed in this section are only metastable.

Transition State Search in Metal Clusters

Generally, in cluster studies, the properties of interest are investigated considering the lowest minimum structure. However, the study of the full potential energy surface (PES) landscape is very important if one aims to understand how for a given

finite system a chemical reaction may occur. A chemical reaction can be described in terms of an energy profile picture with which aspects of the mechanisms, structures, and energies of the reaction are explained. A way to characterize a PES is by the location of its relevant critical points (CPs). In general, stationary points of a function that depends on N variables are reached when the gradient of the given function vanishes. For a multidimensional function such as the PES, several different kinds of stationary points usually exist. These are minima, maxima, and saddle stationary points. These CPs are characterized by the eigenvalue spectrum of the corresponding Hessian matrix. A minimum CP possesses only positive eigenvalues that correspond to positive curvatures, in all principal directions. A maximum possesses only negative eigenvalues that correspond to negative curvatures in all principal directions. Saddle points of different order can be present on a PES, too. A transition state is a saddle point of first order, i.e., it possesses a negative curvature in one principle direction and positive curvatures in all other directions. From a chemical point of view, only minima and first-order saddle points are CPs of interest on a PES because they correspond to stable isomers and to transition states, respectively.

For a molecular system, the reaction coordinate is defined by the minimum energy path on the PES that connects the energy minimum of the reactant with that of the product. For an elementary reaction step, this path has only one maximum which is a first-order saddle point and named the transition state. For metal clusters, the PES of even very simple system consisting of only a small number of atoms can possess a large number of local minimum structures. As a result, a complex network of rearrangement reactions exists in these systems. To unravel such networks, the systematic location of transition states is mandatory. We applied the hierarchical transition state search algorithm (Campo and Köster 2008) for the first time to rearrangement reactions of pure metal clusters, in particular selected sodium clusters. These systems are very challenging because neither the transition states nor the reaction coordinates are chemically intuitive. Therefore, it is usually not guaranteed that the two initial minima structures are indeed connected by an elementary rearrangement step free of intermediates. To this end, we performed intrinsic reaction coordinate (IRC) calculations starting from the located transition states. By the combination of the hierarchical transition state search algorithm with the IRC calculations, reaction intermediates can be found. In this way, complex rearrangement networks can be completely unraveled.

Computational Details

All calculations were performed using the linear combination of Gaussian-type orbital density functional theory (LCGTO-DFT) deMon2k program (Köster et al. 2011a; Geudtner et al. 2011). The Coulomb energy was calculated by the variational fitting procedure proposed by Dunlap, Connolly, and Sabin (Dunlap et al. 1979; Mintmire and Dunlap 1982). The auxiliary density was expanded in primitive Hermite Gaussian functions using the GEN-A2 auxiliary function set (Calaminici et al. 2007a). The exchange-correlation functional was evaluated with this auxiliary density, i.e., the auxiliary density functional theory (ADFT) method was

used (Köster et al. 2004b). The exchange-correlation energy and potential were numerically integrated on an adaptive grid (Krack and Köster 1998; Köster et al. 2004a). The grid accuracy was set to 10^{-5} a.u. in all calculations. The structure optimization and the frequency analysis of the clusters were performed employing the gradient corrected exchange-correlation functional proposed by Perdew, Burke, and Ernzerhof (PBE) (Perdew et al. 1996a) in combination with DFT-optimized double-zeta valence plus polarization (DZVP) basis sets (Godbout et al. 1992).

In order to determine the geometry of the ground state and of corresponding low-lying states of the studied clusters, a very extensive search was performed and several dozens of initial configurations were studied. In this search, structures previously reported by other authors were included as well as many other initial structures which were extracted from BOMD trajectories recorded at a temperature of 250 K. The BOMD trajectories were generated in the canonical ensemble employing a Nosé-Hoover chain thermostat (Nosé 1984; Hoover 1985; Martyna et al. 1992; Gamboa et al. 2010) with seven thermostats in the chain and a coupling frequency of 200 cm^{-1} . For all structures, a full geometry optimization without any symmetry restriction was performed using a quasi-Newton optimization method in delocalized internal coordinates (Reveles and Köster 2004). The convergence was based on the Cartesian gradient and displacement vectors with a threshold of 10^{-4} and 10^{-3} a.u., respectively.

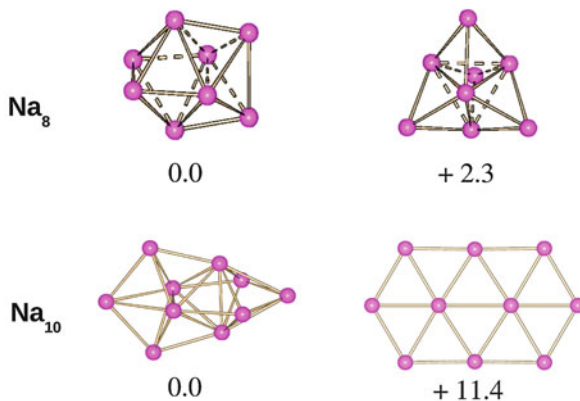
The obtained minima were characterized by frequency analysis. The second derivatives were calculated by numerical differentiation (two-point finite difference) of the analytic energy gradients using a displacement of 0.001 a.u. from the optimized geometry for all $3N$ coordinates. The harmonic frequencies were obtained by diagonalizing the mass-weighted Cartesian force constant matrix.

As already discussed, the transition states were located with the hierarchical transition state finder as implemented in the deMon2k code (Campo and Köster 2008; Geudtner et al. 2011). The only input information consisted of the minimum structures of the sodium clusters involved in the rearrangement. Starting from the obtained transition states, the IRC was calculated, using the approach from Gonzalez and Schlegel (Gonzalez and Schlegel 1989), either to find potential intermediates or to calculate the minimum energy path that connects the two starting minimum structures. Both the IRC and the uphill trust region optimization in the hierarchical transition state search were initialized by Hessian calculations in order to provide the correct eigenvalue spectrum.

Showcase Results

In this section, the results of selected transition state searches in Na_8 and Na_{10} clusters are discussed. The two lowest minimum structures of Na_8 are depicted in Fig. 12 (top), as well as the lowest energy structure and a planar high-energy stable isomer for the Na_{10} cluster (bottom). All minimum structures are in their lowest multiplicity. The relative energies of these structures with respect to the ground-state structure are given in kcal/mol in this figure. The Na_8 ground-state structure is a slightly distorted snub disphenoid formed by a square bipyramidal fragment with two extra Na atoms capping two triangular faces. The first low-lying

Fig. 12 Lowest minima structures of the Na_8 clusters (*top*) and lowest minimum structure with a planar high-energy structure of Na_{10} (*bottom*). The energy differences between the found ground-state structures (*left*) and the other structure (*right*) are in kcal/mol



minimum structure is formed by five slightly distorted tetrahedral fragments and lies 2.3 kcal/mol above the ground-state structure. The hierarchical transition state finder was applied with these two minimum structures as input. As result, a transition state characterized by an imaginary frequency of $\omega = 24i \text{cm}^{-1}$ was found.

Figure 13a illustrates some snapshots along the IRC for this cluster rearrangement reaction. The second structure in Fig. 13a is the transition state structure. The very small energy difference of only 0.2 kcal/mol between the high-energy isomer and the transition state challenges the transition state search algorithm. The hierarchical transition state finder worked flawlessly in the Na_8 rearrangement investigated here. The above discussed rearrangement is, however, not the most common rearrangement in low-temperature Na_8 molecular dynamics. Instead, rearrangements that transfer the minimum structure of Na_8 into a mirror image are typical for the low-temperature dynamics of this cluster. Also this kind of rearrangement can be studied with the hierarchical transition state finder if the atoms of the reactant and product structures are correctly labeled (otherwise the initial alignment algorithm will superimpose the two identical structures). The corresponding IRC is depicted in Fig. 13b. At the starting point (first snapshot of Fig. 13b), the square bipyramidal fragment of the ground-state structure is at the right side of the structure. Following the black point on the energy path, we can see how the structure deforms going uphill to the transition state (second snapshot of Fig. 13). This transition state is characterized by an imaginary frequency of $\omega = 22i \text{cm}^{-1}$. The energy difference between this transition state and the ground-state structure is 0.9 kcal/mol. Once the transition state is formed, the structure deforms then downhill to reach again the ground-state structure configuration at the end point of the energy path (last snapshot of Fig. 13b). At this last point, the square bipyramidal fragment, characteristic of the ground-state structure, is now on the left side of the structure.

The lowest minimum structure we found for the Na_{10} cluster is depicted at the bottom of Fig. 12 (left). We notice that as the cluster size increases, it becomes more complicated to locate the ground-state structure due to the increasing

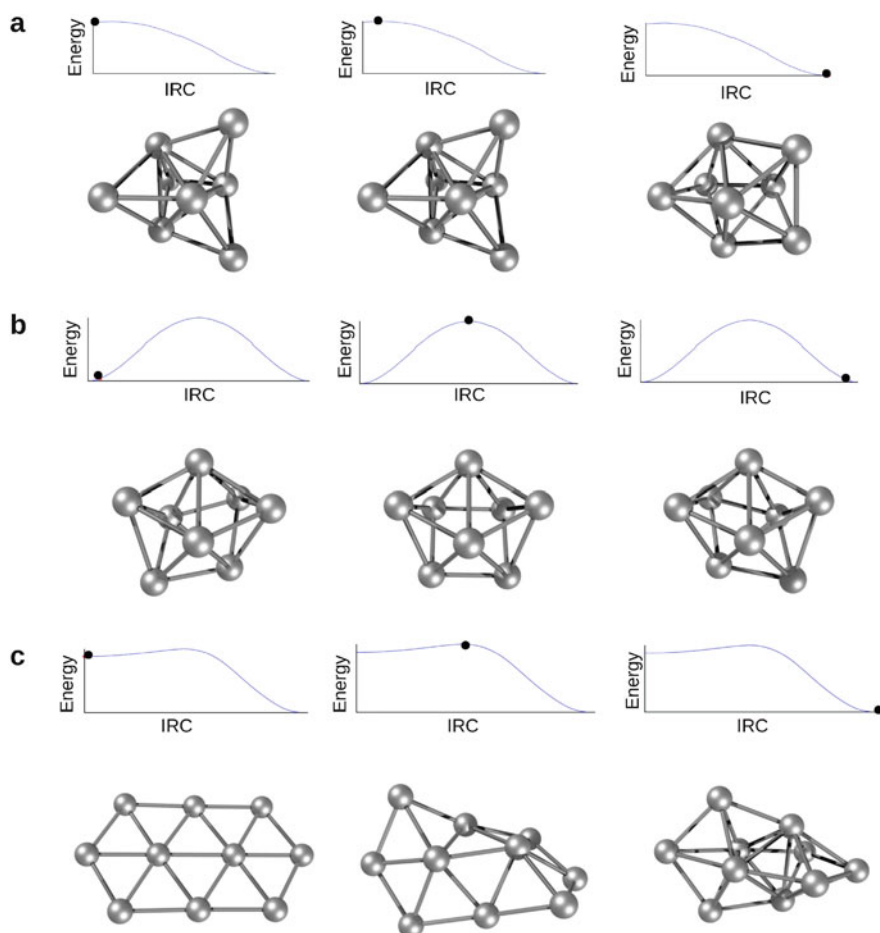


Fig. 13 Snapshots along the IRC for the Na₈ rearrangement between the two lowest minimum structures of this cluster (a), between two mirrored images (b), and rearrangement between the lowest minimum and a planar structure for Na₁₀ (c). The energy is plotted versus the IRC. The structure position is indicated by the *black point* on the energy path

number of low-lying minima. The ground-state structure is built by two pentagonal bipyramidal moieties which are perpendicular to each other. Lying at 11.42 kcal/mol above the ground state, we found a completely planar minimum structure. The hierarchical transition state finder was applied considering this structure and the ground-state structure as input (see Fig. 12 bottom). Due to the shape of the two structures involved as reactant and product, this reaction mechanism is not chemically intuitive. Our result shows that also this rearrangement is free of intermediates. Snapshots along the corresponding IRC are illustrated in Fig. 13c. The activation energy is 1.5 kcal/mol and an imaginary frequency of $\omega = 26i \text{ cm}^{-1}$ characterizes the transition state.

Summary

In this work, a hierarchical transition state algorithm which combines the so-called double-ended interpolation method with the uphill trust region method was applied to search for transition states of selected sodium clusters. These systems are prototypes of finite systems for which transition states are not chemically intuitive. Sodium clusters Na_8 and Na_{10} have been investigated, and selected transition states have been analyzed. All found transition states have been confirmed by the intrinsic reaction coordinate calculations to connect the reactant and product of the corresponding isomerization reaction. The successful application of the hierarchical transition state finder to highly unsymmetrical reaction paths with small activation barriers underlines the robustness of this approach. A particular example is the Na_8 cluster rearrangement. Again it is important to note that no other information than the minimum structures of reactants and products are used in these transition state optimizations. We also showed for Na_8 that with the hierarchical transition state finder transition states between symmetry-equivalent minima can be found if reactant and product structures are correctly labeled. Similar studies for other metal clusters are currently underway in our laboratories.

Conclusion

This review highlights recent developments in auxiliary density functional theory (ADFT) and their implementation in deMon2k. The simplifications associated with ADFT permit an efficient parallel code structure that is suitable for research applications in the nano-regime with chemical accuracy. The presented Born-Oppenheimer molecular dynamics simulation shows that simulation times on the nanosecond time scale can be reached with ADFT. As the applications presented here show, this opens new and exciting perspectives for computational chemistry and material simulations with first-principles methods.

The development of ADFT in deMon2k is also a very educational example for the interplay of theoretical chemistry and computer science. The simplifications of the Kohn-Sham method by employing an atom-centered auxiliary density for the calculation of the Kohn-Sham potential yield simple and very efficient parallel algorithms. A typical example is the SCF acceleration by the MinMax procedure. Because only fitting vectors are involved, it can be applied to systems with hundreds of atoms like the zeolites and fullerenes discussed here. Another example is the formulation of the non-iterative auxiliary density perturbation theory (ADPT) that substitutes the computationally cumbersome CPKS method in ADFT. Similarly, the ADFT-GIAO methodology delivers magnetic properties that are indistinguishable from conventional DFT approaches but with a fraction of the computational cost. With the new variational fitting of exact exchange, almost all chemically relevant exchange-correlation functionals can now be used within the ADFT framework.

The increased use of ADFT in the quantum chemical community over the last years has considerably enhanced our understanding of this fitting approach. While the present authors are responsible for errors or omissions in this small review, the

credit for the advances in ADFT and ADPT is shared with a much wider group cited in the references here and elsewhere.

Acknowledgements Financial support from CONACYT (U48775, 60117-F, CB-179409), ICYTDF (PIFUTP08-87), and CIAM (107310) is gratefully acknowledged. Parts of this review have been realized with the help of the bilateral CONACYT-CNRS project 16871.

Bibliography

- Adamo, C., & Barone, V. (1999). *Journal of Chemical Physics*, *110*, 6158.
- Alberti, A. (1997). *Zeolites*, *19*, 411.
- Alberti, A., Davoli, P., & Vezzalini, G. (1986). *Zeitschrift für Kristallographie*, *175*, 249.
- Almbladh, C. O., & Pedroza, A. C. (1984). *Physical Review A*, *29*, 2322.
- Almlöf, J., Faegri, K., & Korsell, K. (1982). *Journal of Computational Chemistry*, *3*, 385.
- Alvarez-Ibarra, A., Köster, A. M., Zhang, R., & Salahub, D. R. (2012). *Journal of Chemical Theory and Computation*, *8*, 4232.
- Alvarez-Ibarra, A., & Köster, A. M. (2013). *Journal of Chemical Physics*, *139*, 024102.
- Alvarez-Ibarra, A., & Köster, A. M. (2015). *Molecular Physics*, *113*, 3128.
- Andreoni, W. (2007). In *The physics of fullerene-based and fullerene-related materials*. Dordrecht: Kluwer Academic.
- Andzelm, J., Radzio, E., & Salahub, D. R. (1985). *Journal of Computational Chemistry*, *6*, 520.
- Andzelm, J., Russo, N., & Salahub, D. R. (1987). *Journal of Chemical Physics*, *87*, 6562.
- Andzelm, J., & Wimmer, E. (1992). *Journal of Chemical Physics*, *96*, 1280.
- Anquetil, R., Saussey, J. C., & Lavalley, J. C. (1999). *Physical Chemistry Chemical Physics*, *1*, 555.
- Aquilante, F., Pedersen, T. B., de Meras, A. S., & Koch, H. (2006). *Journal of Chemical Physics*, *125*, 174101.
- Aquilante, F., Pedersen, T. B., & Lindh, R. (2007). *Journal of Chemical Physics*, *126*, 194106.
- Baerends, E. J., Ellis, D. E., & Ros, P. (1973). *Chemical Physics*, *2*, 41.
- Bakowies, D., Bühl, M., & Thiel, W. (1995). *Journal of the American Chemical Society*, *117*, 10113.
- Bates, K. R., & Scuseria, G. E. (1998). *Theoretical Chemistry Accounts*, *99*, 29.
- Becke, A. D. (1987). *Journal of Chemical Physics*, *88*, 2547.
- Becke, A. D. (1993). *Journal of Chemical Physics*, *98*, 1372.
- Becke, A. D. (1993). *Journal of Chemical Physics*, *98*, 5648.
- Becke, A. D. (2014). *Journal of Chemical Physics*, *140*, 18A301.
- Belpassi, L., Tarantelli, F., Sgamellotti, A., & Quiney, H. M. (2006). *Journal of Chemical Physics*, *124*, 124104.
- Bergeron, D. E., Castleman, A. W., Jr., Morisato, T., & Khanna, S. N. (2004). *Science*, *304*, 84.
- Bergeron, D. E., Roach, P. J., Castleman, A. W., Jr., Jones, N. O., & Khanna, S. N. (2005). *Science*, *307*, 231.
- Bertran, O., Trickey, S. B., & Torras, J. (2010). *Journal of Computational Chemistry*.
- Binkley, J. S., Pople, J. A., & Dobosh, P. A. (1974). *Molecular Physics*, *28*, 1423.
- Birkenheuer, U., Gordienko, A. B., Nasluzov, V. A., Fuchs-Rohr, M. K., & Rösch, N. (2005). *International Journal of Quantum Chemistry*, *102*, 743.
- Blundell, S. A., Guet, C., & Zope, R. R. (2000). *Physical Review Letters*, *84*, 4826.
- Botalina, O. V., Ioffe, I. N., Sidorov, L. N., Seifert, G., & Vietze, K. (2000). *Journal of the American Chemical Society*, *122*, 9745.
- Bonin, K. D., & Kresin, V. V. (1997). In *Electric-dipole polarizabilities of atoms, molecules and clusters*. Singapore: World Scientific.
- Boys, S. F. (1960). *Reviews of Modern Physics*, *32*, 296.
- Bowlan, J., Liang, A., & de Heer, W. A. (2011). *Physical Review Letters*, *106*, 043401.
- Bühl, M., & Hirsch, A. (2001). *Chemical Reviews*, *101*, 1153.

- Calaminici, P., Jug, K., & Köster, A. M. (1998). *Journal of Chemical Physics*, *109*, 7756.
- Calaminici, P., Jug, K., & Köster, A. M. (1999). *Journal of Chemical Physics*, *111*, 4613
- Calaminici, P., Köster, A. M., Vela, A., & Jug, K. (2000). *Journal of Chemical Physics*, *113*, 2199.
- Calaminici, P., Köster, A. M., Carrington, T., Roy, P. N., Russo, N., & Salahub, D. R. (2001). *Journal of Chemical Physics*, *114*, 4036.
- Calaminici, P., Köster, A. M., & Salahub, D. R. (2003). *Journal of Chemical Physics*, *118*, 4913.
- Calaminici, P., Flores-Moreno, R., & Köster, A. M. (2005). *Computing Letters*, *1*, 164.
- Calaminici, P., Dominguez-Soria, V. D., Geudtner, G., Hernandez-Marin, E., & Köster, A. M. (2006). *Theoretical Chemistry Accounts*, *115*, 221.
- Calaminici, P., Janetzko, F., Köster, A. M., Mejia-Olvera, R., & Zuniga-Gutierrez, B. (2007a). *Journal of Chemical Physics*, *126*, 044108.
- Calaminici, P., Köster, A. M., & Gamboa Martinez, G. U. (2007b). In G. Maroulis & T. Simos (Eds.), *Computational methods in science and engineering, theory and computation: Old problems and new challenges* (Vol. 1). AIP Conference Proceedings, Melville, New York.
- Calaminici, P., Geudtner, G., & Köster, A. M. (2009). *Journal of Chemical Theory and Computation*, *5*, 29.
- Calaminici, P., Carmona-Espíndola, J., Geudtner, G., & Köster, A. M. (2012). *International Journal of Quantum Chemistry*, *112*, 3252.
- Campana, L., Selloni, A., Weber, J., & Goursot, A. (1997). *Journal of Physical Chemistry*, *101*, 9932.
- del Campo, J. M., & Köster, A. M. (2008). *Journal of Chemical Physics*, *129*, 024107.
- Carmona-Espíndola J., Flores-Moreno, R., & Köster, A. M. (2010). *Journal of Chemical Physics*, *133*, 084102.
- Carmona-Espíndola, J., Flores-Moreno, R., & Köster, A. M. (2012). *International Journal of Quantum Chemistry*, *112*, 3461.
- Carmona-Espíndola, J., & Köster, A. M. (2013). *Canadian Journal of Chemistry*, *91*, 795.
- Casida, M. E. (1995). In D. P. Chong (Ed.), *Recent Advances in Density Functional Methods*. Singapore: World Scientific.
- Casida, M. E., Daul, C., Goursot, A., Köster, A. M., Petterson, L. G. M., Proynov, E., St.-Amant, A., Salahub, D. R., Duarte, H., Godbout, N., Guan, J., Jamorski, C., Leboeuf, M., Malkin, V., Malkina, O., Sim, F., & Vela, A. (1996). deMon-KS Version 3.4, deMon Software, Montréal.
- Cerius2 Version 4.10 (2005). Accelrys Inc., San Diego.
- Chacko, S., Kanhere, D. G., & Blundell, S. A. (2005). *Physical Review B*, *71*, 155407.
- Chandrakumar, K. R. S., Ghanty, T. K., & Ghosh, S. K. (2004). *Journal of Chemical Physics*, *120*, 6487.
- Cioslowski, J. (1995). In *Electronic structure calculations on fullerenes and their derivatives*. New York: Oxford University Press.
- Clementi, E. (1963). *Journal of Chemical Physics*, *38*, 996.
- Cruz-Olvera, D., Vasquez, A. T., Geudtner, G., Vásquez-Pérez, J. M., Calaminici, P., & Köster, A. M. (2015). *Journal of Physical Chemistry A*, *119*, 1494.
- Curtiss, L. A., Raghavachari, K., Redfern, P. C., Rassolov, V., & Pople, J. A. (1997). *Journal of Chemical Physics*, *106*, 1063.
- Curtiss, L. A., Raghavachari, K., Redfern, P. C., Rassolov, V., & Pople, J. A. (1998). *Journal of Chemical Physics*, *109*, 7764.
- Curtiss, L. A., Redfern, P. C., Raghavachari, K., & Pople, J. A. (2001). *Journal of Chemical Physics*, *114*, 108.
- Delley, B. (1990). *Journal of Chemical Physics*, *92*, 508.
- Demuth, T., Benco, L., Hafner, J., Toulhouat, H., & Hutschka, F. (2001). *Journal of Chemical Physics*, *114*, 3704.
- Dérouane, E. G., André, J. M., & Lucas, A. A. (1988). *Journal of Catalysis*, *110*, 58.
- Diercksen, G. H. F., & McWeeny, R. (1966). *Journal of Chemical Physics*, *44*, 3554.
- Dirac, P. A. M. (1930). *Proceedings of the Cambridge Philosophical Society*, *26*, 376.
- Ditchfield, R. (1972). *Journal of Chemical Physics*, *56*, 5688.
- Dodds, J. L., McWeeny, R., Raynes, W. T., & Riley, J. P. (1977). *Molecular Physics*, *33*, 611.

- Dominguez-Soria, V. D., Calaminici, P., & Goursot, A. (2007). *Journal of Chemical Physics*, *127*, 154710.
- Dominguez-Soria, V. D., Calaminici, P., & Goursot, A. (2008). In A. Gedeon, P. Massiani, & F. Babonneau (Eds.), *Studies in surface science and catalysis zeolites and related materials: Trends, targets and challenges* (Proceedings of 4th international FEZA conference, Vol. 174, pp. 717). Amsterdam: Elsevier.
- Dominguez-Soria, V. D., Geudtner, G., Morales, J. L., Calaminici, P., & Köster, A. M. (2009). *Journal of Chemical Physics*, *131*, 124102.
- Dominguez-Soria, V. D., Calaminici, P., & Goursot, A. (2011). *The Journal of Physical Chemistry C*, *115*, 6508.
- Dunlap, B. I., & Boettger, J. C. (1996). *Journal of Physics B*, *29*, 4907.
- Dunlap, B. I., & Rösch, (1990). *Advances in Quantum Chemistry*, *21*, 317.
- Dunlap, B. I., Zope, R. R. (2006). *Chemical Physics Letters*, *422*, 451.
- Dunlap, B. I., Connolly, J. W. D., & Sabin, J. R. (1979). *Journal of Chemical Physics*, *71*, 4993.
- Dunlap, B. I., Brenner, D. W., Mintmire, J. W., Mowrey, R. C., & White, C. T. (1991). *Journal of Physical Chemistry*, *95*, 8737.
- Dreizler, R. M., & Gross, E. K. U. (1990). *Density functional theory*. Berlin: Springer.
- Edmiston, C., & Ruedenberg, K. (1963). *Reviews of Modern Physics*, *35*, 457.
- Fermi, E. (1927). *Rendiconti Accademia dei Lincei*, *6*, 602.
- Fermi, E. (1928a). *Zeitschrift für Physik*, *48*, 73.
- Fermi, E. (1928b). *Rendiconti Accademia dei Lincei*, *7*, 342.
- Flores-Moreno, R. (2010). *Journal of Chemical Theory and Computation*, *6*, 48.
- Flores-Moreno, R., & Köster, A. M. (2008). *Journal of Chemical Physics*, *128*, 134105.
- Flores-Moreno, R., & Ortiz, J. V. (2009). *Journal of Chemical Physics*, *131*, 124110.
- Flores-Moreno, R., Melin, J., Ortiz, J. V., & Merino, G. (2008). *Journal of Chemical Physics*, *129*, 224105.
- Fournier, R. (1990). *Journal of Chemical Physics*, *92*, 5422.
- Foster, J. M., & Boys, S. F. (1960). *Reviews of Modern Physics*, *32*, 300.
- Francl, M. M., Petro, W. J., Hehre, W. J., Binkley, J. S., Gordon, M. S., DeFrees, D. J., & Pople, J. A. (1982). *Journal of Chemical Physics*, *77*, 3654.
- Gamboa, G. U., Calaminici, P., & Köster, A. M. (2008). *Journal of Physical Chemistry A*, *112*, 11969.
- Gamboa, G. U., Vásquez-Pérez, J. M., Calaminici, P., & Köster, A. M. (2010). *International Journal of Quantum Chemistry*, *110*, 2172.
- Gaspar, R. (1954). *Acta Physica Academiae Scientiarum Hungaricae*, *3*, 263.
- Gel'fand, I. M., & Fomin, S. V. (1963). *Calculus of variations*. Englewood Cliffs: Prentice Hall.
- Geudtner, G., Janetzko, F., Köster, A. M., Vela, A., & Calaminici, P. (2006). *Journal of Computational Chemistry*, *27*, 483.
- Geudtner, G., Calaminici, P., Carmona-Espíndola, J., del Campo, J. M., Domínguez-Soria, V. D., Flores-Moreno, R., Gamboa, G. U., Goursot, A., Reveles, J. U., Mineva, T., Koster, A. M., Vela, A., Vasquez-Perez, J. M., & Salahub, D. R. (2012). *WIREs: Computational Molecular Science*, *2*, 548.
- Geudtner, G., Calaminici, P., & Köster, A. M. (2013). *Journal of Physical Chemistry C*, *117*, 13210.
- Godbout, N., Salahub, D. R., Andzelm, J., & Wimmer, E. (1992). *Canadian Journal of Physics*, *70*, 560.
- Gonzalez, C., & Schlegel, H. B. (1989). *Journal of Chemical Physics*, *90*, 2154.
- Goursot, A., Fajula, F., Daul, C., & Weber, J. (1998). *Journal of Physical Chemistry*, *92*, 4456.
- Goursot, A., Mineva, T., Vasquez, J. M., Calaminici, P., Köster, A. M., & Salahub, D. R. (2013). *Physical Chemistry Chemical Physics*, *15*, 860.
- Greengard, L., & Rokhlin, V. (1987). *Journal of Computational Physics*, *73*, 325.
- Guan, J. G., Casida, M. E., Köster, A. M., & Salahub, D. R. (1995). *Physical Review B*, *52*, 2184.
- Gusarov, S., Ziegler, T., & Kovalenko, A. (2006). *Journal of Physical Chemistry A*, *110*, 6083.
- Haddon, R. C., Scuseria, G. E., & Smalley, R. E. (1997). *Chemical Physics Letters*, *272*, 38.
- Hall, G. G. (1951). *Proceedings of the Royal Society of London Series A*, *205*, 541.

- Hamel, S., Casida, M. E., & Salahub, D. R. (2001). *Journal of Chemical Physics*, *114*, 7342.
- Heggie, M. I., Terrones, M., Eggen, B. R., Jungnickel, G., Jones, R., Latham, C. D., & Briddon, P. R. (1998). *Physical Review B*, *57* 13339.
- Hehre, W. J., Ditchfield, R., & Pople, J. A. (1972). *Journal of Chemical Physics*, *56*, 2257.
- Hohenberg, P., & Kohn, W., *Physical Review*, *136*, B864 (1964).
- Hollman, D. S., Schaefer, H. F., & Valeev, E. F. (2014). *Journal of Chemical Physics*, *140*, 064109.
- Hoover, W. G. (1985). *Physical Review A*, *31*, 1695.
- Iijima, S. (1980). *Journal of Crystal Growth*, *50*, 675683.
- Ipatov, A., Fouqueau, A., Perez del Valle, C., Cordova, F., Casida, M. E., Köster, A. M., & Vela, A. (2006). *Journal of Molecular Structure: THEOCHEM*, *762*, 179.
- Ito, M., & Saioto, Y. (1985). *Bulletin of the Chemical Society of Japan*, *58*, 3035.
- Itoh, S., Ordejon, P., Drabold, D. A., & Martin, R. M. (1996). *Physical Review B*, *53*, 2132.
- Jacobs, P. A., Martens, G. A. (1987). In *Synthesis in highsilica aluminosilicate zeolites* (p. 321). Amsterdam: Elsevier.
- Jamorski, C., Casida, M. E., & Salahub, D. R. (1996). *Journal of Chemical Physics*, *104*, 5134.
- Janetzko, F., Köster, A. M., & Salahub, D. R. (2008). *Journal of Chemical Physics*, *128*, 024102.
- Johnson, K. H. (1966). *Journal of Chemical Physics*, *45*, 3085.
- Johnson, K. H., & Smith, F. C. (1970). *Physical Review Letters*, *24*, 139.
- Johnson, K. H., & Smith, F. C. (1971). *Chemical Physics Letters*, *10*, 219.
- Johnson, K. H., & Smith, F. C. (1972). *Physical Review B*, *5*, 831.
- Johnson, K. H., & Messmer, R. P. (1974). *Journal of Vacuum Science and Technology*, *11*, 236.
- Kadish, K. M., & Ruoff, R. S. (2007). In *Fullerenes: Chemistry, physics, and technology*. New York: Wiley.
- Karne, A. S., Valal, N., Pal, S., Vázquez-Pérez, J. M., Köster, A. M., & Calaminici, P. (2015). *Chemical Physics Letters*, *635*, 168.
- Knickelbein, M. B. (2001). *Journal of Chemical Physics*, *115*, 5957.
- Knickelbein, M. B. (2003). *Journal of Chemical Physics*, *118*, 6230.
- Knickelbein, M. B. (2004). *Journal of Chemical Physics*, *120*, 10450.
- Komornicki, A., & Fitzgerald, G. (1993). *Journal of Chemical Physics*, *98*, 1398.
- Knight W.D., Clemenger, K., de Heer, W. A., Saunders, W. A., Chou, M. Y., & Cohen, M. L. (1984). *Physical Review Letters*, *52*, 2141.
- Knight, W. D., Clemenger, K., de Heer, W. A., & Saunders, W. A. (1985). *Physical Review B*, *31*, 2539.
- Kohn, W., & Sham, L. J. (1965). *Physical Review*, *140*, A1133.
- Köster, A. M. (1996). *Journal of Chemical Physics*, *104*, 4114.
- Köster, A. M. (1998). Habilitation thesis, Universität Hannover.
- Köster, A. M. (2003). *Journal of Chemical Physics*, *118*, 9943.
- Köster, A. M., Calaminici, P., Gómez, Z., & Reveles, J. U. (2002). In K. Sen (Ed.), *Reviews of modern quantum chemistry, a celebration of the contribution of Robert G. Parr*. River Edge: World Scientific.
- Köster, A. M., Goursot, A., & Salahub, D. R. (2003). In J. McCleverty, T. J. Meyer, & B. Lever (Eds.), *Comprehensive coordination chemistry-II, from biology to nanotechnology* (Vol. 1). Amsterdam: Elsevier.
- Köster, A. M., Flores-Moreno, R., & Reveles, J. U. (2004a). *Journal of Chemical Physics*, *121*, 681.
- Köster, A. M., Reveles, J. U., & del Campo, J. M. (2004b). *Journal of Chemical Physics*, *121*, 3417.
- Köster, A. M., del Campo, J. M., Janetzko, F., & Zuniga-Gutierrez, B. (2009). *Journal of Chemical Physics*, *130*, 114106.
- Köster, A. M., Geudtner, G., Calaminici, P., Casida, M. E., Dominguez-Soria, V. D., Flores-Moreno, R., Gamboa, G. U., Goursot, A., Heine, T., Ipatov, A., Janetzko, F., del Campo, J. M., Reveles, J. U., Vela, A., Zuniga-Gutierrez, B., Salahub, D. R. (2011a). deMon2k Version 3, The deMon developers, Cinvestav, Mexico City, See also <http://www.demon-software.com>

- Köster, A. M., Calaminici, P., Orgaz, E., Roy, D. R., Reveles, J. U., & Khanna, S. N. (2011b). *Journal of the American Chemical Society*, *133*, 12192.
- Köster, A. M., Geudtner, G., Calaminici, P., Casida, M. E., Flores-Moreno, R., Gamboa, G. U., Goursot, A., Janetzko, F., Reveles, J. U., Vela, A. Zuniga-Gutierrez, Salahub, D. R., & Trickey, S. B. (2011c). The deMon2k User's Guide, Version 3.0. Online available at <http://www.demon-software.com>
- Krack, M., & Köster, A. M. (1998). *Journal of Chemical Physics*, *108*, 3226.
- Krishnamurty, S., Heine, T., & Goursot, A. (2003). *Journal of Physical Chemistry B*, *104*, 5728.
- Krishnamurty, S., Stefano, M., Mineva, T., Bégu, S., Devoisselle, J. M., Goursot, A., Zhu, R., & Salahub, D. R. (2008). *ChemPhysChem*, *9*, 2321.
- Krishnamurty, S., Stefano, M., Mineva, T., Bégu, S., Devoisselle, J. M., Goursot, A., Zhu, R., & Salahub, D. R. (2008). *Journal of Physical Chemistry B*, *112*, 13433.
- Kronik, L., Vasiliev, I., & Chelikowsky, J. R. (2000). *Physical Review B*, *62*, 9992.
- Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F., & Smalley, R. E. (1985). *Nature (London)*, *318*, 162.
- Kroto, H. W., & McKay, K. (1988). *Nature*, *331*, 328.
- Kudin, K. N., & Scuseria, G. E. (1998). *Chemical Physics Letters*, *283*, 61.
- Kümmel, S., Akola, J., & Manninen, M. (2000). *Physical Review Letters*, *84*, 4826.
- Laikov, D. N. (1997). *Chemical Physics Letters*, *281*, 151.
- de la Lande, A., Salahub, D. R., & Köster, A. M. (2013). In S. K. Ghosh & P. K. Chatteraj (Eds.), *Concepts and methods in modern theoretical chemistry* (Electronic structure and reactivity, Vol. 1). Boca Raton: CRC Press.
- Lev, B., Zhang, R., Salahub, D. R., & Noskov, S. Y. *Journal of Computational Chemistry*, (2010).
- Levy, M. (1979). *Proceedings of the National Academy of Sciences USA*, *76*, 6062.
- Levy, M., & Perdew, J. P. (1985). In R. M. Dreizler (Ed.), *Density functional methods in physics*. da Providencia J. New York: Plenum.
- Li, H., & Jensen, G. H. (2002). *Theoretical Chemistry Accounts*, *107*, 211.
- London, F. (1937). *Journal de Physique et le Radium*, *8*, 397.
- Malkin, V. G., Malkina, O. L., & Salahub, D. R. (1993). *Chemical Physics Letters*, *204*, 80.
- Malkin, V. G., Malkina, O. L., & Salahub, D. R. (1993). *Chemical Physics Letters*, *204*, 87.
- Malkin, V. G., Malkina, O. L., Casida, M. E., & Salahub, D. R. (1994). *Journal of the American Chemical Society*, *116*, 5898.
- Manzer, S. F., Epifanovsky, E., & Head-Gordon, M., *Journal of Chemical Theory and Computation*, *11*, 518 (2015).
- Marie, O., Thibault-Starzyk, F., & Lavalley, J. C. (2000). *Physical Chemistry Chemical Physics*, *2*, 5341.
- Marie, O., Massiani, P., & Thibault-Starzyk, F. (2004). *Journal of Physical Chemistry B*, *108*, 5073.
- Martyna, G. J., Klein, M. L., & Tuckerman, M. (1992). *Journal of Chemical Physics*, *97*, 2635.
- McWeeny, R. (1962). *Physical Review*, *126*, 1028.
- McWeeny, R. (2001). *Methods of molecular quantum mechanics*, 2nd reprinting. London: Academic.
- McWeeny, R., & Diercksen, G. H. F. (1968). *Journal of Chemical Physics*, *49*, 4852.
- McWeeny, R., Dodds, J. L., & Sadlej A. J. (1977). *Molecular Physics*, *34*, 1779.
- Medel, V. M., Reber, A. C., Chauhan, V., Sen, P., Köster, A. M., Calaminici, P., & Khanna, S. N. (2014). *Journal of the American Chemical Society*, *136*, 8229.
- Meier, W. M. (1961). *Zeitschrift für Kristallographie*, *115*, 439.
- Mejia-Rodriguez, D., & Köster, A. M. (2014). *Journal of Chemical Physics*, *141*, 124114.
- Mejia-Rodriguez, D., Huang, X., del Campo, J. M., & Köster, A. M. (2015a). In J. R. Sabin & R. Cabrera-Trujillo (Eds.), *Advances in quantum chemistry* (Vol. in press). New York: Academic.
- Mejia-Rodriguez, D., Delgado-Venegas, R. I., Calaminici, P., & Köster, A. M. (2015b). *Journal of Chemical Theory and Computation*, *11*, 1493.
- Merlot, P., Kjærgaard, T., Helgaker, T., Lindh, R., Aquilante, F., Reine, S., & Pedersen, T. B. (2013). *Journal of Computational Chemistry*, *34*, 1486.

- Messmer, R. P., Tucker, C. W., & Johnson, K. H. (1975). *Chemical Physics Letters*, 36, 423.
- Messmer, R. P., Salahub, D. R., & Davenport, J. W. (1978). *Chemical Physics Letters*, 57, 29.
- Mintmire, J. W., & Dunlap, B. I. (1982). *Physical Review A*, 25, 88.
- Mintmire, J. W., Sabin, J. R., & Trickey, S. B. (1982). *Physical Review B*, 26, 1743.
- Molof, R. W., Schwartz, H. L., Miller, T. H., & Bederson, B. (1974a). *Physical Review A*, 10, 1131.
- Molof, R. W., Miller, T. M., Schwartz, H. L., Benderson, B., & Park, J. T. (1974b). *Journal of Chemical Physics*, 61, 1816.
- Nosé, S. (1984). *Journal of Chemical Physics*, 81, 511.
- Obara, S., & Saika, A. (1986). *Journal of Chemical Physics*, 84, 3963.
- Parr, R. G., & Yang, W. (1989). *Density-functional theory of atoms and molecules*. New York: Oxford University Press.
- Passaglia, E. (1975). *Contributions to Mineralogy and Petrology*, 50, 65.
- Pearson, R. G. (1973). *Journal of the American Chemical Society*, 85, 3533.
- Perdew, J. P., Burke, K., & Ernzerhof, M. (1996a). *Physical Review Letters*, 77, 3865.
- Perdew, J. P., Ernzerhof, M., & Burke, K. (1996b). *Journal of Chemical Physics*, 105, 9982.
- Petersen, H. G., Soelvason, D., Perram, J. W., & Smith E.R. (1994). *Journal of Chemical Physics*, 101, 8870.
- Pipek, J., & Mezey, P. G. (1989). *Journal of Chemical Physics*, 90, 4916.
- Politzer, P. (1987). *Journal of Chemical Physics*, 86, 1072.
- Polly, R., Werner H.-J., Manby, F. R., & Knowles, P. J. (2004). *Molecular Physics*, 102, 2311.
- Pople, J. A., & Nesbet, R. K. (1954). *Journal of Chemical Physics*, 22, 571.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in FORTRAN* (2nd ed.). Cambridge: Cambridge University Press.
- Rayane, D., Allouche, A. R., Benichou, E., Antoine, R., Aubert-Frecon, M., Dugourd Ph., Broyer, M., Ristori, C., Chandezon, F., Hubert, B. A., & Guet, C. (1999). *European Physical Journal D: Atomic, Molecular, Optical and Plasma Physics*, 9, 243.
- Raynes, W. T., Dodds, J. L., McWeeny, R., & Riley, J. P. (1977). *Molecular Physics*, 33, 611.
- Reveles, J. U., & Köster, A. M. (2004). *Journal of Computational Chemistry*, 25, 1109.
- Roothaan, C. C. J. (1951). *Reviews of Modern Physics*, 23, 69.
- Roothaan, C. C. J. (1960). *Reviews of Modern Physics*, 32, 179.
- Rozanska, X., van Santen, R. A., Demuth, T., Hutschka, F., & Hafner, J. (2003). *Journal of Physical Chemistry B*, 107, 1309.
- Rozanska, X., Barbosa, L. A. M. M., & van Santen, R. A. (2005). *Journal of Physical Chemistry B*, 109, 2203.
- Rungtsirisakun, R., Jansang, B., Pantu, P., & Limtrakul, J. (2005). *Journal of Molecular Structure*, 240, 733.
- Salahub, D. R. (1978). *Journal of the Chemical Society, Chemical Communications*, 385.
- Salahub, D. R., Weber, J., Goursot, A., Köster, A. M., & Vela, A. (2005). In C. E. Dykstra, G. Frenking, K. S. Kim, & G. Scuseria (Eds.), *Theory and applications of the computational chemistry: The first 40 years*. Amsterdam: Elsevier.
- Salahub, D. R., Noskov, S. Y., Lev, B., Zhang, R., Ngo, V., Goursot, A., Calaminici, P., Köster, A. M., Alvarez-Ibarra, A., Mejía-Rodríguez, D., Řezáč J., Cailliez, F., & de la Lande, A. (2015). *Molecules*, 20, 4780.
- Sambe, H., & Felton, R. H. (1975). *Journal of Chemical Physics*, 62, 1122.
- Sauer, S. P. A. (2011). *Molecular electromagnetism: A computational chemistry approach*. New York: Oxford University Press.
- Saunders, V. R. (1983). In G. H. F. Diercksen & S. Wilson (Eds.), *Methods in computational physics* (pp. 1). Dordrecht: Reidel.
- Schwarz, K. (1972). *Physical Review B*, 5, 2466.
- Schlenker, J. L., Pluth, J. J., & Smith, J. V. (1979). *Materials Research Bulletin*, 14, 849.
- Scuseria, G. E. (1995). *Chemical Physics Letters*, 243, 193.
- Scuseria, G. E. (1996). *Science*, 271, 942.
- Seifert, G., Vietze, K., & Schmidt, R. (1996). *Journal of Physics B*, 29, 5183.

- Shao, N., Gao, Y., Yoo, S., An, W., & Zeng, X. C. (2006). *Journal of Physical Chemistry A*, *110*, 7672.
- Shao, N., Gao, Y., & Zeng, X. C. (2007). *Journal of Physical Chemistry C*, *111*, 17671.
- Shedge, S. V., Carmona-Espíndola, J., Pal, S., & Köster, A. M. (2010). *Journal of Physical Chemistry A*, *114*, 2357.
- Shedge, S. V., Pal, S., & Köster, A. M. (2011). *Chemical Physics Letters*, *510*, 185.
- Shedge, S. V., Pal, S., & Köster, A. M. (2012). *Chemical Physics Letters*, *552*, 146.
- Sim, F., Salahub, D. R., & Chin, S. (1992). *International Journal of Quantum Chemistry*, *43*, 463.
- Simoncic, P., & Armbruster, T. (2004). *American Mineralogist*, *89*, 421.
- Slater, J. C. (1951). *Physical Review*, *81*, 385.
- Smirnov, K., & Thibault-Starzyk, F. (1999). *Journal of Physical Chemistry B*, *104*, 8595.
- Sodt, A., & Head-Gordon, M. (2008). *Journal of Chemical Physics*, *128*, 104106.
- St-Amant, A., & Salahub, D. R. (1990). *Chemical Physics Letters*, *169*, 387.
- Stephens, P. J., Devlin, F. J., Chabalowski, C. F., & Frisch, M. (1994). *Journal of Chemical Physics*, *89*, 11623.
- Strain, M. C., Scuseria, G. E., & Frisch, M. J. (1996). *Science*, *271*, 51.
- Thibault-Starzyk, F., Travert, A., & Saussey, J. C. (1998). *Topics in Catalysis*, *6*, 11.
- Thomas, L. H. (1927). *Proceedings of the Cambridge Philosophical Society*, *23*, 542.
- Tikhonov, G., Kasperovich, V., Wong, K., & Kresin, V. V. (2001). *Physical Review A*, *64*, 063202.
- Trickey, S. B., Müller-Plate, F., Dierksen, G. H. F., & Boettger, J. C. (1992). *Physical Review B*, *45*, 4460.
- Trickey, S. B., Alford, J. A., & Boettger, J. C. (2004). In J. Leszczynski (Ed.), *Computational Materials Science, Theoretical and Computational Chemistry* (Vol. 15, pp. 171). Amsterdam: Elsevier.
- Triguero, L., & Pettersson, L. G. M. (1998). *Surface Science*, *398*, 70.
- Triguero, L., Pettersson, L. G. M., & Agren, H. (1998). *Journal of Physical Chemistry A*, *102*, 10599.
- Ugarte, D. (1992). *Nature*, *359*, 707709.
- Ugarte, D. (1995). *Carbon*, *33*, 989993.
- Vahtras, O., Almlöf, J., & Feyereisen, M. W. (1993). *Chemical Physics Letters*, *213*, 514.
- Valerio, G., Goursoot, A., Vetrivel, R., Malkina, O., & Malkin, V. (1998). *Journal of the American Chemical Society*, *120*, 11426.
- Vásquez-Pérez, J. M., Gamboa Martínez, G. U., Köster, A. M., & Calaminici, P. (2009). *Journal of Chemical Physics*, *131*, 124126.
- Vásquez-Pérez, J. M., Calaminici, P., & Köster, A. M. (2013). *Computational and Theoretical Chemistry*, *1021*, 229.
- Vásquez-Pérez, J. M., Gamboa, G. U., Mejía-Rodríguez, D., Alvarez-Ibarra, A., Geudtner, G., Calaminici, P., & Köster, A. M. (2015). *The Journal of Physical Chemistry Letters*, *6*, 4646.
- Velde, G. T., Bickelhaupt, F. M., Baerends, E. J., Guerra, C. F., Van Gisbergen, S. J. A., Snijders, J. G., & Ziegler, T. (2001). *Journal of Computational Chemistry*, *22*, 931.
- Verluis, L., & Ziegler, T. (1988). *Journal of Chemical Physics*, *88*, 322.
- Vos, A. M., Rozanska, X., Schoonheydt, R. A., van Santen, R. A., Hutschka, F., & Hafner, J. (2001). *Journal of the American Chemical Society*, *123*, 2799.
- Vosko, S. H., Wilk, L., & Nusair, M. (1980). *Canadian Journal of Physics*, *58*, 1200.
- Watson, M. A., Handy, N. C., & Cohen, A. J. (2003). *Journal of Chemical Physics*, *119*, 6475.
- Weber, J., Berthou, H., & Jorgensen, C. K. (1977). *Chemical Physics Letters*, *45*, 1.
- Wei, D. Q., & Salahub, D. R. (1994). *Journal of Chemical Physics*, *101*, 7633.
- Wei, D. Q., & Salahub, D. R. (1997). *Journal of Chemical Physics*, *106*, 6086.
- Wei, D. Q., Proynov, E. I., & Milet, A., Salahub, D. R. (2000). *Journal of Physical Chemistry A*, *104*, 2384.
- White, C. A., & Head-Gordon, M. (1994). *Journal of Chemical Physics*, *101*, 6593.
- White, C. A., Johnson, B. G., Gill, P. M. W., & Head-Gordon, M. (1995). *Chemical Physics Letters*, *230*, 8.

- Xu, C. H., & Scuseria, G. E. (1996). *Chemical Physics Letters*, 262, 219.
- Xiao, L., & Wang, L. (2004). *Chemical Physics Letters*, 392, 452.
- Yang, D. S., Zgierski, M. Z., Berces, A., Hackett, P. A., Roy, P. N., Martinez, A., Carrington, T., Salahub, D. R., Fournier, R., Pang, T., & Chen, C. F. (1996). *Journal of Chemical Physics*, 105, 10663.
- Yang, Y., Trafford, K., Kresnawahjuesa, O., Sepa, J., Gorte, R. J., & White, D. (2001). *Journal of Physical Chemistry B*, 105, 1935.
- York, D., Lu, J. P., & Yang, W. (1994). *Physical Review B*, 49, 8526.
- Zhao, Q., & Parr, R. G. (1992). *Physical Review A*, 46, 2337.
- Zhao, Q., & Parr, R. G. (1993). *Journal of Chemical Physics*, 98, 543.
- Zhao, Q., Morrison, R. C., & Parr, R. G. (1994). *Physical Review A*, 50, 2138.
- Zhao, Y., & Truhlar, D. G. (2006). *Theoretical Chemistry Accounts*, 120, 215.
- Zope, R. R., Baruah, T., Pederson, M. R., & Dunlap, B. I. (2008). *Physical Review B*, 77, 115452.
- Zuniga-Gutierrez, B., Geudtner, G., & Köster, A. M. (2011). *Journal of Chemical Physics*, 134, 124108.
- Zuniga-Gutierrez, B., Geudtner, G., & Köster, A. M. (2012). *Journal of Chemical Physics*, 137, 094113.
- Zuniga-Gutierrez, B., Camacho Gonzalez, M., Simon Bastida, P., Bendana Castillo, A., Calaminici, P., & Köster, A. M. (2015). *Journal of Physical Chemistry A*, 119, 1469.
- Zuniga-Gutierrez, B., Camacho Gonzalez, M., Simon Bastida, P., Bendana Castillo, A., Calaminici, P., & Köster, A. M. (2015). *Journal of Chemical Physics*, 143, 104103.

Tao Zeng and Mariusz Klobukowski

Contents

Introduction	862
Free Software	864
GAMESS-US	864
Firefly	866
GAMESS-UK	867
Dalton	868
NWChem	870
ORCA	871
PSI3	872
ACES II	873
ACES III	874
CFOUR	874
COLUMBUS	875
MPQC	875
Commercial Software	876
Gaussian	876
Molcas	876
Q-Chem	876
Turbomole	877
Molpro	877
Jaguar	877
PQS	877
Spartan	877
HyperChem	877
Ampac 9	877
Bibliography	877

T. Zeng (✉)
Department of Chemistry, Carleton University, Ottawa, ON, Canada
e-mail: toby.zeng@carleton.ca

M. Klobukowski
Department of Chemistry, University of Alberta, Edmonton, AB, Canada
e-mail: mariusz.klobukowski@ualberta.ca

Abstract

This chapter reviews most of the widely used nonrelativistic quantum chemistry program packages. Considering that information about availability and capabilities of the free quantum chemistry programs is more limited than that of the commercial ones, the authors concentrated on the free programs. More specifically, the reviewed programs are free for the academic community. Features of these programs are described in detail. The capabilities of each free program can generally be categorized into five fields: independent electron model; electron correlation treatment; excited state calculation; nuclear dynamics including gradient and hessian; and parallel computation. Examples of input files for the Møller–Plesset calculation of formaldehyde are presented for most of the free programs to illustrate how to create the input files. The main contributors of each free program and their institutions are also introduced, with a brief history of program development if available. All the key references of the cited algorithms and the hyperlinks of the home page of each program (both free and commercial) are given in this review for the interested readers. As the most important information of every cited free program's documentation has been extracted here, it is appropriate to consider this chapter to be the manual of manuals.

Introduction

Computational quantum chemistry has been experiencing recently a period of tremendous growth. As the capabilities of computer hardware increased, so did the appetites of researchers for modeling tools. Many computer packages became available, both commercially and as free programs, the latter usually requiring a simple registration procedure to obtain and use the code. The free access is usually restricted to researchers from academic institutions and the programs are sometimes unsupported; however, an associated users' discussion forum often accompanies the program. In this chapter, we will review the software that is currently available, focusing on packages that are familiar to the reviewers. In addition to outlining the unique capabilities of each program, we added the URL links to the home pages to allow the readers to follow the latest developments of each code.

The programs that use only the density functional theory and the programs for relativistic calculations at the level of Dirac theory and beyond are reviewed in other chapters.

All the reviewed programs carry out fundamental tasks of a computational chemist or a computational molecular physicist: calculation of energy for various Hamiltonians; evaluation of gradients of energy (needed to locate stationary points on the potential energy surface); evaluation of the energy hessian (required to analyze the character of the located stationary point, identify local minima and saddle points, and perform vibrational frequency calculation); and evaluation of basic properties (population analysis, dipole moments). The components of the programs include basis set libraries and pseudopotentials.

At the Hartree-Fock level (Hartree 1928; Fock 1930), the energies for closed-shell systems are evaluated using the restricted Hartree-Fock (RHF) method (Hall and Lennard-Jones 1951; Roothaan 1951). For the open-shell molecules, there are several methods that are available in most programs: the unrestricted Hartree-Fock (UHF) method (Pople and Nesbet 1954), several variants of the restricted open-shell Hartree-Fock (ROHF) method (Hsu et al. 1976; McWeeny and Dierksen 1968), and the generalized valence bond (GVB) method (Bobrowicz and Schaefer 1977).

The wavefunction based on a single-configuration Hartree-Fock method is often inadequate to describe bonding in a molecule, and in such a case several multi-configuration methods are used; the approach is usually called multiconfiguration self-consistent field (MCSCF) (Olsen et al. 1983; Roos 1983, 1994; Schmidt and Gordon 1998; Shepard 1987; Werner 1987), and its practical implementations are the full orbital reaction space (FORS) (Ruedenberg et al. 1982a, b, c) and complete active space (CAS) (Roos 1987).

Beyond Hartree-Fock, the energies (and wavefunctions) may be improved at several levels: various orders of the perturbation method at the Møller–Plesset level (1934) (MP2, MP3, MP4); for open-shell systems we could use either unrestricted MP2 (Pople et al. 1976) or one of several variants of the perturbation method based on the ROHF wavefunction: the Z-averaged perturbation theory (ZAPT) (Lee and Jayatilaka 1993; Lee et al. 1994) and RMP (Knowles et al. 1980; Lauderdale et al. 1991); configuration interaction (CI) method (Brooks and Schaefer 1979; Ivanic and Ruedenberg 2001).

Very accurate energies may be obtained using the coupled-cluster theory (CC) (Paldus 2005; Shavitt and Bartlett 2009).

Excited states may be studied using the general post-Hartree-Fock methods listed above, or some specialized techniques, such as configuration interaction with single substitutions (CIS) (Foresman et al. 1992), time-dependent density functional theory (TDDFT) (Dreuw and Head-Gordon 2005; Elliott et al. 2009), and equations-of-motion coupled cluster (EOM-CC) (Kowalski and Piecuch 2004; Włoch et al. 2005).

Solvent effects may be treated using several models: self-consistent reaction field (SCRF) (Karelson et al. 1986, 1993; Kirkwood 1934; Tapia and Goscinski 1975), polarizable continuum model (PCM) (Cammi and Tomasi 1995; Miertuš et al. 1981; Tomasi and Persico 1994; Tomasi et al. 2005), surface and simulation of volume polarization for electrostatics (SS(V)PE) (Chipman 1997, 2000, 2002), and conductor-like screening model (COSMO) (Baldrige and Klamt 1997; Klamt 1995; Klamt and Schüürmann 1993).

Even though the programs described in this chapter are referred to as “non-relativistic” (i.e., using single-component wavefunction), for molecules containing heavy atoms it is necessary to include at least scalar relativistic effects. A popular method of adding relativistic corrections is based on the formalism developed by Douglas and Kroll (1974) and by Hess (1986, 1989), and is usually referred to as DKH n (where n denotes the order of the method).

Some scalar relativistic effects are included implicitly in calculations if pseudopotentials for heavy atoms are used to mimic the presence of core electrons; there are several families of pseudopotentials available: the effective core potentials

(ECP) (Cundari and Stevens 1993; Hay and Wadt 1985; Kahn et al. 1976; Stevens et al. 1984), energy-adjusted pseudopotentials (Cao and Dolg 2006; Dolg 2000; Peterson 2003; Peterson et al. 2003), averaged relativistic effective potentials (AREP) (Hurley et al. 1986; LaJohn et al. 1987; Ross et al. 1990), model core potentials (MCP) (Klobukowski et al. 1999), and ab initio model potentials (AIMP) (Huzinaga et al. 1987).

For very large molecular systems several programs offer semiempirical Hamiltonians (Stewart 1990), including Modified Neglect of Differential Overlap (MNDO) (Dewar and Thiel 1977), Semi-Ab initio Model 1 (SAM1) (Dewar et al. 1993), Austin Model 1 (AM1) (Dewar et al. 1985), Parametric Method 3 (PM3) (Stewart 1989), and Parametric Method 6 (PM6) (Stewart 2007) (the PM3 parameterization is available in the MOPAC2009 program which is freely available to academics, see <http://openmopac.net/downloads.html>). Zerner's modification of the Intermediate Neglect of Differential Overlap approach (ZINDO) (Ridley and Zerner 1973) (see also Zerner 1991) is available in the ArgusLab program <http://www.arguslab.com/index.htm>.

The present Guide will focus on those programs that are freely available to academic community. In describing the capabilities of the programs we heavily borrowed from the available documentation. Commercial programs for ab initio computational chemistry (including several that employ semiempirical methods) will be mentioned only very briefly, guiding the reader to relevant home pages on the Internet.

Several collections of references to quantum chemistry software may be found on the Internet http://en.wikipedia.org/wiki/Quantum_chemistry_computer_programs.

Free Software

Free access to computer programs discussed in the present section is sometimes restricted to academic researchers. Please consult the Internet links for each program for specific restrictions.

GAMESS-US

GAMESS-US (GAMESS = General Atomic and Molecular Electronic Structure System) is one of the early programs for the large-scale ab initio calculations in the public domain (Gordon and Schmidt 2005; Schmidt et al. 1993). The origins of GAMESS-US go back to the National Resources for Computations in Chemistry (NRCC), where the original GAMESS was assembled before 1981 (Dupuis et al. 1980). Since 1982 the development of GAMESS-US was carried out in the research group of Mark Gordon, first at North Dakota State and then at Iowa State University, with several research groups across the world making their contributions. The code is a good illustration of the care and effort that go into the development of any modern ab initio code: the October 2010 (Release 2) version has 769,161 lines of code and 221,147 lines of comments.

The capabilities of GAMESS-US are listed below:

Analytic gradients and Hessians for RHF, ROHF, UHF, GVB, and MCSCF wavefunctions that are used to locate stationary points on the potential energy surface and identify their character (local minimum or transition state).

MP2 energy and gradients for RHF, ROHF, and UHF wavefunctions.

CI energy for very general CI excitation schemes (Ivanic 2003a, b; Ivanic and Ruedenberg 2001).

Several methods of evaluating the CC energy, such as CC with single and double excitations (CCSD), with perturbative triple excitations (CCSD(T)), and including the latest completely renormalized (CR) CR-CC(2,3) method (Piecuch and Włoch 2005; Piecuch et al. 2002); numerical gradients for the CC methods are available.

Excited states may be treated using the CIS, TDDFT, multiconfigurational quasidgenerate perturbation theory (MCQDPT) (Nakano 1993a, b), and equation-of-motion coupled-cluster (EOM-CC) methodologies.

There are many ubiquitous density functionals available.

Extensive basis set libraries.

Pseudopotentials: both effective core potentials and model core potentials are available.

The intrinsic reaction coordinate method allows for checking whether the path from a transition state connects the two minima that correspond to the required reactant and product.

Anharmonic vibrational analysis via the vibrational self-consistent field (VSCF) method (Chaban et al. 1999).

Many molecular properties may be calculated (static polarizability and hyperpolarizability; frequency-dependent polarizability; electric moments; electric field, and electric field gradient).

Solvent effects may be studied via continuum solvation methods using effective fragment potentials (EFP) (Adamovic et al. 2003; Day et al. 1996; Gordon et al. 2007; Jensen et al. 1984), polarizable continuum model (PCM), surface and simulation of volume polarization for electrostatics (SS(V)PE), conductor-like screening model (COSMO), and self-consistent reaction field (SCRF).

A unique feature of the program is its ability to carry out all-electron calculations based on the Fragment Molecular Orbital (FMO) method (Fedorov and Kitaura 2007, 2009).

Several possibilities for carrying out multireference perturbation calculations; spin-orbit coupling may be treated as a perturbation.

For spin-orbit coupling CI wave functions, two-component natural orbitals (natural spinors) may be formed and analyzed (Zeng et al. 2011a, b)

Scalar relativistic effects may be included at the levels of DKH2, DKH3, normalized elimination of small components (NESC) (Dyall 2002), and relativistic elimination of small components (RESC) (Nakajima and Hirao 1999) methods.

Semiempirical Hamiltonians (AM1, PM3, and MNDO) are available.

The input style is illustrated below for the MP2 energy calculation for formaldehyde H_2CO at the experimental geometry $r(\text{CH}) = 1.116 \text{ \AA}$, $r(\text{CO}) = 1.208 \text{ \AA}$. The

basis set is the correlation consistent cc-pVDZ basis (Dunning 1989), used in the form of spherical Gaussian functions. C_{2v} point group symmetry is used. In this example, the geometric structure of formaldehyde is defined in terms of internal coordinates (bond lengths, bond angles, and dihedral angles) in the representation known as Z-matrix. For H_2CO , the first line in the Z-matrix definition specifies the starting atom, in this case oxygen O. The second line, C 1 1.208, defines the second atom, carbon, which is connected to the first atom, and the CO bond length equals 1.208 Å. The first hydrogen atom is defined on the third line, H 2 1.116 1 121.75; this H atom is connected to the second atom (carbon), the CH bond length equals 1.116 Å, and the bond angle H–C(2)–O(1) equals 121.75°. Finally, the fourth line, H 2 1.116 1 121.75 3 180.0, defines the second hydrogen atom, which is connected to the second atom (C) with the CH bond length of 1.116 Å, the bond angle H–C(2)–O(1) of 121.75°, and the dihedral angle H(4)–C(2)–O(1)–H(3) of 180.0°.

```
$contrl mplevl=2 runtyp=energy coord=zmt ispher=1 $end
$basis gbasis=ccd $end
$data
H2CO at experimental geometry (CRC, p. 9-36)
cnv 2

O
C 1 1.208
H 2 1.116 1 121.75
H 2 1.116 1 121.75 3 180.0
$end
```

The home page of the program is <http://www.msg.ameslab.gov/GAMESS/>.

Firefly

Formerly known as PC-GAMESS, the Firefly was developed on the basis of GAMESS-US, ending with the 1999 version. The program has been extensively modified, with large sections written to increase the efficiency of the program. The program is available as binaries for Windows-based computers as well as the Macintosh computers. The development of the code took place at the lab of A. A. Granovsky (Moscow State University). Some interesting features of the program are

The Møller–Plesset energy corrections may be calculated up to the fourth order (MP2, MP3, and MP4).

Compression of files allows for very large calculations to be carried out; the compression is accomplished via packing of integrals, matrix elements, and their indices reduce file size and input/output (I/O) time.

Faster MCQDPT algorithm and availability of XMCQDPT implementation.

Integral code allows for fast evaluation of generally contracted basis sets (Raffenetti 1973), such as atomic natural orbitals (ANO) (Almlöf 1987).

Sample input file is illustrated below for the same MP2 energy calculation. At present, the program's internal library does not have the cc-pVDZ basis set; in order to avoid the explicit (and thus rather lengthy) definition of this basis, the double-zeta basis set of Dunning and Hay (1977) is used instead. This input file also illustrates an alternative definition of the geometric structure of a molecule: in terms of Cartesian coordinates (according to Firefly's defaults, in Ångströms), with only the symmetry-nonequivalent atoms defined. Each line defining atomic Cartesian coordinates specifies also the atom name and its nuclear charge.

```
$contrl mplevl=2 runtyp=energy ds=1 $end
$basis gbasis=DH $end
$data
H2CO at experimental geometry (CRC, p. 9~36)
cnv 2

O   8.0   0.0000000000   0.0000000000   ~0.6036077039
C   6.0   0.0000000000   0.0000000000   0.6043922961
H   1.0   0.9489930831   0.0000000000   1.1916470349
$end
```

Home page <http://classic.chem.msu.su/gran/gamess/index.html>.

GAMESS-UK

This code also originated from the NRCC code (Dupuis et al. 1980) and has been extensively modified by the Daresbury Laboratory (UK) group that included M. F. Guest, J. H. van Lenthe, J. Kendrick, K. Schöffel, P. Sherwood, and R. J. Harrison. A good summary of the code's features has been published (Guest et al. 2005). A summary of the program's capabilities is given below:

Gradients and Hessians for many energies (RHF, ROHF, UHF, MCSCF, CASSCF). Energies of the excited states and corresponding transition moments may be calculated using the multireference CI method (MR-DCI). For the studies of electronically excited states, direct random-phase approximation (RPA) and multiconfigurational linear response (MCLR) (Fuchs et al. 1993) excitation energies and oscillator strengths are also available.

Ionization potentials may be calculated using the Green's function methodology (Cederbaum and Domcke 1977) and the two-particle-hole Tamm-Dancoff method (Schirmer and Cederbaum 1978).

MP2 (energies and gradients) and MP3 (energies).

Multireference second-order perturbation method (CASPT2 method); multireference MP3 method (MR-MP3) is also available.

CCSD and CCSD(T) energies.

Many density functionals for the density functional theory calculations.

Many properties may be evaluated, among them atomic charges (derived from Mulliken and Löwdin population analysis, natural population analysis (Read and Weinhold 1983; Read et al. 1985), and electrostatic potentials), polarizabilities, hyperpolarizabilities, and magnetizabilities.

Semiempirical Hamiltonians (MNDO, MINDO/3, AM1, PM3, PM5, and MNDO-d) are available.

Pseudopotentials (Hay and Wadt 1985; Roy et al. 2008) may be employed, with many pseudopotentials included in the library.

The effects of the solvent may be modeled using the direct reaction field (DRF) method (de Vries et al. 1995; Duijnen and de Vries 1996).

Relativistic effects may be treated using the zeroth-order regular approximation (ZORA) (Faas et al. 1995).

Sample input file for the H₂CO calculation at the MP2/cc-pVDZ level of theory is shown below; notice that the symmetry of the molecule is determined by the program.

```
title
  H2CO at experimental geometry (CRC, p. 9-36)
zmatrix angstrom
  O
  C 1 1.208
  H 2 1.116  1 121.75
  H 2 1.116  1 121.75  3 180.0
end
harmonic
basis cc-pVDZ
scftype mp2
enter
```

Home page is <http://www.cfs.dl.ac.uk/gamess-uk/index.shtml>.

Dalton

To quote from the authors of the program, “the Dalton program system is designed to allow convenient, automated determination of a large number of molecular properties based on an HF, density functional theory (DFT), MP2, coupled cluster, or MCSCF reference wave function.” Given the great flexibility of Dalton’s computational capabilities, the authors describe it as directed at experts in ab initio calculations. The program has been developed at the University of Oslo.

A very short list of Dalton’s capabilities is shown below; for those interested in the evaluation of molecular properties we suggest that Dalton’s manual be consulted (<http://www.kjemi.uio.no/software/dalton/dalton.html>).

Complete Active Space (CAS) or Restricted Active Space (RAS) wave functions.
 Calculation of excited states and core hole states.
 Several coupled-cluster (CC) wave functions.
 MP2 and explicitly correlated R12-MP2 triples models.
 Dalton can calculate a large variety of molecular properties at many levels of theory:
 linear, quadratic, and cubic frequency-dependent response properties.
 Linear second-order polarization propagator approach (SOPPA) (Nielsen et al. 1980).
 NMR properties (magnetizabilities, nuclear shieldings, and all contributions to nuclear spin-spin coupling constants).
 EPR properties (electronic g -tensor, hyperfine coupling tensor, and zero-field splitting tensor).
 Circular dichroism properties (electronic circular dichroism, vibrational circular dichroism, and Raman optical activity).
 Gauge-origin independent magnetic properties.
 Solvent effects may be modeled via the self-consistent reaction field (SCRf) to arbitrary order in the multipole expansion.
 First-order properties for ground states: dipole and quadrupole moments, second moments of the electronic charge distribution and electric field gradients at the nuclei, as well as scalar-relativistic one-electron corrections (Darwin and mass velocity).
 Relativistic two-electron Darwin correction to the ground state.
 Excitation energies may be calculated for several levels of the coupled-cluster wavefunctions.
 Density functional theory.
 Effective core potentials.

Input file for the calculation of the MP2/cc-pVDZ energy of H₂CO is shown below; the symmetry of the molecule is determined by the program.

```

BASIS
CC-pVDZ
Z-matrix input

  4  0
ZMAT
O  1  8.0
C  2  1  1.208  6.0
H  3  2  1.116  1  121.75  1.0      .HF
H  4  2  1.116  1  121.75  3  180.0  0  1.0  .MP2
                                         *MP2 INPUT
**DALTON INPUT                          .MP2 FROZEN
-RUN WAVE FUNCTION                       2  0  0  0
**WAVE FUNCTIONS                          **END OF INPUT

```

Home page <http://www.kjemi.uio.no/software/dalton/dalton.html>.

NWChem

The program has been developed in the Molecular Sciences Software Group at the Pacific Northwest National Laboratory (Richland, WA, USA). The authors of the program state that “NWChem is a computational chemistry package designed to run on high-performance parallel supercomputers. Code capabilities include the calculation of molecular electronic energies and analytic gradients using Hartree-Fock self-consistent field (SCF) theory, Gaussian density function theory (DFT), and second-order perturbation theory. For all methods, geometry optimization is available to determine energy minima and transition states.” Furthermore, they add that “Classical molecular dynamics capabilities provide for the simulation of macromolecules and solutions, including the computation of free energies using a variety of force fields.” Details of the NWChem program system were described (Kendall et al. 2000).

Features of the program include

Spin-orbit density functional theory.

Electron correlation energies may be calculated using many methods, including linearized coupled-cluster doubles (LCCD), coupled-cluster doubles (CCD), linearized coupled-cluster singles and doubles (LCCSD), coupled-cluster singles and doubles (CCSD), coupled-cluster singles, doubles, and triples (CCSDT), coupled-cluster singles, doubles, and active triples (CCSDTA), coupled-cluster singles, doubles, triples, and quadruples (CCSDTQ), CCSD and perturbative connected triples (CCSD(T)), CCSD and perturbative connected triples (CCSD[T]), completely renormalized CCSD[T] method (CR-CCSD[T]), completely renormalized CCSD(T) method (CR-CCSD(T)), CCSD and perturbative locally renormalized CCSD(T) correction (LR-CCSD(T)), quadratic configuration interaction singles and doubles (QCISD), configuration interaction singles and doubles (CISD), configuration interaction singles, doubles, and triples (CISDT), configuration interaction singles, doubles, triples, and quadruples (CISDTQ), second-, third-, and fourth-order Møller–Plesset perturbation theory (MP2, MP3, and MP4).

Multiconfiguration SCF wavefunctions.

Selected configuration interaction with perturbation correction.

Effective core potentials that allow to include both the scalar relativistic effects and spin-orbit effects (Pacios and Christiansen 1985).

Relativistic corrections may be added at the levels of DKH2 and ZORA.

Solvent effects may be described using COSMO (Baldrige and Klamt 1997; Klamt 1995; Klamt and Schüürmann 1993).

For the studies of excited states, the CIS, time-dependent Hartree-Fock (TDHF), and TDDFT methods are available.

Anharmonic vibrational analysis may be carried out using the VSCF method (Chaban et al. 1999).

Sample input file for the H₂CO calculation at the MP2/cc-pVDZ level of theory is shown below; the symmetry of the molecule is determined by the program. The Cartesian coordinates of all atoms must be defined, and the atomic symbol must be used as atom name, eliminating the need to define nuclear charges.

```
echo
start h2co
geometry units angstrom
  O      0.00000      0.00000      -0.60361
  C      0.00000      0.00000      0.60439
  H      0.94899      0.00000      1.19165
  H     -0.94899      0.00000      1.19165
end
basis
  O library cc-pvdz
  C library cc-pvdz
  H library cc-pvdz
end
task mp2 energy
mp2
  freeze core 2
end
```

Home page <http://www.emsl.pnl.gov/capabilities/computing/nwchem/>.

ORCA

The principal developer of the code has been Frank Neese, first at the Max Planck Institute for Bioinorganic Chemistry in Mülheim and then at the University of Bonn. In the words of the program's author, "ORCA is a flexible, efficient and easy-to-use general purpose tool for quantum chemistry with specific emphasis on spectroscopic properties of open-shell molecules. It features a wide variety of standard quantum chemical methods ranging from semiempirical methods to DFT to single- and multireference correlated ab initio methods. It can also treat environmental and relativistic effects." An abbreviated list of program's features includes

Calculation of MP2 and RI-MP2 energies and gradients.

Coupled cluster methods.

CASSCF wavefunction.

Scalar relativistic corrections at the levels of Douglas-Kroll, ZORA, and the infinite-order relativistic approximation (IORA) (Dyall and van Lenthe 1999).

Excited states energies may be calculated using the CIS, CIS(D), and TDDFT methods.

Multireference CI and perturbation theory.

Many properties may be calculated, including spin-orbit coupling, hyperfine and quadrupole couplings, the EPR g -tensor, and zero-field splitting.

Sample input file for calculation of the MP2/cc-pVDZ energy of H₂CO is shown below (the symmetry of the molecule is determined by the program).

```
! MP2 RHF cc-pVDZ
* int 0 1
C(1)  0 0 0 0.00  0.0  0.00
O(2)  1 0 0 1.208  0.0  0.00
H(3)  1 2 0 1.116 121.75  0.00
H(3)  1 2 3 1.116 121.75 180.00
*
```

Home page <http://www.thch.uni-bonn.de/tc/orca/>.

PSI3

In the description of the designers of the code, C.D. Sherrill, T.D. Crawford, and E.F. Valeev (from Georgia Institute of Technology and Virginia Tech), “the PSI3 suite of quantum chemical programs is designed for efficient, high-accuracy calculations of properties of small to medium-sized molecules. The package’s current capabilities include a variety of Hartree-Fock, coupled cluster, complete-active-space self-consistent-field, and multi-reference configuration interaction models.” PSI3, written in the programming languages C and C++ (rather than in the Fortran language, which has been traditionally used in computational quantum chemistry), was recently reviewed (Crawford et al. 2007).

Some interesting features of the program are

Arbitrarily high angular momentum levels in integrals and derivative integrals (up to m -type functions have been tested).

Coupled cluster methods including CC2, CCSD, CCSD(T), and CC3 with RHF, ROHF, UHF, and Brueckner orbitals.

Determinant-based CI including CASSCF, RAS-CI, and full CI.

MP2 and MP2-R12 methods.

Methods for excited state: CIS, CIS(D), random phase approximation (RPA), EOM-CCSD, and CC3.

Analytic energy gradients for CCSD based on RHF, ROHF, and UHF orbitals.

Coupled cluster linear response methods for static and dynamic polarizabilities and optical rotation.

Diagonal Born–Oppenheimer correction (DBOC) for RHF, ROHF, UHF, and CI wave functions.

Sample input file for the H₂CO calculation at the MP2/cc-pVDZ level of theory is shown below (symmetry of the molecule is determined by the program).

```
psi:(
  label = "H2CO at experimental geometry (CRC, p. 9-36)"
  wfn = mp2
  reference = rhf
  basis = "cc-pvDZ"
  zmat = (
    O
    C 1 1.208
    H 2 1.116 1 121.75
    H 2 1.116 1 121.75 3 180.0
  )
)
```

Home page <http://www.psicode.org/>.

ACES II

The program ACES II (Advanced Concepts in Electronic Structure) was developed in the research group of R.J Bartlett at the Quantum Theory Project (QTP) of the University of Florida in Gainesville. The program's backbone was written in the early 1990s and has undergone extensive developments since that time. ACES II focused on the coupled-cluster theory (Paldus 2005) and its applications (Bartlett 2005; Bartlett and Musiał 2007).

Single point energy calculations may be carried out for the following cases:

Independent particle models RHF, UHF, and ROHF.

Correlation methods utilizing RHF and UHF reference determinants, including MBPT(2), MBPT(3), SDQ-MBPT(4), MBPT(4), CCD, CCSD, CCSD(T), CCSD + TQ*(CCSD), CCSD(TQ), CCSDT-1, CCSDT-2, CCSDT-3, QCISD, QCISD(T), QCISD(TQ), UCCS(4), UCCSD(4), CID, and CISD.

Correlation methods that can use ROHF reference determinants include MBPT(2), CCSD, CCSDT, CCSD(T), CCSDT-1, CCSDT-2, and CCSDT-3.

Correlation methods that can use quasi-restricted Hartree-Fock (QRHF) or Brueckner orbital reference determinants include CCSD, CCSDT, CCSD(T), CCSDT-1, CCSDT-2, and CCSDT-3.

Two-determinant CCSD calculations for open-shell singlet state.

Equation-of-motion CCSD calculation of dynamic polarizabilities and of NMR spin-spin coupling constants.

Equation-of-motion CCSD calculations of excitation energies.

Kohn-Sham DFT density functional methods combined with a wide selection of density functionals.

Analytical gradients may be calculated for

Independent particle models RHF, UHF, and ROHF.

Correlation methods utilizing RHF and UHF reference determinants, including MBPT(2), MBPT(3), SDQ-MBPT(4), MBPT(4), CCD, CCSD, CCSD + T(CCSD), CCSD(T), CCSDT-1, CCSDT-2, CCSDT-3, QCISD, QCISD(T), UCC(4), UCCSD(4), CID, and CISD.

Correlation methods that can utilize ROHF reference determinants, including MBPT(2), CCSD, and CCSD(T).

Correlation methods that can utilize QRHF reference determinants, including CCSD.

Two-determinant CCSD calculations for open-shell singlet state based on QRHF orbitals.

EOM-CCSD for excited states.

TD-CCSD analytical derivatives.

Analytical Hessians may be carried out for the independent particle models RHF, UHF, and ROHF.

Home page <http://www.qtp.ufl.edu/ACES/>.

ACES III

The program represents a significant advance in parallel processing and resulted from completely rewriting ACES II for efficiency in parallel calculations. The parallel technology based on the super instruction assembly language (SIAL) that marked the progress from ACES II to ACES III was discussed by Bartlett and coworkers (Lotrich et al. 2008).

Home page <http://www.qtp.ufl.edu/ACES/>.

CFOUR

The four Cs in the name of the program stand for Coupled-Cluster techniques for Computational Chemistry. The predecessor of this program was known as the Mainz-Austin-Budapest version of ACES II which was replaced by CFOUR in April 2009 (see <http://www.aces2.de/>). At present, the development of the program is continued at the University of Mainz (J. Gauss), at the University of Texas at Austin (J.F. Stanton and M.E. Harding), and at Eötvös Loránd University (P.G. Szalay).

Some of the features of the program are

NMR chemical shifts.

Many properties, including nuclear spin-rotation constants, magnetizabilities, rotational g -tensors, indirect spin-spin coupling constants.

Electronically excited states may be calculated via CIC, CIS(D), and several EOM-CC methods; the ionized and electron-attached states may be described via EOM-CC methods.

Anharmonic force-field calculations may be carried out.

Vibrationally averaged properties.

Relativistic corrections.

Diagonal Born–Oppenheimer corrections.

Automated additivity and basis-set extrapolation schemes.

Basis sets up to *i*-type Gaussian functions.

Effective core potentials.

Home page <http://www.cfour.de/>.

COLUMBUS

COLUMBUS originated in 1980 at the Ohio State University and was developed by I. Shavitt, H. Lischka (University of Vienna), and R. Shepard (Battelle Columbus Laboratories). The program was designed to carry out a variety of multireference calculations on ground and excited states of molecules (Lischka et al. 2006). Specific features include

Geometry optimization to find minima and saddle points.

Searches for minima on the crossing seam (conical intersections).

General MCSCF with quadratic convergence including state averaging.

Direct MR-CISD (single and double excitations), using an arbitrary set of reference configurations.

Multireference averaged coupled-pair-functional (MR-ACPF) and multi-reference average quadratic coupled-cluster (MR-AQCC) with size-extensivity corrections.

Spin-orbit CI (based on graphical unitary group approach) (Yabushita et al. 1999).

Search for conical intersections may be done at the MRCI level.

For excited states, MCSCF, MR-CISD, and MR-AQCC-LRT transition moments may be calculated.

Finite field method for second-order properties (polarizabilities).

Home page <http://www.univie.ac.at/columbus/>.

MPQC

The Massively Parallel Quantum Chemistry (MPQC) program is an object-oriented code that was written in C++. The work on the code has been initiated by C.L. Janssen at the Sandia National Laboratories (USA).

The capabilities of MPQC include

Energies and gradients are available for closed shell, unrestricted and general restricted open shell Hartree-Fock and density functional methods.

Energies may be calculated for the second-order open-shell perturbation theory.

Energies and gradients may be evaluated for the second-order closed-shell Møller–Plesset perturbation theory (MP2).

MP2 energies of closed-shell systems, including an R12 correlation factor and using an auxiliary basis set, are supported.

Home page <http://www.mpqc.org/index.php>.

Commercial Software

For the details of commercial software the reader is referred to the links listed below for the most up-to-date information about each program. For completeness, both ab initio and semiempirical programs are included.

Gaussian

One of the first programs for large-scale calculations in the commercial domain, dating its origins to 1970, is currently in its 2009 edition.¹

Please see home page <http://www.gaussian.com/>.

Molcas

Please see home page <http://www.teokem.lu.se/molcas/introduction.html>

Q-Chem

Please see home page <http://www.q-chem.com/>

¹The original 13,370 lines of Gaussian 70 code were released to general public via the now defunct Quantum Chemistry Program Exchange (QCPE). Historic information about QCPE may be found on http://www.ccl.net/ccl/qcpe/QCPE_removed/. QCPE offered as the first ab initio program Polyatom (Version I with 3,275 lines of code) was made available by Csizmadia et al. in 1964. It is worth mentioning in passing that the fees that QCPE charged for the programs were very modest by today's standards: \$175 for codes greater than 10,000 lines plus \$35 for media and handling. The programs grew as new capabilities were added: in 1974 Polyatom (Version II for IBM 360) grew to 20,000 lines, while the 1980 release of Gaussian (IBM Version II) contained about 60,000 lines of code. The current status of QCPE was explained in a brief note saved in Computational Chemistry List <http://ccl.net/chemistry/resources/messages/2009/06/04.001-dir/index.html>.

Turbomole

Please see home page <http://www.turbomole.com/>

Molpro

Please see home page <http://www.molpro.net/>

Jaguar

Please see home page <https://www.schrodinger.com/products/14/7/>.

PQS

Parallel Quantum Solutions. Please see home page <http://www.pqs-chem.com/software.shtml>

Spartan

Please see home page <http://www.wavefun.com/>

HyperChem

Please see home page <http://www.hyper.com/>

Ampac 9

Please see home page <http://www.semichem.com/>

Acknowledgments The work was partly supported by the Research Grant No. G12121041441 to MK from the Natural Sciences and Engineering Research Council of Canada.

Bibliography

- Adamovic, I., Freitag, M. A., & Gordon, M. S. (2003). Density functional theory based effective fragment potential method. *The Journal of Chemical Physics*, 118, 6725–6732.
- Almlöf, J. (1987). General contraction of Gaussian basis sets. I. Atomic natural orbitals for first- and second-row atoms. *The Journal of Chemical Physics*, 86, 4070–4077.

- Baldrige, K., & Klamt, A. (1997). First principles implementation of solvent effects without outlying charge error. *The Journal of Chemical Physics*, *106*, 6622–6633.
- Bartlett, R. J. (2005). How and why coupled-cluster theory became the pre-eminent method in an ab initio quantum chemistry. In C. E. Dykstra, G. Frenking, K. S. Kim, & G. E. Scuseria (Eds.), *Theory and applications of computational chemistry: The first forty years* (pp. 1191–1221). Amsterdam/Boston: Elsevier.
- Bartlett, R. J., & Musiał, M. (2007). Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics*, *79*, 291.
- Bobrowicz, F. W., & Schaefer, H. F., III. (1977). The self-consistent field equations for generalized valence bond and open-shell Hartree-Fock wave functions. In H. F. Schaefer III (Ed.), *Methods of electronic structure theory (modern theoretical chemistry)* (Vol. 3, pp. 79–127). New York/London: Plenum.
- Brooks, B. R., & Schaefer, H. F., III. (1979). The graphical unitary group approach to the electron correlation problem. Methods and preliminary applications. *The Journal of Chemical Physics*, *70*, 5092–5106.
- Cammi, R., & Tomasi, J. (1995). Remarks on the use of the apparent surface charges (ASC) methods in solvation problems: Iterative versus matrix-inversion procedures and the renormalization of the apparent charges. *Journal of Computational Chemistry*, *16*, 1449–1458.
- Cao, X., & Dolg, M. (2006). Relativistic energy-consistent ab initio pseudopotentials as tools for quantum chemical investigations of actinide systems. *Coordination Chemistry Reviews*, *250*, 900–910.
- Cederbaum, L. S., & Domcke, W. (1977). Theoretical aspects of ionization potentials and photoelectron spectroscopy: A Green's function approach. *Advances in Chemical Physics*, *36*, 205–344.
- Chaban, G. M., Jung, J. O., & Gerber, R. B. (1999). Ab initio calculation of anharmonic vibrational states of polyatomic systems: Electronic structure combined with vibrational self-consistent field. *The Journal of Chemical Physics*, *111*, 1823–1829.
- Chipman, D. M. (1997). Charge penetration in dielectric models of solvation. *The Journal of Chemical Physics*, *106*, 10194–10206.
- Chipman, D. M. (2000). Reaction field treatment of charge penetration. *The Journal of Chemical Physics*, *112*, 5558–5565.
- Chipman, D. M. (2002). Comparison of solvent reaction field representations. *Theoretical Chemistry Accounts*, *107*, 80–89.
- Crawford, T. D., Sherrill, C. D., Valeev, E. F., Fermann, J. T., King, R. A., Leininger, M. L., Brown, S. T., Janssen, C. L., Seidl, E. T., Kenny, J. P., & Allen, W. D. (2007). PSI3: An open-source ab initio electronic structure package. *Journal of Computational Chemistry*, *28*, 1610–1616.
- Csizmadia, I. G., Harrison, M. C., Moskowitz, J. W., Seung, S., Sutcliffe, B. T., & Barrett, M. P. (1964). POLYATOM. *Quantum Chemistry Program Exchange*, *11*, 47.
- Cundari, T. R., & Stevens, W. J. (1993). Effective core potential methods for the lanthanides. *The Journal of Chemical Physics*, *98*, 5555–5565.
- Day, P. N., Jensen, J. H., Gordon, M. S., Webb, S. P., Stevens, W. J., Krauss, M., Garmer, D., Basch, H., & Cohen, D. (1996). An effective fragment method for modeling solvent effects in quantum mechanical calculations. *The Journal of Chemical Physics*, *105*, 1968–1986.
- de Vries, A. H., van Duijnen, P. T., Juffer, A. H., Rullmann, J. A. C., Dijkman, J. P., Merenga, H., & Thole, B. T. (1995). Implementation of reaction field methods in quantum chemistry computer codes. *Journal of Computational Chemistry*, *16*, 37–55.
- Dewar, M. J. S., & Thiel, W. (1977). Ground states of molecules. 38. The MNDO method. Approximations and parameters. *Journal of the American Chemical Society*, *99*, 4899–4906.
- Dewar, M. J. S., Zoebisch, E. G., Healy, E. F., & Stewart, J. J. P. (1985). AM1: A new general purpose quantum mechanical molecular model. *Journal of the American Chemical Society*, *107*, 3902–3909.
- Dewar, M. J. S., Jie, C., & Yu, J. (1993). SAM1; the first of a new series of general purpose quantum mechanical molecular models. *Tetrahedron*, *49*, 5003–5038.

- Dolg, M. (2000). Effective core potentials. In J. Grotendorst (Ed.), *Modern methods and algorithms of quantum chemistry* (Vol. 1, pp. 479–508). Jülich: John von Neumann Institute for Computing.
- Douglas, M., & Kroll, N. M. (1974). Quantum electrodynamic corrections to the fine structure of helium. *Annals of Physics*, 82, 89–155.
- Dreuw, A., & Head-Gordon, M. (2005). Single-reference ab initio methods for the calculation of excited states of large molecules. *Chemical Reviews*, 105, 4009–4037.
- Duijnen, P. T. V., & de Vries, A. H. (1996). Direct reaction field force field: A consistent way to connect and combine quantum-chemical and classical descriptions of molecules. *International Journal of Quantum Chemistry*, 60, 1111–1132.
- Dunning, T. H. (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90, 1007–1023.
- Dunning, T. H., & Hay, P. J. (1977). Gaussian basis sets for molecular calculations. In H. F. Schaefer III (Ed.), *Methods of electronic structure theory (modern theoretical chemistry)* (Vol. 3, pp. 1–27). New York/London: Plenum.
- Dupuis, M., Spangler, D., & Wendoloski, J. (1980). *NRCC software catalog* (Vol. 1, Program No. QG01 GAMESS Tech. rep.) Berkeley: National Resource for Computations in Chemistry, University of California.
- Dyall, K. G. (2002). A systematic sequence of relativistic approximations. *Journal of Computational Chemistry*, 23, 786–793.
- Dyall, K. G., & van Lenthe, E. (1999). Relativistic regular approximations revisited: An infinite-order relativistic approximation. *The Journal of Chemical Physics*, 111, 1366–1372.
- Elliott, P., Furche, F., & Burke, K. (2009). Excited states from time-dependent density functional theory. *Reviews in Computational Chemistry*, 26, 91–166.
- Faas, S., Snijders, J. G., van Lenthe, J. H., van Lenthe, E., & Baerends, E. J. (1995). The ZORA formalism applied to the Dirac-Fock equation. *Chemical Physics Letters*, 246, 632–640.
- Fedorov, D. G., & Kitaura, K. (2007). Extending the power of quantum chemistry to large systems with the fragment molecular orbital method. *The Journal of Physical Chemistry. A*, 111, 6904–6914.
- Fedorov, D. G., & Kitaura, K. (2009). *The fragment molecular orbital method: Practical applications to large molecular systems*. Boca Raton: CRC Press.
- Fock, V. A. (1930). Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems. *Zeitschrift für Physik*, 61, 126–148.
- Foresman, J. B., Head-Gordon, M., Pople, J. A., & Frisch, M. J. (1992). Toward a systematic molecular orbital theory for excited states. *The Journal of Physical Chemistry*, 96, 135–149.
- Fuchs, C., Bonačić-Koutecký, V., & Koutecký, J. (1993). Compact formulation of multiconfigurational response theory. Applications to small alkali metal clusters. *The Journal of Chemical Physics*, 98, 3121–3140.
- Gordon, M. S., & Schmidt, M. W. (2005). Advances in electronic structure theory: GAMESS a decade later. In C. E. Dykstra, G. Frenking, K. S. Kim, & G. E. Scuseria (Eds.), *Theory and applications of computational chemistry: The first forty years* (pp. 1167–1189). Amsterdam: Elsevier.
- Gordon, M. S., Slipchenko, L. V., Li, H., & Jensen, J. H. (2007). The effective fragment potential: A general method for predicting intermolecular interactions. In D. Spellmeyer & R. Wheeler (Eds.), *Annual reports in computational chemistry* (Vol. 3, pp. 177–193). Amsterdam: Elsevier.
- Guest, M. F., Bush, I. J., van Dam, H. J. J., Sherwood, P., Thomas, J. M. H., van Lenthe, J. H., Havenith, R. W. A., & Kendrick, J. (2005). The GAMESS-UK electronic structure package: Algorithms, developments and applications. *Molecular Physics*, 103, 719–747.
- Hall, G. G., & Lennard-Jones, J. (1951). The molecular orbital theory of chemical valency. III. Properties of molecular orbitals. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 202, 155–165.
- Hartree, D. R. (1928). The wave mechanics of an atom with a non-coulomb central field. Part I. Theory and methods. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24, 89–110.

- Hay, P. J., & Wadt, W. R. (1985). Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *The Journal of Chemical Physics*, *82*, 299–310.
- Hess, B. A. (1986). Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators. *Physical Review A*, *33*, 3742–3748.
- Hess, B. A. (1989). Revision of the Douglas-Kroll transformation. *Physical Review A*, *39*, 6016–6017.
- Hsu, H., Davidson, E. R., & Pitzer, R. M. (1976). An SCF method for hole states. *The Journal of Chemical Physics*, *65*, 609–613.
- Hurley, M. M., Pacios, L. F., Christiansen, P. A., Ross, R. B., & Ermler, W. C. (1986). Ab initio relativistic effective core potentials with spin-orbit operators. II. K through Kr. *The Journal of Chemical Physics*, *84*, 6840–6853.
- Huzinaga, S., Seijo, L., Barandiarán, Z., & Klobukowski, M. (1987). The ab initio model potential method. Main group elements. *The Journal of Chemical Physics*, *86*, 2132–2145.
- Ivanic, J. (2003a). Direct configuration interaction and multiconfigurational self-consistent-field method for multiple active spaces with variable occupations. I. Method. *The Journal of Chemical Physics*, *119*, 9364–9376.
- Ivanic, J. (2003b). Direct configuration interaction and multiconfigurational self-consistent-field method for multiple active spaces with variable occupations. II. Application to oxoMn(salen) and N₂O₄. *The Journal of Chemical Physics*, *119*, 9377–9385.
- Ivanic, J., & Ruedenberg, K. (2001). Identification of deadwood in configuration spaces through general direct configuration interaction. *Theoretical Chemistry Accounts*, *106*, 339–351.
- Jensen, J. H., Day, P. N., Gordon, M. S., Basch, H., Cohen, D., Garmer, D. R., Krauss, M., & Stevens, W. J. (1984). An effective fragment method for modeling intermolecular hydrogen bonding-effects on quantum mechanical calculations. In D. A. Smith (Ed.), *Modeling the hydrogen bond* (ACS symposium, Vol. 569, pp. 139–151). New York: ACS.
- Kahn, L. R., Baybutt, P., & Truhlar, D. G. (1976). Ab initio effective core potentials: Reduction of all-electron molecular structure calculations to calculations involving only valence electrons. *The Journal of Chemical Physics*, *65*, 3826–3853.
- Karelson, M. M., Katritzky, A. R., & Zerner, M. C. (1986). Reaction field effects on the electron distribution and chemical reactivity of molecules. *International Journal of Quantum Chemistry*, *30*, 521–527.
- Karelson, M., Tamm, T., & Zerner, M. C. (1993). Multicavity reaction field method for the solvent effect description in flexible molecular systems. *The Journal of Physical Chemistry*, *97*, 11901–11907.
- Kendall, R. A., Aprà, E., Bernholdt, D. E., Bylaska, E. J., Dupuis, M., Fann, G. I., Harrison, R. J., Ju, J., Nichols, J. A., Nieplocha, J., Straatsma, T. P., Windus, T. L., & Wong, A. T. (2000). High performance computational chemistry: An overview of NWChem a distributed parallel application. *Computer Physics Communications*, *128*, 260–283.
- Kirkwood, J. G. (1934). Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *The Journal of Chemical Physics*, *2*, 351–361.
- Klamt, A. (1995). Conductor-like screening model for real solvents: A new approach to the quantitative calculation of solvation phenomena. *The Journal of Physical Chemistry*, *99*, 2224–2235.
- Klamt, A., & Schüürmann, G. (1993). COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of the Chemical Society, Perkin Transactions*, *2*, 799–805.
- Klobukowski, M., Huzinaga, S., & Sakai, Y. (1999). Model core potentials: Theory and applications. In J. Leszczynski (Ed.), *Computational chemistry: Reviews of current trends* (Vol. 3, pp. 49–74). Singapore: World Scientific.
- Knowles, P. J., Andrews, J. S., Amos, R. D., Handy, N. C., & Pople, J. A. (1980). Restricted Møller-Plesset theory for open-shell molecules. *Chemical Physics Letters*, *186*, 130–136.

- Kowalski, K., & Piecuch, P. (2004). New coupled-cluster methods with singles, doubles, and noniterative triples for high accuracy calculations of excited electronic states. *The Journal of Chemical Physics*, *120*, 1715–1738.
- LaJohn, L. A., Christiansen, P. A., Ross, R. B., Atashroo, T., & Ermler, W. C. (1987). Ab initio relativistic effective core potentials with spin-orbit operators. III. Rb through Xe. *The Journal of Chemical Physics*, *87*, 2812–2824.
- Lauderdale, W. J., Stanton, J. F., Gauss, J., Watts, J. D., & Bartlett, R. J. (1991). Many-body perturbation theory with a restricted open-shell Hartree-Fock reference. *Chemical Physics Letters*, *187*, 21–28.
- Lee, T. J., & Jayatilaka, D. (1993). An open-shell restricted Hartree-Fock perturbation theory based on symmetric spin orbitals. *Chemical Physics Letters*, *201*, 1–10.
- Lee, T. J., Rendell, A. P., Dyall, K. G., & Jayatilaka, D. (1994). Open-shell restricted Hartree-Fock perturbation theory: Some considerations and comparisons. *The Journal of Chemical Physics*, *100*, 7400–7409.
- Lischka, H., Shepard, R., Shavitt, I., Pitzer, R. M., Dallos, M., Müller, Th., Szalay, P. G., Brown, F. B., Ahlrichs, R., Böhm, H. J., Chang, A., Comeau, D. C., Gdanitz, R., Dachsel, H., Ehrhardt, C., Ernzerhof, M., Höchtel, P., Irlé, S., Kedziora, G., Kovar, T., Parasuk, V., Pepper, M. J. M., Scharf, P., Schiffer, H., Schindler, M., Schüler, M., Seth, M., Stahlberg, E. A., Zhao, J.-G., Yabushita, S., Zhang, Z., Barbatti, M., Matsika, S., Schuurmann, M., Yarkony, D. R., Brozell, S. R., Beck, E. V., & Blaudeau, J.-P. (2006). *COLUMBUS, an ab initio electronic structure program, release 5.9.1*.
- Lotrich, V., Flocke, N., Ponton, M., Yau, A. D., Perera, A., Deumens, E., & Bartlett, R. J. (2008). Parallel implementation of electronic structure energy, gradient, and hessian calculations. *The Journal of Chemical Physics*, *128*, 194104/1–15.
- McWeeny, R., & Dierksen, G. H. F. (1968). Self-consistent perturbation theory. II. Extension to open shells. *The Journal of Chemical Physics*, *49*, 4852–4856.
- Miertuš, S., Scrocco, E., & Tomasi, J. (1981). Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects. *Chemical Physics*, *55*, 117–129.
- Møller, Ch., & Plesset, M. S. (1934). Note on an approximation treatment for many-electron systems. *Physical Review*, *46*, 618–622.
- Nakajima, T., & Hirao, K. (1999). A new relativistic theory: A relativistic scheme by eliminating small components (RESC). *Chemical Physics Letters*, *302*, 383–391.
- Nakano, H. (1993a). MCSCF reference quasidegenerate perturbation theory with Epstein-Nesbet partitioning. *The Journal of Chemical Physics*, *99*, 7983–7992.
- Nakano, H. (1993b). Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions. *Chemical Physics Letters*, *207*, 372–378.
- Nielsen, E. S., Jørgensen, P., & Oddershede, J. (1980). Transition moments and dynamic polarizabilities in a second order polarization propagator approach. *The Journal of Chemical Physics*, *73*, 6238–6246.
- Olsen, J., Yeager, D. L., & Jørgensen, P. (1983). Optimization and characterization of a multiconfigurational self-consistent field (MCSCF) state. *Advances in Chemical Physics*, *54*, 1–176.
- Pacios, L. F., & Christiansen, P. A. (1985). Ab initio relativistic effective potentials with spin-orbit operators. I. Li through Ar. *The Journal of Chemical Physics*, *82*, 2664–2671.
- Paldus, J. (2005). The beginnings of coupled-cluster theory: An eyewitness account. In C. E. Dykstra, G. Frenking, K. S. Kim, & G. E. Scuseria (Eds.), *Theory and applications of computational chemistry: The first forty years* (pp. 115–147). Amsterdam: Elsevier.
- Peterson, K. A. (2003). Systematically convergent basis sets with relativistic pseudopotentials. I. Correlation consistent basis sets for the post-d group 13–15 elements. *The Journal of Chemical Physics*, *119*, 11099–11112.

- Peterson, K. A., Figgen, D., Goll, E., Stoll, H., & Dolg, M. (2003). Systematically convergent basis sets with relativistic pseudopotentials. II. Small-core pseudopotentials and correlation consistent basis sets for the post-d group 16–18 elements. *The Journal of Chemical Physics*, *119*, 11113–11123.
- Piecuch, P., & Włoch, M. (2005). Renormalized coupled-cluster methods exploiting left eigenstates of the similarity-transformed Hamiltonian. *The Journal of Chemical Physics*, *123*, 224105/1–10.
- Piecuch, P., Kucharski, S. A., Kowalski, K., & Musiał, M. (2002). Efficient computer implementation of the renormalized coupled-cluster methods: The R-CCSD[T], R-CCSD(T), CR-CCSD[T], and CR-CCSD(T) approaches. *Computer Physics Communications*, *149*, 7196.
- Pople, J. A., & Nesbet, R. K. (1954). Self-consistent orbitals for radicals. *The Journal of Chemical Physics*, *22*, 571–572.
- Pople, J. A., Binkley, J. S., & Seeger, R. (1976). Theoretical models incorporating electron correlation. *International Journal of Quantum Chemistry*, *S10*, 1–19.
- Raffenetti, R. C. (1973). General contraction of Gaussian atomic orbitals: Core, valence, polarization, and diffuse basis sets; Molecular integral evaluation. *The Journal of Chemical Physics*, *58*, 4452–4458.
- Read, A. E., & Weinhold, F. (1983). Natural bond orbital analysis of near-Hartree-Fock water dimer. *The Journal of Chemical Physics*, *78*, 4066–4073.
- Read, A. E., Weinstock, R. B., & Weinhold, F. (1985). Natural population analysis. *The Journal of Chemical Physics*, *83*, 735–746.
- Ridley, J. E., & Zerner, M. C. (1973). Intermediate neglect of differential overlap techniques for spectroscopy: Pyrrole and the azines. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling*, *32*, 111–134.
- Roos, B. O. (1983). The multiconfiguration SCF method. In G. H. F. Diercksen & S. Wilson (Eds.), *Methods in computational molecular physics* (pp. 161–187). Dordrecht: Reidel.
- Roos, B. O. (1987). The CASSCF method and its application in electronic structure calculations. *Advances in Chemical Physics*, *69*, 339–445.
- Roos, B. O. (1994). The multiconfiguration SCF theory. In B. O. Roos (Ed.), *Lecture notes in quantum chemistry* (Vol. 58, pp. 177–254). Berlin: Springer.
- Roothaan, C. C. J. (1951). New developments in molecular orbital theory. *Reviews of Modern Physics*, *23*, 69–89.
- Ross, R. B., Powers, J. M., Atashroo, T., Ermler, W. C., LaJohn, L. A., & Christiansen, P. A. (1990). Ab initio relativistic effective core potentials with spin-orbit operators. IV. Cs through Rn. *The Journal of Chemical Physics*, *93*, 6654–6670.
- Roy, L. E., Hay, P. J., & Martin, R. L. (2008). Revised basis sets for the LANL effective core potentials. *Journal of Chemical Theory and Computation*, *4*, 1029–1031.
- Ruedenberg, K., Schmidt, M. W., Gilbert, M. M., & Elbert, S. T. (1982a). Are atoms intrinsic to molecular electronic wavefunctions? I. The FORS model. *Chemical Physics*, *71*, 41–49.
- Ruedenberg, K., Schmidt, M. W., Gilbert, M. M., & Elbert, S. T. (1982b). Are atoms intrinsic to molecular electronic wavefunctions? II. Analysis of FORS orbitals. *Chemical Physics*, *71*, 51–64.
- Ruedenberg, K., Schmidt, M. W., Gilbert, M. M., & Elbert, S. T. (1982c). Are atoms intrinsic to molecular electronic wavefunctions? III. Analysis of FORS configurations. *Chemical Physics*, *71*, 65–78.
- Schirmer, J., & Cederbaum, L. S. (1978). The two-particle-hole Tamm-Dancoff approximation (2ph-TDA) equations for closed-shell atoms and molecules. *Journal of Physics B: Atomic and Molecular Physics*, *11*, 1889–1900.
- Schmidt, M. W., & Gordon, M. S. (1998). The construction and interpretation of MCSCF wavefunctions. *Annual Review of Physical Chemistry*, *49*, 233–266.
- Schmidt, M. W., Baldrige, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S., Windus, T. L., Dupuis, M., & Montgomery, J. J. A. (1993). General atomic and molecular electronic structure system. *Journal of Computational Chemistry*, *14*, 1347–1363.

- Shavitt, I., & Bartlett, R. J. (2009). *Many-body methods in chemistry and physics: MBPT and coupled-cluster theory*. Cambridge, UK: Cambridge University Press.
- Shepard, R. (1987). The MCSCF method. *Advances in Chemical Physics*, 69, 63–200.
- Stevens, W. J., Basch, H., & Krauss, M. (1984). Compact effective potentials and efficient shared-exponent basis sets for the first- and second-row atoms. *The Journal of Chemical Physics*, 81, 6026–6033.
- Stewart, J. J. P. (1989). Optimization of parameters for semiempirical methods. I. Method. *Journal of Computational Chemistry*, 10, 209–220.
- Stewart, J. J. P. (1990). MOPAC: A semiempirical molecular orbital program. *Journal of Computer-Aided Molecular Design*, 4, 1–103.
- Stewart, J. J. P. (2007). Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular Modeling*, 13, 1173–1213.
- Tapia, O., & Goscinski, O. (1975). Self-consistent reaction field-theory of solvent effects. *Molecular Physics*, 29, 1653.
- Tomasi, J., & Persico, M. (1994). Molecular interactions in solution: An overview of methods based on continuous distributions of the solvent. *Chemical Reviews*, 94, 2027–2094.
- Tomasi, J., Mennucci, B., & Cammi, R. (2005). Quantum mechanical continuum solvation models. *Chemical Reviews*, 105, 2999–3093.
- Werner, H. (1987). Matrix formulated direct MCSCF and multiconfiguration reference CI methods. *Advances in Chemical Physics*, 69, 1–62.
- Włoch, M., Gour, J. R., Kowalski, K., & Piecuch, P. (2005). Extension of renormalized coupled-cluster methods including triple excitations to excited electronic states of open-shell molecules. *The Journal of Chemical Physics*, 122, 214107/1–15.
- Yabushita, S., Zhang, Z., & Pitzer, R. M. (1999). Spin-orbit configuration interaction using the graphical unitary group approach and relativistic core potential and spin-orbit operators. *The Journal of Physical Chemistry*, 103, 5791–5800.
- Zeng, T., Fedorov, D. G., Schmidt, M. W., & Klobukowski, M. (2011a). Two-component natural spinors for two-step spin-orbit coupled wave functions. *The Journal of Chemical Physics*, 134, 214107-1–214107-9.
- Zeng, T., Fedorov, D. G., Schmidt, M. W., & Klobukowski, M. (2011b). Effects of spin-orbit coupling on covalent bonding and the Jahn-Teller effect are revealed with the natural language of spinors. *The Journal of Chemical Theory and Computation*, 7, 2864–2875.
- Zerner, M. C. (1991). Semiempirical molecular orbitals methods. In K. B. Lipkowitz & D. B. Boyd (Eds.), *Reviews in computational chemistry* (Vol. 2, pp. 313–365). New York: VCH Publishers.

Paweł Tecmer, Katharina Boguslawski, and Dariusz Kędziera

Contents

Origin of Relativistic Effects and Their Importance in Chemistry	886
Relativistic Quantum Chemistry Made Simpler: Two-Component Hamiltonians	889
Methods Based on Eliminating the Small Component	894
Unitary Transformation of the Dirac Equation	899
A One-Step DKH Method	911
Algebraic Alternatives: The X2C Family of Methods	914
Concluding Remarks on Two-Component Hamiltonians	916
Two-Component Hamiltonians in Practice	917
Where to Find Two-Component Hamiltonians	917
Which Basis Sets to Use for Relativistic Calculations	918
Which Two-Component Hamiltonians to Choose: The Good, the Bad, and the Ugly ...	920
When Do We Need Spin–Orbit Coupling?	921
Concluding Remarks	923
Bibliography	923

Abstract

In this chapter, we briefly discuss the theoretical foundations of relativistic two-component methods used in quantum chemistry calculations. Specifically, we focus on two groups of methods. These are (i) methods based on the elimination of the small component, such as the zeroth-order regular approximation (ZORA),

P. Tecmer (✉) • K. Boguslawski (✉)

Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Toruń, Poland

e-mail: ptecmer@fizyka.umk.pl, ptecmer@gmail.com; k.boguslawski@fizyka.umk.pl

D. Kędziera (✉)

Department of Chemistry and Photochemistry of Polymers, Faculty of Chemistry, Nicolaus Copernicus University, Toruń, Poland

e-mail: teodar@chem.umk.pl

the first-order regular approximation (FORA), and the normalized elimination of small component (NESC) formalisms, and (ii) approaches that use a unitary transformation to decouple the electronic and positronic states such as the Douglas–Kroll–Hess (DKH) and the infinite-order two-component (IOTC) Hamiltonians. Furthermore, we describe the algebraic approach to IOTC and scrutinize pure algebraic schemes that paved the way to the eXact 2-Component (X2C) Hamiltonians taking advantage of the nonsymmetric algebraic Riccati equation (nARE). Finally, we assess the accuracy of the aforementioned methods in calculating core and valence properties of heavy-element compounds and discuss some challenging examples of computational actinide chemistry.

Origin of Relativistic Effects and Their Importance in Chemistry

Computational chemistry aims at modeling the electronic structure of atoms and molecules and at predicting their properties. This task is usually accomplished by solving the electronic Schrödinger equation:

$$\hat{H}_S|\Psi\rangle = E|\Psi\rangle, \quad (1)$$

which describes the nonrelativistic movement of electrons in the field of the nuclei. Solving the above equation gives us the electronic wavefunction $|\Psi\rangle$ of the investigated atom or molecule and the electronic energy E as the target quantity, which is the eigenvalue of the Hamiltonian \hat{H}_S (Kędziera and Kaczmarek-Kędziera 2012) for the eigenvector $|\Psi\rangle$. The (electronic) Hamiltonian \hat{H}_S can be generally partitioned into one- and two-electron terms:

$$\hat{H}_S = \hat{H}_{S,1} + \hat{H}_{S,2}. \quad (2)$$

This common partitioning will be particularly useful when we will “design” relativistic approximations or analyze relativistic effects in the following sections. In Eq. (2), the one-electron part for a molecule comprising M nuclei and N electrons can be further partitioned in two contributions:

$$\hat{H}_{S,1} = - \sum_A^M \sum_i^N \frac{Z_A}{|\hat{\mathbf{r}}_i - \hat{\mathbf{R}}_A|} - \frac{1}{2} \sum_i^N \hat{\nabla}_i^2, \quad (3)$$

where the first term contains the electron–nuclear attraction for nuclei A with charge Z_A and electron–nucleus distance $|\hat{\mathbf{r}}_i - \hat{\mathbf{R}}_A|$ and the second term accounts for the kinetic energy of the electrons. Throughout this chapter, we will use the conventional notation of quantum chemistry, where capital letters indicate nuclei, while small letters are used to index electrons. The two-electron part of Eq. (2) comprises the electron–electron repulsion (Coulomb interaction):

$$\hat{H}_{S,2} = \sum_{i < j}^N \frac{1}{|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|}, \quad (4)$$

where $|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|$ is the electron–electron distance, sometimes abbreviated as $|\hat{r}_{ij}|$ or \hat{r}_{ij} . Both the one- and two-electron Hamiltonian in Eqs. (3) and (4) are written in (Hartree) atomic units (a.u.). This convention will be used through the whole chapter, unless specified otherwise.

Although the nonrelativistic electronic Schrödinger equation (*cf.* Eq. (1)) is a reliable model to theoretically describe the properties of light atoms and molecules, it is an inappropriate model when we move down in the periodic table, toward heavy elements. Well-known examples where the Schrödinger model is insufficient and fails in elucidating the properties of atoms and molecules are the goldish color of gold (unlike most other metals that are silver) and the bond-length shortening in the Au–H molecule, which cannot be explained using nonrelativistic theories (Pyykkö 1988; Autschbach 2012). Some basic ideas on which type of relativistic effects need to be incorporated into nonrelativistic models can be deduced from the Bohr model of hydrogen-like atoms. For that purpose, let us consider Bohr’s formula for the average speed of an electron v_n in hydrogen-like atoms

$$v_n = \frac{Z}{n} \quad (5)$$

and for the average distance of an electron from the nucleus r_n in hydrogen-like atoms

$$r_n = \frac{n^2}{mZ}, \quad (6)$$

where n denotes the main quantum number and Z and m are the charge of the nucleus and the mass of the electron, respectively. When we move toward heavier nuclei, i.e., with increasing charge Z , we increase the average speed of the electron (the numerator increases faster than the denominator in Eq. (5)). For some atomic charge onward, v_n becomes a significant fraction of the speed of light c , which takes on an approximate value of $c \approx 137$ a.u. This is the maximum speed at which all massless particles, light, and gravitational waves travel in vacuum. The considerable increase in the average velocity of the electron in heavy elements is illustrated in the following example.

Example 1. For the hydrogen atom ($Z_A = 1$ and $n = 1$), the average speed of the electron is $v_n = 1$ a.u.¹ For the 1s electron in the C^{5+} cation ($Z_A = 6$ and $n = 1$), we have $v_n = 6$ a.u., which is approximately 4% of the speed of light c , while for

¹The atomic unit of speed is defined as $\frac{a_0 E_h}{\hbar} = \alpha c$, which corresponds to $2.18769126277 \times 10^6 \frac{\text{m}}{\text{s}}$.

the U^{91+} cation, the average speed of the 1 s electron increases to $v_n \approx 92$ a.u. and reaches approximately 67 % of c .

The increasing average speed of the electron requires modification of the electron mass in accordance with the theory of relativity (Dyall and Faegri 2007; Reiher and Wolf 2014):

$$m = \frac{m_0}{\sqrt{1 - \frac{v_n^2}{c^2}}}, \quad (7)$$

where m_0 is the rest mass, i.e., the mass of the particle at rest. The modified value for the electron mass will change the average distance of the electron from the nucleus as given in Eq. (6). Substituting the above equation for the electron mass into Eq. (6) and inserting Eq. (5) for the average speed of an electron lead to the following relativistic equation for r_n :

$$r_n^{\text{rel}} = \frac{n^2 \sqrt{1 - \frac{Z^2}{n^2 c^2}}}{m_0 Z}. \quad (8)$$

In nonrelativistic quantum mechanics, the electron mass equals the rest mass ($m = m_0$) and Eq. (6) remains valid. Since $v_n < c$, the denominator in Eq. (7) will always be less than or equal to 1. Thus, the electron mass is always greater or equal to the rest mass and increases with increasing v_n . Therefore, we obtain the following inequality:

$$r_n > r_n^{\text{rel}}, \quad (9)$$

which explains the origin of the shortening (contraction) of bond lengths in heavy-element-containing molecules.

In nonrelativistic theory, the average speed of a 1 s electron is proportional to the atomic charge Z (see Eq. (5)). If we want to anticipate the importance of relativistic effects in heavy elements, we can simply calculate the ratio $\frac{v_n}{c}$ for the 1 s-shell.

Example 2. The average speed of the 1 s electron in the uranium atom ($Z = 92$) equals 67 % of c ($\frac{92}{137} = 0.67$). Using Eq. (7), the mass of the (moving) 1 s electron thus becomes $\approx 1.35m_0$. This corresponds to a radial shrinkage of 25 % (compare Eqs. (6) and (8)).

These simple examples illustrate that nonrelativistic quantum mechanics based on the Schrödinger equation is inadequate for heavy elements like the uranium atom. Applications of relativistic quantum chemistry explained many phenomena observed in heavy-element chemistry that could not be understood using nonrelativistic quantum mechanics. These are the contraction of s and p orbitals, the expansion of d and f orbitals, and singlet-triplet excitations with

nonzero intensities (due to spin–orbit coupling). There are thus numerous examples that demonstrate that heavy elements have to be described using relativistic quantum mechanics, i.e., using the Dirac equation (Dirac 1928a, b) instead of the Schrödinger picture.

Relativistic Quantum Chemistry Made Simpler: Two-Component Hamiltonians

Including relativistic corrections in the theoretical model improves a quantum-mechanical description of heavy elements compared to nonrelativistic methods. Unfortunately, a reliable relativistic description of quantum systems is not a straightforward task and remains a remarkable challenge. The technical and mathematical difficulties that come along with a fully relativistic treatment of the movement of electrons will be addressed in the following sections.

By analogy to nonrelativistic theory, the relativistic Hamiltonian \hat{H}_R can be written as a sum of one- and two-electron parts:

$$\hat{H}_R = \hat{H}_{R,1} + \hat{H}_{R,2}. \quad (10)$$

The one-electron part $\hat{H}_{R,1}$ is simply a sum of the one-electron Dirac Hamiltonian \hat{H}_D for all electrons in the atom or molecule. For one electron, the Dirac Hamiltonian \hat{H}_D can be written in closed form. For the H atom, we have

$$\hat{H}_D = \hat{\beta}c^2 + \hat{V} + c\hat{\alpha} \cdot \hat{p}, \quad (11)$$

where c indicates the velocity of light, $\hat{p} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$ is the momentum operator, and \hat{V} is the Coulomb potential (electron–nuclear interaction). In addition, we also have a vector $\hat{\alpha}$ with elements $(\hat{\alpha}_x, \hat{\alpha}_y, \hat{\alpha}_z)$ as well as $\hat{\beta}$, all being 4×4 matrices. It is important to stress that according to Dirac’s theory (Dyall and Faegri 2007; Reiher and Wolf 2014), the matrices $\hat{\alpha}_i$ and $\hat{\beta}$ have to fulfill the following conditions:

$$\hat{\alpha}_i^2 = \hat{\beta}^2 = \hat{I}_4 \quad (12)$$

and

$$\hat{\alpha}_i\hat{\alpha}_j + \hat{\alpha}_j\hat{\alpha}_i = \hat{\alpha}_i\hat{\beta} + \hat{\beta}\hat{\alpha}_i = \hat{0}_4, \quad \text{for } i \neq j, \quad (13)$$

where \hat{I}_4 and $\hat{0}_4$ are the 4×4 identity and zero matrices, respectively.

Among possible choices of $\hat{\alpha}_i$ and $\hat{\beta}$, the most common ones are

$$\hat{\alpha}_i = \begin{pmatrix} 0 & \hat{\sigma}_i \\ \hat{\sigma}_i & 0 \end{pmatrix}, \quad \hat{\beta} = \begin{pmatrix} \hat{I}_2 & 0 \\ 0 & -\hat{I}_2 \end{pmatrix}, \quad (14)$$

where σ_i are the (two-dimensional) Pauli spin matrices

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (15)$$

It is convenient to use the abovementioned structures of $\hat{\alpha}_i$ and $\hat{\beta}$ and rewrite the Dirac Hamiltonian \hat{H}_D as

$$\hat{H}_D = \begin{pmatrix} (\hat{V} + c^2)\hat{I}_2 & c\hat{\sigma} \cdot \hat{\mathbf{p}} \\ c\hat{\sigma} \cdot \hat{\mathbf{p}} & (\hat{V} - c^2)\hat{I}_2 \end{pmatrix}, \quad (16)$$

with \hat{I}_2 being the 2×2 identity matrix. This leads to the following eigenvalue problem:

$$\begin{pmatrix} \hat{V}\hat{I}_2 & c\hat{\sigma} \cdot \hat{\mathbf{p}} \\ c\hat{\sigma} \cdot \hat{\mathbf{p}} & (\hat{V} - 2c^2)\hat{I}_2 \end{pmatrix} \begin{pmatrix} \psi^L(\mathbf{r}) \\ \psi^S(\mathbf{r}) \end{pmatrix} = \epsilon \begin{pmatrix} \psi^L(\mathbf{r}) \\ \psi^S(\mathbf{r}) \end{pmatrix}, \quad (17)$$

known as the four-component form of the *Dirac equation*, where

$$\epsilon = E - c^2 \quad (18)$$

denotes the adjusted (scaled) Dirac energy. The wavefunction resulting from Eq. (17),

$$\psi(\mathbf{r}) = \begin{pmatrix} \psi^L(\mathbf{r}) \\ \psi^S(\mathbf{r}) \end{pmatrix}, \quad (19)$$

after taking the normalization condition,

$$\langle \psi(\mathbf{r}) | \psi(\mathbf{r}) \rangle = \langle \psi^L(\mathbf{r}) | \psi^L(\mathbf{r}) \rangle + \langle \psi^S(\mathbf{r}) | \psi^S(\mathbf{r}) \rangle = 1, \quad (20)$$

has a four-component structure and differs therefore from the nonrelativistic wavefunction.

The two-component elements in Eq. (19) are called the large $\psi^L(\mathbf{r})$ and small $\psi^S(\mathbf{r})$ component, respectively. This naming convention originates from the fact that, for the electronic solutions, the large component has a larger norm than the small component. We refer the interested reader to relativistic quantum chemistry textbooks for more details (see, e.g., Schwerdtfeger (2002), Dyall and Faegri (2007), and Reiher and Wolf (2014)). Here, we mention only two key features of the Dirac equation. First, the Dirac equation describes the behavior of both electrons and positrons (the counterpart of the electron with an opposite charge). Second, with the $\hat{\sigma}$ matrices embedded into the Dirac equation, the electron spin emerges naturally, while in nonrelativistic quantum mechanics, the existence of the electron spin is

postulated and included a posteriori into the model. In the following, we will drop the “ $\hat{}$ ” sign above all operators for simplicity. The notation will be evident from the context.

Let us have a closer look at Eq. (17). We can rewrite Eq. (17) as a set of two coupled equations:

$$V\psi^L(\mathbf{r}) + c(\boldsymbol{\sigma} \cdot \mathbf{p})\psi^S(\mathbf{r}) = \epsilon\psi^L(\mathbf{r}) \quad (21)$$

$$c(\boldsymbol{\sigma} \cdot \mathbf{p})\psi^L(\mathbf{r}) + (V - 2c^2)\psi^S(\mathbf{r}) = \epsilon\psi^S(\mathbf{r}) \quad (22)$$

Now, only consider Eq. (22), which can be written as (using matrix–vector multiplication and neglecting all magnetic fields),

$$c(\boldsymbol{\sigma} \cdot \mathbf{p})\psi^L(\mathbf{r}) - 2c^2\psi^S(\mathbf{r}) + V\psi^S(\mathbf{r}) = \epsilon\psi^S(\mathbf{r}). \quad (23)$$

Bringing all terms containing the small component on one side and the remaining terms on the other side yields

$$(2c^2 + \epsilon - V)\psi^S(\mathbf{r}) = c(\boldsymbol{\sigma} \cdot \mathbf{p})\psi^L(\mathbf{r}). \quad (24)$$

From the above equation, we can deduce a relation between the large and small component. After multiplying both sides by $\frac{1}{2c^2}$ and dividing by the prefactor of $\psi^S(\mathbf{r})$, we get a mathematical relation between the small and large component function, the so-called *atomic balance*,

$$\psi^S(\mathbf{r}) = \frac{1}{2c} \frac{1}{1 - \frac{V-\epsilon}{2c^2}} (\boldsymbol{\sigma} \cdot \mathbf{p})\psi^L(\mathbf{r}). \quad (25)$$

Unfortunately, we are dealing with finite-size basis sets in quantum chemistry and this relation cannot be strictly fulfilled. As a consequence, difficulties with the representation of the operator $\boldsymbol{\sigma} \cdot \mathbf{p}$ may occur, which can further lead to unphysical energies lying below the Dirac energy levels. To illustrate this issue, it is worth to analyze the formula resulting from substituting Eq. (25) into Eq. (21),

$$\left(V + \frac{1}{2}(\boldsymbol{\sigma} \cdot \mathbf{p}) \frac{1}{1 - \frac{V-\epsilon}{2c^2}} (\boldsymbol{\sigma} \cdot \mathbf{p}) \right) \psi^L(\mathbf{r}) = \epsilon\psi^L(\mathbf{r}). \quad (26)$$

Equation (26) is a pseudo-eigenvalue equation with a structure similar to the Schrödinger equation written in a basis of Pauli spinors,

$$(V + \frac{1}{2}(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{p}))\psi_{\text{nr}}(\mathbf{r}) = E_{\text{nr}}\psi_{\text{nr}}(\mathbf{r}), \quad (27)$$

where the subscript nr denotes the nonrelativistic (but two-component) form of the wavefunction. The similarity between the relativistic and the nonrelativistic

framework will become more evident after introducing the so-called Dirac's relation (Reiher and Wolf 2014):

$$\frac{(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{p})}{2} = \frac{p^2}{2}. \tag{28}$$

Now, if we take the limit $c \rightarrow \infty$, we should obtain two identical equations, that is, Eq. (26) should be equal to Eq. (27). Unfortunately, this is not the case as we need to change the analytical derivation into its algebraic counterpart (algebraization of the problem) by introducing some (approximate) expansion for the expression between the $\boldsymbol{\sigma} \cdot \mathbf{p}$ operators in Eq. (26) (i.e., introducing a finite basis set). By using orthonormal basis functions approximating $\psi^L(\mathbf{r})$ and $\psi^S(\mathbf{r})$,

$$\psi^L(\mathbf{r}) \approx |\chi^L\rangle C^L, \tag{29}$$

$$\psi^S(\mathbf{r}) \approx |\chi^S\rangle C^S, \tag{30}$$

we obtain the algebraic representation of Eq. (26),

$$\langle \chi^L | \left(V + \frac{1}{2}(\boldsymbol{\sigma} \cdot \mathbf{p})|\chi^S\rangle\langle \chi^S| \frac{1}{1 - \frac{V-\epsilon}{2c^2}}|\chi^S\rangle\langle \chi^S|(\boldsymbol{\sigma} \cdot \mathbf{p}) \right) | \chi^L \rangle C^L = \langle \chi^L | \chi^L \rangle C^L \epsilon, \tag{31}$$

which after taking the limit $c \rightarrow \infty$ results in

$$\langle \chi^L | \left(V + \frac{1}{2}(\boldsymbol{\sigma} \cdot \mathbf{p})|\chi^S\rangle\langle \chi^S|(\boldsymbol{\sigma} \cdot \mathbf{p}) \right) | \chi^L \rangle C^L = \langle \chi^L | \chi^L \rangle C^L \epsilon. \tag{32}$$

The problem with the nonrelativistic limit can be thus resolved if $\langle \chi^L | p^2 | \chi^L \rangle$ will equal $\langle \chi^L | \boldsymbol{\sigma} \cdot \mathbf{p} | \chi^S \rangle \langle \chi^S | \boldsymbol{\sigma} \cdot \mathbf{p} | \chi^L \rangle$. This equality is true if the basis $|\chi^S\rangle$ spans the whole space of small component functions, which, however, is difficult to achieve in calculations. In most cases, the aforementioned equality is not true.

If we introduce the complement of the small component basis $|\chi^S\rangle$ denoted as $|\bar{\chi}^S\rangle$, we can recover the equality relation,

$$\langle \chi^L | p^2 | \chi^L \rangle = \langle \chi^L | \boldsymbol{\sigma} \cdot \mathbf{p} | \chi^S \rangle \langle \chi^S | \boldsymbol{\sigma} \cdot \mathbf{p} | \chi^L \rangle + \langle \chi^L | \boldsymbol{\sigma} \cdot \mathbf{p} | \bar{\chi}^S \rangle \langle \bar{\chi}^S | \boldsymbol{\sigma} \cdot \mathbf{p} | \chi^L \rangle. \tag{33}$$

Since the operator $\boldsymbol{\sigma} \cdot \mathbf{p}$ is Hermitian, the matrix $\langle \chi^L | \boldsymbol{\sigma} \cdot \mathbf{p} | \bar{\chi}^S \rangle \langle \bar{\chi}^S | \boldsymbol{\sigma} \cdot \mathbf{p} | \chi^L \rangle$ is positive semi-definite. The incompleteness of the $|\chi^S\rangle$ basis (i.e., the missing second terms on the right-hand side of Eq. (33)) leads to

$$\langle \chi^L | \boldsymbol{\sigma} \cdot \mathbf{p} | \chi^S \rangle \langle \chi^S | \boldsymbol{\sigma} \cdot \mathbf{p} | \chi^L \rangle \leq \langle \chi^L | p^2 | \chi^L \rangle. \tag{34}$$

And thus, even after taking the nonrelativistic limit, the resulting Dirac energies might be artificially lower than the energies obtained from the Schrödinger equation which can be attributed to lower (relativistic) kinetic energies.

Therefore, it is particularly important in relativistic quantum chemistry calculations to use basis sets that are suitable for the large component (which should lead to nonrelativistic functions in the nonrelativistic limit) but that also allow us to represent the small component in a well-balanced way. How can we ensure that the small component functions are well balanced in practical calculations? This is not an easy task, primarily because the small and large component functions have to be of different type. For instance, the large component s -type functions have to balance the small component p -type functions, while functions of p -type have to balance functions of s - and d -type, functions of d -type have to balance functions of p - and f -type, etc. Satisfying these conditions results in large basis set sizes, typically much larger than in nonrelativistic calculations, which makes calculations susceptible to linear dependencies in the basis set. These difficulties motivate us to eliminate the small component functions right from the start. In the following, we will describe different methods that aim at solving the Dirac equation using large component functions only. Before scrutinizing the so-called two-component methods, we will briefly discuss the two-electron part of the relativistic Hamiltonian.

The exact (Lorentz-invariant) form of the operator for the electron–electron interaction is unknown and we are forced to use inaccurate approximations. Two approximations are commonly used in relativistic quantum chemistry. These are (i) the standard nonrelativistic two-electron operator as defined in Eq. (4), which is added to the relativistic one-electron term and leads to the Dirac–Coulomb (DC) Hamiltonian, and (ii) the Coulomb term augmented by the Breit operator:

$$O(i, j)^{\text{Breit}} = -\frac{c\boldsymbol{\alpha}_i \cdot c\boldsymbol{\alpha}_j}{2c^2 r_{ij}} - \frac{(c\boldsymbol{\alpha}_i \mathbf{r}_{ij}) \cdot (c\boldsymbol{\alpha}_j \mathbf{r}_{ij})}{2c^2 r_{ij}^3}. \quad (35)$$

Including the Breit operator into the two-electron part of the Hamiltonian leads to the Dirac–Coulomb–Breit (DCB) Hamiltonian (Autschbach 2012),

$$H_{R,2}^{\text{DCB}} = \sum_{i < j}^N \left(\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + O(i, j)^{\text{Breit}} \right), \quad (36)$$

which, in addition to the charge–charge interaction (through the Coulomb term), also accounts for the current–current interaction. Although the nonrelativistic Coulomb interaction is straightforward to evaluate compared to the DCB expression, the resulting DC Hamiltonian is not even close to being Lorentz invariant. Adding the Breit interaction results in a Hamiltonian that is *approximately* Lorentz invariant (Reiher and Wolf 2014). Yet, the approximate Lorentz invariance of the DCB Hamiltonian is usually sufficient for chemical applications.

In summary, solving the four-component Dirac equation is difficult, primarily because of the desire to have large and small component functions that are well balanced. Furthermore, the necessity to include both small and large component basis functions at the same time significantly increases the computational cost. On the other hand, when solving the Dirac equation, we obtain solutions for both

electrons and positrons. While the former are of interest in chemistry, the latter are irrelevant for chemical purposes. Furthermore, the Dirac equation (either written in its full four-component form or abbreviated by only two components) poses technical problems for conventional (usually nonrelativistic) quantum chemistry software packages, primarily because the Dirac Hamiltonian differs so much from the Schrödinger Hamiltonian that quantum chemistry codes require a fundamental rewrite to facilitate implementing the Dirac equation. Thus, the Dirac Hamiltonian is usually not supported in mainstream software packages, and relativistic effects are typically included as corrections to the Schrödinger Hamiltonian. Such corrections can be easily implemented in nonrelativistic software packages. In the following, we will discuss some popular approaches that aim at “simplifying” the Dirac Hamiltonian in a form that can be handled by nonrelativistic codes. All methods presented in this chapter transform the four-component Dirac equation into a Hamiltonian with at most two components (Mastalerz and Reiher 2008; Reiher 2012; Wolf et al. 2002; Kędziera and Barysz 2004; Reiher 2006; Liu 2010; Saue 2011).

Methods Based on Eliminating the Small Component

One family of methods tries to simplify the (four-component) Dirac equation by eliminating the small component (ESC) $\psi^S(\mathbf{r})$ from Eq. (17), thus resulting in an eigenvalue equation for the large component $\psi^L(\mathbf{r})$ only. To eliminate the small component from the Dirac equation, we start with the atomic balance relation between the small and large component functions (Eq. (25), but slightly rewritten):

$$\psi^S(\mathbf{r}) = \left(1 + \frac{\epsilon - V}{2c^2}\right)^{-1} \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{2c} \psi^L(\mathbf{r}), \quad (37)$$

which has the nonrelativistic limit, i.e., the limit for $c \rightarrow \infty$,

$$\lim_{c \rightarrow \infty} 2c \psi^S(\mathbf{r}) = (\boldsymbol{\sigma} \cdot \mathbf{p}) \psi^L(\mathbf{r}). \quad (38)$$

The first factor in Eq. (37) is a multiplicative operator that depends on the energy (the eigenvalue of the Dirac equation) and the position of the electron encoded in the potential V . It is convenient for us to introduce an abbreviation for this operator

$$K(\epsilon, \mathbf{r}) = \left(1 + \frac{\epsilon - V}{2c^2}\right)^{-1}. \quad (39)$$

The relation between the large and small component of Eq. (37) as well as the multiplicative operator $K(\epsilon, \mathbf{r})$ will serve as a starting point for approximations to the (four-component) Dirac equation. Specifically, Eq. (37) suggests that, given a solution for the large component $\psi^L(\mathbf{r})$, we can immediately determine a solution for the small component $\psi^S(\mathbf{r})$. Hence, we can *eliminate* the small component from

the Dirac equation and solve the resulting equations only for the large component. Substituting the relation for the small component $\psi^S(\mathbf{r})$ (Eq. (37)) into Eq. (21), we arrive at a *two-component* equation for the large component of the wavefunction $\psi^L(\mathbf{r})$,

$$\left(\frac{1}{2}(\boldsymbol{\sigma} \cdot \mathbf{p})K(\epsilon, \mathbf{r})(\boldsymbol{\sigma} \cdot \mathbf{p}) + V \right) \psi^L(\mathbf{r}) = \epsilon \psi^L(\mathbf{r}). \quad (40)$$

This will be our starting point for deriving approximate Hamiltonians that aim at eliminating the small component.

To solve Eq. (40) for the large component, we need to calculate the expression for the multiplicative operator $K(\epsilon, \mathbf{r})$, which is the inverse of a function containing the energy and some potential. In principle, this can be accomplished using a series expansion in the parameter $\frac{\epsilon - V}{2c^2}$. Specifically, K can be rewritten in terms of a geometric series,²

$$K(\epsilon, \mathbf{r}) = \left(1 + \frac{\epsilon - V}{2c^2} \right)^{-1} = \sum_{k=0}^{\infty} \left(\frac{V - \epsilon}{2c^2} \right)^k. \quad (41)$$

However, such a procedure will only work if the expansion parameter $\frac{\epsilon - V}{2c^2}$ is small. Since the denominator ($2c^2$) is a constant, the validity of the series expansion depends on the energy difference $\epsilon - V$ in the numerator. For valence electrons, the numerator will be much smaller than $2c^2$ and the expansion will be still valid. However, in the case of electrons that are close in position to the nuclei, V increases significantly in magnitude, the series expansion breaks down, and the resulting operator expression becomes invalid. This problem implicates singularities in the approximated K function (i.e., the series expansion) near the nuclei. To remedy the problem of singularities and to arrive at a variationally stable operator, we have to find a series expansion that is valid for the whole region of space, that is, for electrons in the valence shell as well as core electrons. One possible choice is the regular approximation (RA) introduced by Chang et al. (1986), which was later popularized by van Lenthe et al. (1994) within the framework of density functional theory. The basic idea of the regular approximation is to rewrite the expression for the K operator Eq. (39) as

$$K(\epsilon, \mathbf{r}) = \left(1 + \frac{\epsilon - V}{2c^2} \right)^{-1} = \left(1 - \frac{V}{2c^2} \right)^{-1} \left(1 + \frac{\epsilon}{2c^2 - V} \right)^{-1} \quad (42)$$

and to choose $\frac{\epsilon}{2c^2 - V}$ as the expansion parameter. Since the numerator ϵ is constant and always much smaller than the denominator ($2c^2 - V$),³ this parameter is small

²The geometric series is defined as $\frac{a}{1-x} = \sum_{k=0}^{\infty} ax^k$, for $|x| < 1$.

³Note that V is negative and hence the denominator is even larger than $2c^2$.

for all regions in space and the (geometric) series expansion is well defined (no singularities occur). Expanding the second factor of Eq. (42) using the geometric series, the K operator becomes

$$K(\epsilon, \mathbf{r}) = \left(1 - \frac{V}{2c^2}\right)^{-1} \sum_{k=0}^{\infty} \left(\frac{\epsilon}{V - 2c^2}\right)^k. \quad (43)$$

After inserting the above series expansion for the K operator into Eq. (40), we obtain a general expression for the regular approximation,

$$\left(\frac{1}{2}(\boldsymbol{\sigma} \cdot \mathbf{p}) \left[\left(1 - \frac{V}{2c^2}\right)^{-1} \sum_{k=0}^{\infty} \left(\frac{\epsilon}{V - 2c^2}\right)^k \right] (\boldsymbol{\sigma} \cdot \mathbf{p}) + V\right) \psi^L(\mathbf{r}) = \epsilon \psi^L(\mathbf{r}). \quad (44)$$

In practice, we cannot include all terms in the series expansion (there are infinitely many), and it is a common procedure to truncate the series expansion after a given order to approximate the relativistic Hamiltonian. This will result in different flavors of the regular approximation. The most popular approaches include only the first as well as the first and second term of the expansion in Eq. (41).

The most popular and also most simple flavor of the regular approximation includes only the first term of the series expansion (van Lenthe et al. 1996). Since the first term of the sum in Eq. (43) is for $k = 0$, this approximation is called the zeroth-order regular approximation, abbreviated as ZORA (Chang et al. 1986; van Lenthe et al. 1994). Keeping only the first term of Eq. (43), which is equal to 1, the K operator becomes

$$K(\epsilon, \mathbf{r}) = \left(1 - \frac{V}{2c^2}\right)^{-1}. \quad (45)$$

Inserting the above equation into Eq. (40), we obtain the so-called ZORA equation for the large component $\psi_{\text{ZORA}}^L(\mathbf{r})$,

$$\left(\frac{1}{2}(\boldsymbol{\sigma} \cdot \mathbf{p}) \left(1 - \frac{V}{2c^2}\right)^{-1} (\boldsymbol{\sigma} \cdot \mathbf{p}) + V\right) \psi_{\text{ZORA}}^L(\mathbf{r}) = \epsilon \psi_{\text{ZORA}}^L(\mathbf{r}), \quad (46)$$

where the expression in parentheses is the ZORA Hamiltonian. In practical calculations, one usually employs the *scalar relativistic* version of the ZORA equation, where the spin-orbit interaction is not taken into account, which reads

$$\left(\frac{1}{2}\mathbf{p} \left(1 - \frac{V}{2c^2}\right)^{-1} \mathbf{p} + V\right) \psi_{\text{ScZORA}}^L(\mathbf{r}) = \epsilon \psi_{\text{ScZORA}}^L(\mathbf{r}), \quad (47)$$

where we have used Dirac's relation Eq. (28).

The advantages of ZORA are the normalization of the ZORA wavefunction $\psi_{\text{ZORA}}^L(\mathbf{r})$ and the absence of the energy ϵ in the denominator of the approximate K operator. Although ZORA is the simplest form of the regular approximation, it can account for the most important relativistic effects in the region of the nucleus. Furthermore, it excellently reproduces valence shell properties compared to the fully relativistic Dirac equation.

The disadvantage of the ZORA Hamiltonian is the *gauge invariance* problem. If the potential V is changed by adding a constant shift, i.e., some constant background potential C , so that the new potential reads $V' = V + C$, it does not shift the energy uniformly by this constant C , as it, however, should. This gauge dependence crucially affects properties of core orbitals, for instance, X-ray spectroscopy, but also the optimization of molecular structures. There are, however, ways to remedy the gauge invariance problem (scaling procedures, introduction of model potentials, etc.), which make ZORA a very powerful and efficient relativistic approximation to optimize structures and predict properties of heavy-element compounds (van Lenthe et al. 1993, 1999).

The ZORA Hamiltonian can be improved by including higher-order terms in the series expansion of the K operator (van Lenthe et al. 1994, 1993; Dyall and van Lenthe 1999). If the expansion is truncated after the second term, i.e., $k = 1$, we obtain the so-called first-order regular approximation (FORA), which can be considered as a first-order correction to the ZORA Hamiltonian. Keeping only the first two terms in Eq. (43), the K operator of FORA reads

$$K(\epsilon, \mathbf{r}) = \left(1 - \frac{V}{2c^2}\right)^{-1} \left(1 + \frac{\epsilon}{V - 2c^2}\right), \quad (48)$$

while the FORA equation for the large component $\psi_{\text{FORA}}^L(\mathbf{r})$ takes on the form

$$\left(\frac{1}{2}(\boldsymbol{\sigma} \cdot \mathbf{p}) \left[\left(1 - \frac{V}{2c^2}\right)^{-1} \left(1 + \frac{\epsilon}{V - 2c^2}\right) \right] (\boldsymbol{\sigma} \cdot \mathbf{p}) + V\right) \psi_{\text{FORA}}^L(\mathbf{r}) = \epsilon \psi_{\text{FORA}}^L(\mathbf{r}). \quad (49)$$

In contrast to the ZORA Hamiltonian, which can be used variationally, the FORA Hamiltonian is unbounded from below and can thus only be applied in a perturbative manner. The FORA energy contains terms that are missing from the ZORA Hamiltonian and form a relativistic correction to the ZORA energy. Specifically, the first-order correction term of FORA results in an energy that is correct to $\mathcal{O}(c^{-2})$ and hence removes the $\mathcal{O}(c^{-2})$ error in the ZORA energy (Dyall and van Lenthe 1999; Visscher 2013).

Further improvements over the ZORA and FORA Hamiltonians can be achieved by the infinite-order regular approximation (IORA) (Dyall and van Lenthe 1999; Filatov 2002).

Going back to our starting point when deriving the regular approximation, Eq. (40) is known as the unnormalized elimination of the small component (UESC)

equation.⁴ Although being exact, UESC is not a practical approach, primarily because the energy ϵ appears on both sides of the equation making UESC not an eigenvalue problem (like the Dirac equation). This results from using the strict *atomic balance* relation that depends on the energy. We can remedy this problem by replacing the small component of the wavefunction with a *pseudo-large component* $\phi^L(\mathbf{r})$,

$$\psi^S(\mathbf{r}) = \frac{\boldsymbol{\sigma} \cdot \mathbf{p}}{2c} \phi^L(\mathbf{r}). \quad (50)$$

The above equation can be understood as a modification of Eq. (37). The non-relativistic limit of the pseudo-large component is the large component; note that Eq. (37) has the same limit as in the case of the pseudo-large component:

$$\lim_{c \rightarrow \infty} 2c \psi^S(\mathbf{r}) = (\boldsymbol{\sigma} \cdot \mathbf{p}) \phi^L(\mathbf{r}). \quad (51)$$

Inserting the relation for the pseudo-large component in the two-component form of the Dirac equation (Eq. (17)) and using Dirac's relation given in Eq. (28) will lead us to

$$(V - \epsilon) \psi^L(\mathbf{r}) + \frac{p^2}{2} \phi^L(\mathbf{r}) = 0 \quad (52)$$

$$\frac{p^2}{2} \psi^L(\mathbf{r}) + \left(\frac{1}{4m^2 c^2} (\boldsymbol{\sigma} \cdot \mathbf{p})(V - \epsilon)(\boldsymbol{\sigma} \cdot \mathbf{p}) - \frac{p^2}{2} \right) \phi^L(\mathbf{r}) = 0. \quad (53)$$

The above set of equations are known as the modified Dirac equation (Dyall 1994), a coupled second-order differential equation for the two components of the modified Dirac wavefunction ($\psi^L(\mathbf{r})$, $\phi^L(\mathbf{r})$). Note that Eq. (52) has a structure similar to the nonrelativistic Schrödinger equation. The advantage of the modified Dirac equation is that $\psi^L(\mathbf{r})$ and $\phi^L(\mathbf{r})$ have the same parity, so they can be expanded in the basis set of the same type. This is beneficial in practical calculations since only one basis set for both components is used. The modified Dirac equation (Dyall 1994) was a big step forward in the development of relativistic methods. Equations (52) and (53) form the starting point for eliminating the small (here, pseudo-large) component. An elimination recipe has been developed by Dyall (1997, 1998) and is known as the normalized elimination of the small component (NESC). Specifically, the NESC method gives energy eigenvalues that are correct to $\mathcal{O}(c^{-4})$ with respect to the Dirac eigenvalue, while UESC is correct to only $\mathcal{O}(c^{-2})$. Besides, NESC is free from singularities, which are a common problem for unnormalized methods

⁴The term unnormalized refers to the normalization property of the large component of the wavefunction, which does not fulfill the normalization condition $\int |\psi^L|^2 d\mathbf{r} \neq 1 = \int |\psi^L|^2 d\mathbf{r} + \int |\psi^S|^2 d\mathbf{r}$.

(Zou et al. 2011) (A detailed description of the UESC method can be found in the book of Dyall and Faegri (2007).).

Before we move on to other two-component methods, we should emphasize that Eqs. (52) and (53) are equivalent to the original Dirac Hamiltonian transformed by a matrix

$$U = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\sigma \cdot \mathbf{p}}{2c} \end{pmatrix}. \quad (54)$$

In general, transformations of this kind can be used to derive equations with a more convenient structure. In the following, we will focus on other the two-component methods that are based on a unitary transformation and decompose the Dirac Hamiltonian into a block diagonal form. For better readability, we will omit the explicit dependence of the wavefunction ψ on \mathbf{r} and simply write ψ instead of $\psi(\mathbf{r})$.

Unitary Transformation of the Dirac Equation

When solving the four-component Dirac equation (*cf.* Eq.(17)), it would be advantageous to use as much as possible the standard techniques in nonrelativistic quantum chemistry. Unfortunately, the presence of the positronic states makes it impossible to blindly use the variational principle without any precautions. As positronic states are much lower in energy than electronic states and unbound, i.e., no energy minimum exists, minimizing the energy without any constraints would result in a *variational collapse* (Brown and Ravenhall 1951). Such a collapse would be avoided if we managed to fully *decouple* the electronic and positronic solutions, yielding a set of equations for the electronic states only. This is desirable in quantum chemistry, primarily because we are interested in electronic states only and wish to ignore the positronic states. A decoupling of electronic and positronic solutions can be accomplished by a *unitary transformation* \mathcal{U} of the Dirac Hamiltonian,

$$\tilde{H}_D = \mathcal{U}^\dagger H_D \mathcal{U} = \begin{pmatrix} h_+ & 0 \\ 0 & h_- \end{pmatrix}, \quad (55)$$

that makes the Dirac Hamiltonian H_D block diagonal with respect to the large and small component. In the above equation, we have denoted the unitary transformed Hamiltonian as \tilde{H}_D , while h_+ and h_- are two-component Hamiltonians that act only on electronic and positronic states, respectively. The unitary transformation also modifies the four-component Dirac wavefunction and transforms it into a wavefunction with only two components (for the electronic part only),

$$\mathcal{U}^\dagger \begin{pmatrix} \psi^L \\ \psi^S \end{pmatrix} = \begin{pmatrix} \phi_+ \\ 0 \end{pmatrix}, \quad (56)$$

where ϕ_+ is an eigenfunction of the two-component Hamiltonian h_+ . Thus the unitary transformation separates the large and small component of the untransformed Dirac wavefunction and converts the Dirac Hamiltonian into two separate eigenvalue problems, one for the electronic states (our *target* states) and the positronic states (the ones we wish to ignore). Combining Eqs. (55) and (56), the movement of the electrons is then described by the following equation

$$h_+ \phi_+ = E_+ \phi_+. \quad (57)$$

Note that h_+ is a two-dimensional matrix, while ϕ_+ is a two-dimensional vector. Thus, the above equation has two components. We should emphasize that all approaches mentioned above (methods based on elimination of the small component and methods based on a unitary transformation) aim at eliminating the small component from the electronic part of the Dirac equation. Furthermore, these two kinds of approaches are equivalent in the limit of infinite expansion used in ESC or RA.

Heully and coworkers (1986) proposed a particularly useful form for the transformation \mathcal{U} ,

$$\mathcal{U} = \begin{pmatrix} \Omega_+ & R^\dagger \Omega_- \\ R \Omega_+ & -\Omega_- \end{pmatrix}, \quad (58)$$

where

$$\Omega_+ = (1 + R^\dagger R)^{-1/2}, \quad \Omega_- = (1 + R R^\dagger)^{-1/2}. \quad (59)$$

The structure of Eq. (58) enforces operator R to be unitary, independent of its form, except that the condition

$$R^\dagger = -R \quad (60)$$

must be satisfied. The R operator is determined by requiring the off-diagonal elements of \bar{H}_D to be equal to zero, thus ensuring that the transformed Hamiltonian \bar{H}_D is block diagonal,

$$R(H_1)_{11} + R(H_1)_{12}R - (H_1)_{21} - (H_1)_{22}R = 0. \quad (61)$$

Unfortunately, neither the analytical solution of the above equation nor the form of the unitary transformation \mathcal{U} is known. Finding reliable approximations to R is one of the central problems in the relativistic quantum chemistry community. Even though we do not know the analytical form of R , we can analyze its physical meaning. Substituting Eq. (58) into Eq. (56), we obtain

$$\phi^L = \Omega_+(\psi^L + R^\dagger \psi^S), \quad 0 = \Omega_-(R\psi^L - \psi^S), \quad (62)$$

where the second equation is only satisfied if

$$\psi^S = R\psi^L. \quad (63)$$

Thus, R can be understood as an operator converting the large component of electronic solutions into its small component counterpart. Moreover, using Eq. (63), the normalized two-component electronic function ϕ^L simplifies to

$$\phi^L = \Omega_+^{-1}\psi^L. \quad (64)$$

The exact form of \mathcal{U} is only known for the *free-particle* Dirac Hamiltonian (a single freely moving particle, i.e., neglecting any scalar potentials ($V = 0$) and vector fields ($A = 0$), for instance, the nucleus–electron attraction) and called the *Foldy–Wouthuysen* transformation (Foldy and Wouthuysen 1950). The free-particle Foldy–Wouthuysen transformation (Foldy and Wouthuysen 1950) can be written as

$$\mathcal{U}_{\text{fpFW}} = \begin{pmatrix} A & \alpha AB \\ \alpha AB & -A \end{pmatrix}, \quad (65)$$

where

$$A = \sqrt{\frac{e_p + 1}{2e_p}}, \quad B = b(\boldsymbol{\sigma} \cdot \mathbf{p}), \quad b = \frac{1}{e_p + 1}, \quad e_p = \sqrt{1 + \alpha^2 p^2}, \quad \alpha = 1/c. \quad (66)$$

This transformation brings the free-particle Dirac Hamiltonian,

$$H_{\text{fpD}} = \begin{pmatrix} \mathbf{0}_2 & c(\boldsymbol{\sigma} \cdot \mathbf{p}) \\ c(\boldsymbol{\sigma} \cdot \mathbf{p}) & -2c^2 \mathbf{I}_2 \end{pmatrix}, \quad (67)$$

into block-diagonal form,

$$\mathcal{U}_{\text{fpFW}}^\dagger H_{\text{fpD}} \mathcal{U}_{\text{fpFW}} = \begin{pmatrix} T_p & \mathbf{0}_2 \\ \mathbf{0}_2 & -T_p - 2c^2 \mathbf{I}_2 \end{pmatrix}, \quad (68)$$

where

$$T_p = c^2(e_p - 1). \quad (69)$$

The free-particle Foldy–Wouthuysen transformation (Foldy and Wouthuysen 1950) thus involves square roots of the square root operator and the final block-diagonalized operator looks more complicated than the original Dirac Hamiltonian. Nonetheless, explicitly eliminating the small component results in a Hamiltonian for the positive-energy states (the electronic states) that may have a variationally lower bound, which removes the problem of variational collapse encountered in the

Dirac equation. However, the Foldy–Wouthuysen transformation discussed above is only exact for a constant potential. If we consider the Dirac Hamiltonian for an electron moving in the field of a nucleus like the hydrogen atom (i.e., with a nonconstant potential V), the decoupling operator is not exact, and the resulting transformed Hamiltonian still contains off-diagonal elements. Specifically, the free-particle Foldy–Wouthuysen transformation of the Dirac Hamiltonian, here indicated as H_1 , only reduces the mixing terms between the small and large component to the order of α^2 ,

$$H_1 = \begin{pmatrix} T_p + A(V + \alpha^2 BVB)A & \alpha A[V, B]A \\ \alpha A[B, V]A & -2\alpha^{-2} - T_p + A(V + \alpha^2 BVB)A \end{pmatrix}. \quad (70)$$

Similarly, the relation between the large and small component is modified. After the free-particle Foldy–Wouthuysen transformation, the small component is approximately proportional to the large component as

$$\psi_1^S \approx \alpha^3 \psi_1^L. \quad (71)$$

Because of the above relation, the transformed Hamiltonian H_1 (cf. Eq.(70)) represents a better starting point for subsequent two-component methods than the Dirac Hamiltonian itself.

One of the very first methods using the Foldy–Wouthuysen transformation as an initial transformation of the Dirac Hamiltonian has been introduced to chemistry by Bernard Hess under the name Douglas–Kroll transformation (Douglas and Kroll 1974). The Douglas–Kroll–Hess method (often denoted by DKH n , where n indicates the order of accuracy) is based on the transformation of the H_1 Hamiltonian, introduced by Douglas and Kroll, where the unitary transformation is given by (Douglas and Kroll 1974)

$$U_1 = \sqrt{1 + \mathscr{W}_1^2} + \mathscr{W}_1, \quad (72)$$

with \mathscr{W}_1 being some unknown anti-Hermitian operator that can be determined from Eq. (70) by eliminating off-diagonal terms to first order in the external potential V . The original implementation proposed by Hess (1985) was accurate up to the 2nd order in V . Further development of this method by Nakajima and coworkers (2000) in 2000 led to a relativistic Hamiltonian that is accurate up to the 3rd order in V . In 2002, Wolf and coworkers (2002) introduced a generalized form of the Douglas–Kroll–Hess transformation that facilitated the derivation of a Hamiltonian accurate up to the 5th order. Further improvements of this scheme by Reiher and Wolf (2004a, b) resulted in a transformed Hamiltonian that is accurate up to infinite order.

Below, we briefly describe the derivation and implementation of the DKH3 Hamiltonian. The derivation of higher-order terms can be done in a similar way, but is more tedious and time consuming. For transferability reasons to the original DKH work, we introduce some helpful notations

- \mathcal{E}_k : an even and Hermitian operator
- \mathcal{O}_k : an odd and Hermitian operator
- \mathcal{W}_k : an odd and anti-Hermitian operator

The subscript k indicates the order with respect to V . We should note that elements of the four-component Hamiltonian are written using calligraphic letters, while roman-type letters denote the corresponding two-component ones. Furthermore, a “'” will mark those terms that will be corrected in subsequent transformations. We will also apply this convention to the \mathcal{W}_k operator until its final representation has been found. Using this notation, the (transformed) Hamiltonian from Eq. (70) has the following form:

$$H_1 = \mathcal{E}_0 + \mathcal{E}_1 + \mathcal{O}'_1, \quad (73)$$

where

$$\mathcal{E}_0 = \begin{pmatrix} T_p & \mathbf{0}_2 \\ \mathbf{0}_2 & -2\alpha^{-2} - T_p \end{pmatrix}, \quad (74)$$

$$\mathcal{E}_1 = \begin{pmatrix} A(V + \alpha^2 BVB)A & \mathbf{0}_2 \\ \mathbf{0}_2 & A(V + \alpha^2 BVB)A \end{pmatrix}, \quad (75)$$

and

$$\mathcal{O}'_1 = \begin{pmatrix} \mathbf{0}_2 & \alpha A[V, B]A \\ \alpha A[B, V]A & \mathbf{0}_2 \end{pmatrix}. \quad (76)$$

The key idea of the DKH method is to apply a product of transformations as defined in Eq. (72) in order to gradually eliminate odd terms \mathcal{O}_k in the Hamiltonian. The DKH transformation can thus be understood as an order-by-order diagonalization. The transformed Hamiltonian of $(n + 1)$ th order has the general form

$$\begin{aligned} H_{n+1} &= U_n^\dagger U_{n-1}^\dagger \dots U_2^\dagger U_1^\dagger H_1 U_1 U_2 \dots U_{n-1} U_n \\ &= \sum_{i=0}^{2n+1} \mathcal{E}_i + \sum_{i=2n+2}^{\infty} \mathcal{E}'_i + \sum_{i=n+1}^{\infty} \mathcal{O}'_i. \end{aligned} \quad (77)$$

If we are interested in the DKH3 Hamiltonian, we have to include all \mathcal{E}_k operators for $k = 0, 1, 2, 3$. Thus, we have to transform the Hamiltonian H_1 only once to determine the DKH3 Hamiltonian (if $n = 1$, i runs over 0, 1, 2, 3 and we obtain a complete form of the operator \mathcal{E}_3). The most general (but not exclusive) ansatz to construct a unitary transformation is to use a power series expansion of an (anti-Hermitian) operator. In so doing, the U_1 transformation can be written as

$$U_1 = a_0 \mathbf{1}_4 + a_1 \mathcal{W}'_1 + a_2 \mathcal{W}'_1{}^2 + a_3 \mathcal{W}'_1{}^3 + \dots, \quad (78)$$

where the coefficients a_i correspond to the coefficients of a Taylor series expansion of Eq. (72). These coefficients are chosen in such a way that U_1 is as close to unitary as possible. This condition is particularly important when deriving higher-order DKH corrections and originates from the fact that the operators U_i can be represented in other ways than shown in Eq. (72). These different operator representations implicate some arbitrariness when choosing the expansion coefficients a_i .

The unitary condition for Eq. (78) reads

$$U_1^\dagger U_1 = \mathbf{1}_4 = a_0^2 \mathbf{1}_4 + (2a_0 a_2 - a_1^2) \mathscr{W}'_1{}^2 + (2a_0 a_4 - 2a_1 a_3 + a_2^2) \mathscr{W}'_1{}^4 + \dots \quad (79)$$

This condition is fulfilled if the parameter a_0 is equal to 1, while all remaining coefficients in front of the different powers of \mathscr{W}'_1 are equal to 0, that is,

$$a_0^2 = 1, \quad (80)$$

$$2a_0 a_2 - a_1^2 = 0, \quad (81)$$

$$2a_0 a_4 - 2a_1 a_3 + a_2^2 = 0, \quad (82)$$

etc. It is convenient to express the even coefficients (a_2, a_4 , etc.) using the odd coefficients (a_1, a_3 , etc.)

$$a_0^2 = 1, \quad (83)$$

$$a_2 = \frac{1}{2} \frac{a_1^2}{a_0}, \quad (84)$$

$$a_4 = a_0 a_1 a_3 - \frac{1}{8} a_0 a_1^4. \quad (85)$$

These relations will be particularly useful for simplifying higher-order corrections to DKH.

Our primary goal is to find the DKH3 Hamiltonian. To achieve this goal, we have to determine the operator \mathcal{E}_3 . This means that the transformation U_1 has to be accurate up to $\mathscr{W}'_1{}^2$. Therefore, the H_2 Hamiltonian can be obtained using the following approximate unitary transformation:

$$U_1 \approx a_0 \mathbf{1} + a_1 \mathscr{W}'_1 + a_2 \mathscr{W}'_1{}^2. \quad (86)$$

The transformed H_2 Hamiltonian, $H_2 = U_1^\dagger H_1 U_1$, can be expressed as

$$H_2 \approx \mathcal{E}_0 + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 + \mathcal{E}_4 + \mathcal{E}_5 \dots + \mathcal{O}_1 + \mathcal{O}_2 + \mathcal{O}_3, \quad (87)$$

where

$$\mathcal{E}_2 = -a_0 a_1 [\mathscr{W}'_1, \mathcal{O}'_1] + \frac{1}{2} a_1^2 [\mathscr{W}'_1, [\mathscr{W}'_1, \mathcal{E}_0]] \quad (88)$$

$$\mathcal{E}_3 = \frac{1}{2}a_1^2[\mathcal{W}'_1, [\mathcal{W}'_1, \mathcal{E}_1]] \quad (89)$$

$$\mathcal{O}'_1 = \mathcal{O}'_1 - a_0a_1[\mathcal{W}'_1, \mathcal{E}_0]. \quad (90)$$

Note that the brackets $[\]$ are used to denote commutators, that is, $[A, B] = AB - BA$, while the brackets $\{ \}$ indicate anticommutators, that is, $\{A, B\} = AB + BA$. The remaining (prime) terms are not complete, that is, block diagonal, and will be modified in a subsequent DKH transformation step. Now, we have to eliminate the first odd term \mathcal{O}'_1 , that is, we have to ensure that $\mathcal{O}'_1 = 0$. This elimination procedure defines the \mathcal{W}'_1 operator,

$$[\mathcal{W}'_1, \mathcal{E}_0] = \frac{a_0}{a_1}\mathcal{O}'_1. \quad (91)$$

Substituting the above equation into Eq. (88) further simplifies the expression for \mathcal{E}_2 ,

$$\mathcal{E}_2 = -\frac{1}{2}a_0a_1[\mathcal{W}'_1, \mathcal{O}'_1]. \quad (92)$$

The dependence on the coefficients a_0 and a_1 in the above equation can be eliminated by substituting

$$\mathcal{W}_1 = a_0a_1\mathcal{W}'_1, \quad (93)$$

which according to Eq. (91) leads to

$$[\mathcal{W}_1, \mathcal{E}_0] = \mathcal{O}'_1. \quad (94)$$

Using this definition of the \mathcal{W}_1 operator, we can write for \mathcal{E}_2

$$\mathcal{E}_2 = -\frac{1}{2}[\mathcal{W}_1, \mathcal{O}'_1], \quad (95)$$

and for \mathcal{E}_3

$$\mathcal{E}_3 = \frac{1}{2}[\mathcal{W}_1, [\mathcal{W}_1, \mathcal{E}_1]]. \quad (96)$$

Consistent with our notation, the abovementioned operators \mathcal{E}_2 and \mathcal{E}_3 are 4×4 matrices. We can now derive corrections for the two-component operators acting on the electronic states, which will be denoted (now using roman-type letters) as E_i . We have

$$E_0 = T_p, \quad (97)$$

and

$$E_1 = A(V + \alpha^2 BVB)A. \quad (98)$$

Furthermore, the operators \mathcal{O}'_1 and \mathcal{W}_1 can be expressed using two-component operators

$$\mathcal{O}'_1 = \begin{pmatrix} \mathbf{0}_2 & O_1 \\ -O_1 & \mathbf{0}_2 \end{pmatrix}, \quad \mathcal{W}_1 = \begin{pmatrix} \mathbf{0}_2 & W_1 \\ W_1 & \mathbf{0}_2 \end{pmatrix}, \quad (99)$$

where

$$O_1 = \alpha A[V, B]A, \quad (100)$$

and the operator W_1 is given by

$$\{W_1, e_p\} = -\alpha^2 O_1. \quad (101)$$

Finally, we can bring the operators \mathcal{E}_i to a block diagonal form (which will give us the DKH3 Hamiltonian), where the operators acting on the electronic states (upper block) are separated from the operators acting on the positronic states:

$$\mathcal{E}_0 = \begin{pmatrix} E_0 & \mathbf{0}_2 \\ \mathbf{0}_2 & -E_0 - 2c^2 \mathbf{I}_2 \end{pmatrix}, \quad \mathcal{E}_1 = \begin{pmatrix} E_1 & \mathbf{0}_2 \\ \mathbf{0}_2 & E_1 \end{pmatrix}, \quad (102)$$

$$\mathcal{E}_2 = \begin{pmatrix} E_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & -E_2 \end{pmatrix}, \quad \mathcal{E}_3 = \begin{pmatrix} E_3 & \mathbf{0}_2 \\ \mathbf{0}_2 & E_3 \end{pmatrix}, \quad (103)$$

with

$$E_2 = \frac{1}{2}\{W_1, O_1\} \quad (104)$$

and

$$E_3 = \frac{1}{2}[W_1, [W_1, E_1]]. \quad (105)$$

Higher orders of the DKH Hamiltonian can be constructed according to a snap-together modular building block system. Such an approach was developed by Reiher and Wolf and led to the development of the infinite-order DKH method (Reiher and Wolf 2004a, b). However, higher-order DKH Hamiltonians have a more complex structure due to additional parameters that have to be determined to ensure that the transformation matrices are unitary. In practical applications, corrections higher than the 14th order are not used because the time required to calculate these higher-order corrections (for the one-electron part of the Hamiltonian) becomes larger than

an actual four-component calculation, while the differences from the reference Dirac energies are negligible. Besides, no matter which DKH order we are considering, the problem will become algebraic at some point. Both the Foldy–Wouthuysen and DKH Hamiltonians are well defined in momentum space. The construction of the matrix representation of such Hamiltonians thus requires a Fourier transformation from \mathbf{r} to \mathbf{p} space. Most probably, one reason for the popularity of the DKH method is the scheme proposed by Hess (1985) that allows us to obtain matrix representations of the operator products in different spaces. We will elaborate on this issue by means of the matrix representation of the E_2 operator.

Let $|\chi\rangle$ denote some orthogonal basis with functions depending on \mathbf{r} . After introducing such a basis set, we can diagonalize the matrix representation of the operator p^2 ,

$$\mathbf{G}^\dagger \langle \chi | p^2 | \chi \rangle \mathbf{G} = \mathbf{p}^2, \quad (106)$$

where \mathbf{G} is an orthogonal and \mathbf{p}^2 a diagonal matrix, respectively. After diagonalization, we can introduce a new basis

$$|\mathbf{k}\rangle = |\chi\rangle \mathbf{G}, \quad (107)$$

that is an eigenvector of the operator p^2 represented in terms of functions dependent on \mathbf{r} . The important point now is that, within this new basis $|\mathbf{k}\rangle$ (which is \mathbf{r} dependent), we can represent operators that are functions of p^2 ,

$$\langle \mathbf{k} | f(p^2) | \mathbf{k} \rangle = f. \quad (108)$$

Furthermore, using the approximate resolution of identity,

$$|\mathbf{k}\rangle \langle \mathbf{k}| \approx 1 \quad (109)$$

we can derive the representations of the operator products defined in \mathbf{r} and \mathbf{p} space. Specifically, in the DKH method, we have to define the general matrix representation of two kinds of operators, namely:

$$X = f(p^2)g(r)h(p^2), \quad (110)$$

and

$$Y = (\boldsymbol{\sigma} \cdot \mathbf{p}) f(p^2)g(r)h(p^2)i(r)j(p^2)(\boldsymbol{\sigma} \cdot \mathbf{p}). \quad (111)$$

In the first case, we have to substitute Eq. (109) between the operators dependent on p^2 and r :

$$\langle \mathbf{k} | X | \mathbf{k} \rangle = \langle \mathbf{k} | f(p^2) | \mathbf{k} \rangle \langle \mathbf{k} | g(r) | \mathbf{k} \rangle \langle \mathbf{k} | h(p^2) | \mathbf{k} \rangle \quad (112)$$

$$= f \mathbf{G}^\dagger \langle \chi | g(r) | \chi \rangle \mathbf{G} h, \quad (113)$$

where in going from the first to the second equality, we have used Eqs.(108) and (106). We cannot, however, use the same procedure for the second case as the p^2 -dependent part will be linear in the operator product $(\boldsymbol{\sigma} \cdot \mathbf{p})$, that is, we will have odd terms. Instead, operator Y can be rearranged using

$$\frac{(\boldsymbol{\sigma} \cdot \mathbf{p})(\boldsymbol{\sigma} \cdot \mathbf{p})}{p^2} = 1, \quad (114)$$

to the following form

$$Y = f(p^2)(\boldsymbol{\sigma} \cdot \mathbf{p})g(r)(\boldsymbol{\sigma} \cdot \mathbf{p})h(p^2)\frac{1}{p^2}(\boldsymbol{\sigma} \cdot \mathbf{p})i(r)(\boldsymbol{\sigma} \cdot \mathbf{p})j(p^2), \quad (115)$$

which results in the following matrix representation

$$\langle \mathbf{k} | Y | \mathbf{k} \rangle = \mathbf{f} \mathbf{G}^\dagger \langle \boldsymbol{\chi} | (\boldsymbol{\sigma} \cdot \mathbf{p})g(r)(\boldsymbol{\sigma} \cdot \mathbf{p}) | \boldsymbol{\chi} \rangle \mathbf{G} \mathbf{h} \frac{1}{p^2} \mathbf{G}^\dagger \langle \boldsymbol{\chi} | (\boldsymbol{\sigma} \cdot \mathbf{p})i(r)(\boldsymbol{\sigma} \cdot \mathbf{p}) | \boldsymbol{\chi} \rangle \mathbf{G} \mathbf{j}. \quad (116)$$

Integrals of the type $\langle \boldsymbol{\chi} | (\boldsymbol{\sigma} \cdot \mathbf{p})g(r)(\boldsymbol{\sigma} \cdot \mathbf{p}) | \boldsymbol{\chi} \rangle$, which appear in the above equations, are computed using the Hermitian property of the operator $(\boldsymbol{\sigma} \cdot \mathbf{p})$,

$$\langle \boldsymbol{\chi} | (\boldsymbol{\sigma} \cdot \mathbf{p})g(r)(\boldsymbol{\sigma} \cdot \mathbf{p}) | \boldsymbol{\chi} \rangle = \langle (\boldsymbol{\sigma} \cdot \mathbf{p}) \boldsymbol{\chi} | g(r) | (\boldsymbol{\sigma} \cdot \mathbf{p}) \boldsymbol{\chi} \rangle. \quad (117)$$

Now, we are all set to derive a matrix representation of the operator E_2 in the basis $\boldsymbol{\chi}$. Recall that the E_2 operator has the following form

$$E_2 = \frac{1}{2}(W_1 O_1 + O_1 W_1). \quad (118)$$

From the property of the operators \mathscr{W}_1 (anti-Hermitian) and \mathscr{O}_1 (Hermitian), we can deduce that (cf. Eq. (99))

$$W_1^\dagger = -W_1 \quad \text{and} \quad O_1^\dagger = O_1. \quad (119)$$

Since $(W_1 O_1)^\dagger = O_1 W_1$, it is, thus, sufficient to derive only one of the two possible products of Eq.(118). First, let us introduce a temporary operator Y defined as follows:

$$Y = W_1 O_1, \quad (120)$$

for which Eq. (118) becomes

$$E_2 = \frac{1}{2}(Y + Y^\dagger). \quad (121)$$

As the W_1 and O_1 operators are odd, their matrix representations should be obtained with two basis sets of different parity. However, in two-component methods, we are interested in using only the large component. A remedy to this problem is obvious: we can insert Eq. (114) between the W_1 and O_1 operators and bring the operator Y to a new form,

$$Y = W_1(\boldsymbol{\sigma} \cdot \mathbf{p}) \frac{1}{p^2} (\boldsymbol{\sigma} \cdot \mathbf{p}) O_1. \quad (122)$$

As a result, we have to determine matrices of two even operators,

$$W = W_1(\boldsymbol{\sigma} \cdot \mathbf{p}) \quad \text{and} \quad O = (\boldsymbol{\sigma} \cdot \mathbf{p}) O_1. \quad (123)$$

The W operator, which defines the \mathcal{U}_1 transformation, can be determined from the following condition (*cf.* Eq. (101)):

$$W e_p + e_p W = -\alpha^2 O. \quad (124)$$

In practice, we have, thus, to derive the matrix representation of the operator O only,

$$O = \alpha A (V b p^2 - b(\boldsymbol{\sigma} \cdot \mathbf{p}) V (\boldsymbol{\sigma} \cdot \mathbf{p})) A. \quad (125)$$

Given the matrix representation of an arbitrary operator Γ in the basis \mathbf{k}

$$\Gamma = \langle \mathbf{k} | \Gamma | \mathbf{k} \rangle, \quad (126)$$

we can write for the matrix representation of O

$$O = \alpha A \left(G \langle \chi | V | \chi \rangle G^\dagger b p^2 - b G \langle (\boldsymbol{\sigma} \cdot \mathbf{p}) \chi | V | (\boldsymbol{\sigma} \cdot \mathbf{p}) \chi \rangle \right) A. \quad (127)$$

Now, Eq. (124) can be rewritten in the same basis

$$W e_p + e_p W = -\alpha^2 O. \quad (128)$$

Finding a solution to the above equation is straightforward as the matrix e_p is diagonal. Using matrix representation, we obtain

$$W_{(i,j)} e_{p(j,j)} + e_{p(i,i)} W_{(i,j)} = -(\alpha^2 O^\dagger)_{(i,j)}, \quad (129)$$

with matrix elements of W defined as

$$W_{(i,j)} = -\alpha^2 \frac{O_{(i,j)}}{e_{p(i,i)} + e_{p(j,j)}} \quad (130)$$

Finally, after we have determined the matrix representation of the operators W and O , we can derive the matrix representation of the Y operator,

$$Y = W \frac{1}{p^2} O, \quad (131)$$

which can be used to calculate the matrix representation of the operator E_2 in the basis $|\mathbf{k}\rangle$,

$$E_2 = \frac{1}{2}(Y + Y^\dagger). \quad (132)$$

Since we intended to obtain a matrix representation of E_2 in basis χ , we have to reverse the transformation defined in Eq. (107). This results in

$$\langle \chi | E_2 | \chi \rangle = \mathbf{G} E_2 \mathbf{G}^\dagger. \quad (133)$$

Matrix representations of other corrections and W_i operators can be determined in a similar way.

We should emphasize that the second-order DKH (DKH2) Hamiltonian is used in the majority of applications of the DKH method. Higher-order DKH Hamiltonians (DKH n with $n = 3, 4, \dots$) are supported in some quantum chemistry computer programs but are less frequently used (Nakajima and Hirao 2000; van Wüllen 2004; Reiher and Wolf 2004a, b). The reason for the prevalent use of the DKH2 Hamiltonian is that the accuracy of the second-order approximation is usually sufficient for chemical applications, resulting in negligible deviations from a four-component reference calculation (see section “Two-Component Hamiltonians in Practice” for a detailed discussion on the performance of two-component Hamiltonians). Furthermore, a higher-order transformation only reduces the error in the external potential V , but does not affect the error in the electron–electron interaction. The error in the electron–electron interaction is caused by the fact that in a conventional DKH transformation, only the one-electron part of the Hamiltonian has been replaced by the DKH transformed expressions, while the (nonrelativistic) two-electron part remains unchanged. However, transforming the nonrelativistic Coulomb interaction, or the even more complicated Breit interaction (Eq. (35)), results in a considerably complicated expression for the two-electron part of the Hamiltonian. Since this term is difficult to evaluate efficiently, the two-electron interaction is not altered in conventional applications of the DKH method. The different treatment of one- and two-electron terms is, however, inconsistent and gives rise to the so-called *picture change* (Dyall and Faegri 2007; Kello and Sadlej 1997).

A picture change error is always encountered when only some operators are transformed to the DKH picture, while others are left untouched. For instance, if the Hamiltonian is augmented by other operators (dipole moment operator, etc.), these additional operators need to be transformed first before molecular properties can be evaluated. Skipping this transformation and using the original form of the operator to evaluate its expectation value lead to non-negligible errors, the so-called

picture change errors (Kello and Sadlej 1997). Therefore, to achieve results that agree well with reference (four-component) Dirac calculations, the corresponding operators have to be transformed to the DKH picture to eliminate (or minimize) the picture change error.

A One-Step DKH Method

One may think that DKH n is the method we have been looking for: it describes only the electronic part of the spectrum and the electronic wavefunction can be optimized as in nonrelativistic calculations, that is, we don't have to modify our quantum chemistry code substantially. However, higher orders of the DKH n Hamiltonian are computationally demanding, primarily because we have to perform a series of product transformations \mathcal{U}_i that have to be expressed in higher powers of \mathcal{H}_i (longer expansions). Even though we can utilize algorithms to obtain higher-order corrections to DKH, the method is nonetheless computationally expensive. A solution to this problem is to replace the product of transformations by a one-step procedure. A particularly useful scheme was proposed by Heully and coworkers (1986) that uses a transformation \mathcal{U} as defined in Eq. (58) in terms of an anti-Hermitian, unitary operator R . The arbitrariness of the operator R , however, does not deteriorate the ansatz for \mathcal{U} but allows us to construct different two-component Hamiltonians starting with the same general expression. Yet, not every form of the operator R leads to a block-diagonal form of the Hamiltonian.

If \mathcal{U} is chosen as in Eq. (58), the transformed electronic Hamiltonian is given by

$$h_+ = \Omega_+ \left((H_1)_{11} + R^\dagger (H_1)_{21} + (H_1)_{12} R + R^\dagger (H_1)_{22} R \right) \Omega_+, \quad (134)$$

which can be rewritten using the matrix elements of H_1 (Eq. (70)) and the notation defined in Eqs. (97), (98), and (100)

$$h_+ = \Omega_+ \left(E_0 + E_1 + [R, O_1] - R(-E_0 + E_1 - 2\alpha^{-2})R \right) \Omega_+. \quad (135)$$

We can completely eliminate the off-diagonal blocks of the Hamiltonian H_1 if

$$O_1 + \alpha^{-2}(e_p R + R e_p) + R E_1 - E_1 R + R O_1 R = 0 \quad (136)$$

is satisfied. We now seek a method analogous to DKH n but with a unitary transformation as given by Eq. (58). First, we expand the operator R into a series (Kędziera 2005),

$$R = \sum_{i=1}^{\infty} R_i, \quad (137)$$

where the index i denotes the contribution of the i th order in the external potential V . The first and main advantage of this approach over the standard DKH

transformation is that we don't need to parameterize any coefficients in front of the R_i operators. Substituting Eq. (137) into Eq. (136) gives

$$O_1 + \alpha^{-2} \sum_{i=1}^{\infty} (e_p R_i + R_i e_p) + \sum_{i=1}^{\infty} (R_i E_1 - E_1 R_i) = 0 + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} R_i O_1 R_j. \quad (138)$$

After rearranging with respect to orders in V , we get

$$\begin{aligned} 0 = & O_1 + \alpha^{-2}(e_p R_1 + R_1 e_p) \\ & + \alpha^{-2}(e_p R_2 + R_2 e_p) + (R_1 E_1 - E_1 R_1) \\ & + \alpha^{-2}(e_p R_3 + R_3 e_p) + (R_2 E_1 - E_1 R_2) + R_1 O_1 R_1 \\ & + \alpha^{-2}(e_p R_4 + R_4 e_p) + (R_3 E_1 - E_1 R_3) + R_1 O_1 R_2 + R_2 O_1 R_1 \\ & + \alpha^{-2}(e_p R_5 + R_5 e_p) + (R_4 E_1 - E_1 R_4) + R_1 O_1 R_3 + R_2 O_1 R_2 \\ & + R_3 O_1 R_1 \\ & + \dots \end{aligned} \quad (139)$$

We can write the above equation in closed form

$$e_p R_1 + R_1 e_p = -\alpha^2 O_1 \quad (140)$$

for $i = 1$ and

$$e_p R_i + R_i e_p = -\alpha^2 \left(R_{i-1} E_1 - E_1 R_{i-1} + \sum_{k=1}^{i-2} R_k O_1 R_{i-1-k} \right) \quad (141)$$

for $i \geq 2$. The above equations can be solved up to arbitrary order in V , which guarantees accuracy of the electronic Hamiltonian. As in the case of the W_1 operator for DKH, the R operator must be multiplied by $\sigma \cdot p$ to make it even. For that purpose, we now introduce temporary operators Y_i ,

$$Y_i = (\sigma \cdot p) R_i. \quad (142)$$

We can determine the operators Y_i from

$$e_p Y_1 + Y_1 e_p = -\alpha^2 O \quad (143)$$

for $i = 1$ and

$$e_p Y_i + Y_i e_p = -\alpha^2 \left(Y_{i-1} E_1 - (\sigma \cdot p) E_1 (\sigma \cdot p) Y_{i-1} + \sum_{k=1}^{i-2} Y_k O p^{-2} Y_{i-1-k} \right) \quad (144)$$

for $i \geq 2$. These equations can be rewritten in a more elegant way by introducing some auxiliary operators Z_i , with

$$Z_1 = O \quad (145)$$

and

$$Z_i = Y_{i-1}E_1 - (\boldsymbol{\sigma} \cdot \mathbf{p})E_1(\boldsymbol{\sigma} \cdot \mathbf{p})Y_{i-1} + \sum_{k=1}^{i-2} Y_k O p^{-2} Y_{i-1-k} \quad (146)$$

for $i \geq 2$. Now, we can obtain the expressions for Y_i in closed form which is similar to Eq. (124),

$$e_p Y_i + Y_i e_p = -\alpha^2 Z_i \quad (147)$$

with the following matrix expansion

$$Y_{i,(k,l)} = -\alpha^2 \frac{Z_{i,(k,l)}}{e_{p(k,k)} + e_{p(l,l)}}. \quad (148)$$

If we are interested in numerical accuracy rather than its analytical form, we can use an iterative approach to Eq. (136) suggested by Sadlej and coworkers (2002). This method is known as the infinite-order two-component (IOTC) approach. In the IOTC method, Eq. (136) is rewritten in terms of the operator Y ,

$$O + \alpha^{-2}(e_p Y + Y e_p) + Y E_1 + (\boldsymbol{\sigma} \cdot \mathbf{p}) E_1 (\boldsymbol{\sigma} \cdot \mathbf{p}) p^{-2} Y + Y O p^{-2} Y = 0. \quad (149)$$

After some rearrangements, the above equation can be directly used in an iterative scheme

$$e_p Y_{k+1} + Y_{k+1} e_p = -\alpha^2 (O + Y E_1 + (\boldsymbol{\sigma} \cdot \mathbf{p}) E_1 (\boldsymbol{\sigma} \cdot \mathbf{p}) p^{-2} Y + Y O p^{-2} Y), \quad (150)$$

where we have the initial condition

$$e_p Y_0 + Y_0 e_p = -\alpha^2 O \quad (151)$$

to start the iteration.

In practice, finding the solutions to these equations is equivalent to finding the matrix representation of operator Y and, as a consequence, the electronic Hamiltonian. Such a procedure is rather ordinary in quantum chemistry as we often elaborate on mathematical formulas in favor of their algebraization. This is exactly what we do when we compute expectation values (like energies) in any quantum chemistry software package.

To conclude, the Foldy–Wouthuysen transformation yields an initial Hamiltonian that can be used to derive various electronic two-component Hamiltonians.

Specifically in this subsection, we have discussed only a couple of approaches that are based on an expansion with respect to the external potential V . There also exist two-component methods that are based on an expansion with respect to the parameter α , like the Barysz–Sadlej–Snijders (BSS) approach (Barysz et al. 1997). This family of methods is, however, not as effective as methods based on an expansion with respect to V due to slow convergence or presence of singularities, unless the complete expansion is used (Reiher and Wolf 2014). Nonetheless, the Foldy–Wouthuysen transformation is a key step in improving the accuracy of two-component methods. We should note that, in quantum chemistry calculations, we have to ensure the elimination of odd operators, linear in $(\boldsymbol{\sigma} \cdot \boldsymbol{p})$. So far, this has been realized at the two-component level. It is, however, also possible to eliminate odd operators at the Hamiltonian level, that is, using transformations that are similar to those used in NESCF.

Algebraic Alternatives: The X2C Family of Methods

If we are not interested in analytic formulas of operators, but rather in their matrix representation, we have to solve Eq. (61) in its corresponding matrix form:

$$\mathbf{R}(\mathbf{H})_{11} + \mathbf{R}(\mathbf{H})_{12}\mathbf{R} - (\mathbf{H})_{21} - (\mathbf{H})_{22}\mathbf{R} = 0, \quad (152)$$

where \mathbf{H}_{ij} is a block representation of the Hamiltonian matrix H in a given basis set. In mathematics, there is an identical quadratic equation known as the nonsymmetric algebraic Riccati equation (nARE) (Kędziera 2006; Kędziera and Barysz 2007)

$$-\mathbf{X}\mathbf{C}\mathbf{X} + \mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{D} + \mathbf{B} = \mathbf{0}, \quad (153)$$

where \mathbf{X} is an unknown matrix and the coefficients \mathbf{A} , \mathbf{B} , \mathbf{C} , \mathbf{D} are matrices (the individual sizes of all elements must lead to an $n \times n$ matrix for $\mathbf{0}$). The solutions of Eq. (153) are connected to the eigenvectors of a matrix \mathbf{M} defined by

$$\mathbf{M} = \begin{bmatrix} \mathbf{D} & -\mathbf{C} \\ -\mathbf{B} & \mathbf{A} \end{bmatrix}. \quad (154)$$

When comparing Eqs. (152) and (153), we observe that the solutions of Eq. (152) depend on the eigenvectors of the Hamiltonian matrix \mathbf{H} , that is, the matrix representation of the Hamiltonian. Before we write down the form of the solutions of \mathbf{R} , let us examine the eigenproblem of \mathbf{H} ,

$$\mathbf{H}\mathbf{V} = \mathbf{V}\mathbf{E}, \quad (155)$$

where \mathbf{V} is the matrix of eigenvectors and \mathbf{E} is a diagonal matrix containing the eigenvalues. In our case, we can organize the eigenvalues in such a way that we

split them into two subsets, \mathbf{E}_+ for electronic and \mathbf{E}_- for positronic energies, respectively. According to this division, we can organize the eigenvectors as follows:

$$\mathbf{V} = [\mathbf{V}_+, \mathbf{V}_-] = \left[\begin{pmatrix} \mathbf{V}_{11} \\ \mathbf{V}_{21} \end{pmatrix}, \begin{pmatrix} \mathbf{V}_{12} \\ \mathbf{V}_{22} \end{pmatrix} \right]. \quad (156)$$

Now, \mathbf{V}_+ corresponds to \mathbf{E}_+

$$\mathbf{H}\mathbf{V}_+ = \mathbf{V}_+\mathbf{E}_+, \quad (157)$$

while \mathbf{V}_- corresponds to \mathbf{E}_-

$$\mathbf{H}\mathbf{V}_- = \mathbf{V}_-\mathbf{E}_-. \quad (158)$$

Using the properties of \mathbf{V}_+ , we obtain

$$\mathbf{H}_{11}\mathbf{V}_{11} + \mathbf{H}_{12}\mathbf{V}_{21} = \mathbf{V}_{11}\mathbf{E}_+, \quad (159)$$

and

$$\mathbf{H}_{21}\mathbf{V}_{11} + \mathbf{H}_{22}\mathbf{V}_{21} = \mathbf{V}_{21}\mathbf{E}_+. \quad (160)$$

Since the \mathbf{V}_{ij} blocks are assumed to be invertible, we can now determine \mathbf{E}_+ from Eq. (159),

$$\mathbf{E}_+ = \mathbf{V}_{11}^{-1}\mathbf{H}_{11}\mathbf{V}_{11} + \mathbf{V}_{11}^{-1}\mathbf{H}_{12}\mathbf{V}_{21}. \quad (161)$$

Inserting this expression into Eq. (160), we obtain

$$\mathbf{H}_{21} + \mathbf{H}_{22}\mathbf{V}_{21}\mathbf{V}_{11}^{-1} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{H}_{11} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{H}_{12}\mathbf{V}_{21}\mathbf{V}_{11}^{-1} = \mathbf{0}. \quad (162)$$

Introducing

$$\mathbf{Y} = \mathbf{V}_{21}\mathbf{V}_{11}^{-1}; \quad (163)$$

we can rewrite the above equation as follows:

$$\mathbf{Y}\mathbf{H}_{11} + \mathbf{Y}\mathbf{H}_{12}\mathbf{Y} - \mathbf{H}_{21} - \mathbf{H}_{22}\mathbf{Y} = \mathbf{0}, \quad (164)$$

that is, we come back to our starting point Eq. (152). Finally, we have a form for the solutions that lead to block diagonalization of \mathbf{H} . By multiplying Eq. (157) from the left side by \mathbf{V}_+ and using the orthogonality properties of \mathbf{V}_+ , we obtain the matrix representation of the electronic two-component Hamiltonian h_+ ,

$$\mathbf{h}_+ = \mathbf{\Omega}_+ \left(\mathbf{H}_{11} + \mathbf{H}_{12}\mathbf{Y}_1 + \mathbf{Y}_1^\dagger\mathbf{H}_{21} + \mathbf{Y}_1^\dagger\mathbf{H}_{22}\mathbf{Y}_1 \right) \mathbf{\Omega}_+, \quad (165)$$

where

$$\mathbf{\Omega}_+ = (\mathbf{1} + \mathbf{Y}_1^\dagger \mathbf{Y}_1)^{-1/2}, \quad (166)$$

which concludes all our efforts.

Now, one may wonder what the benefits are which we get from the solutions of nARE in relativistic quantum chemistry. We were looking for two-component Hamiltonians, but we finished with diagonalizing its four-component counterpart. The reasons for that are manifold:

1. There exists an exact solution of \mathbf{H} , and its form can be obtained within the nARE approach or other methods, like IOTC, or a series expansion of R .
2. Among all the existing solutions of Eq. (152), we have the freedom to choose those solutions that are important for chemistry. Specifically, we are interested in solutions that lead to a two-component Hamiltonian whose spectrum matches the electronic part of the Dirac spectrum.
3. \mathbf{H} can be any arbitrary Hamiltonian, for example, the Dirac Hamiltonian defined in the basis of the large and small components or the modified Dirac Hamiltonian in the basis of the large component only.

To summarize, the nARE approach was originally proposed for the IOTC Hamiltonian (Kędziera 2006; Kędziera and Barysz 2007). Iliáš and Saue (2007) used this approach for the Dirac Hamiltonian for the first time (in the basis of large and small component functions). Their approach is known as the eXact 2-Component (X2C) method (Iliáš and Saue 2007; Sikkema et al. 2009; Liu and Peng 2009). We should also mention that the matrix representation of the Hamiltonian allows us to efficiently correct for picture change errors, not only for one-electron terms but also for two-electron operators (Iliáš and Saue 2007; Sikkema et al. 2009). These advantages promoted the X2C method to a very popular approach among computational chemists and led to its further development. Recently, an efficient approximation to the decoupling matrix has been introduced in the X2C scheme that uses atomic basis sets rather than the complete molecular basis. This feature facilitates accurate studies of large molecules containing heavy elements (Peng and Reiher 2012; Peng et al. 2013). Furthermore, the X2C method has been used to accurately model magnetic properties in heavy-element chemistry (Autschbach et al. 2012).

Concluding Remarks on Two-Component Hamiltonians

Over the past few decades, a number of two-component Hamiltonians has been developed. The first big step toward an accurate description of the electronic part of the Dirac equation emerged with the DKH Hamiltonian, which allowed for reliably studying many-electron systems containing heavy elements for the first time. In the course of time, the ZORA Hamiltonian gained a lot of popularity, primarily because

of its simplicity. These traditional approaches are slowly being replaced by more robust methods, such as IOTC, nARE, and X2C.

While there is still some development going on in the field of relativistic two-component methods, it is, however, mostly focused on approximations and efficient implementations of the already-developed two-component Hamiltonians in quantum chemistry codes, rather than on the development of new approaches (Sikkema et al. 2009; Autschbach et al. 2012; Peng and Reiher 2012; Peng et al. 2013). Modern two-component methods reached the accuracy of the Dirac Hamiltonian leaving no need for further methodological improvements.

Two-Component Hamiltonians in Practice

We will redirect our so far theory-focused discussion about two-component Hamiltonians to some more practical aspects. The following sections summarize useful instructions on quantum chemistry software packages supporting two-component Hamiltonians, in which atomic basis sets should be chosen in molecular calculations and in which two-component Hamiltonians yield the most accurate results for the properties of heavy-element compounds.

Where to Find Two-Component Hamiltonians

Table 1 lists the most popular quantum chemistry codes that support quantum chemical calculations with relativistic two-component Hamiltonians. As indicated in the Table, not every two-component Hamiltonian is easily available in standard quantum chemistry software packages. Most programs support only one or two types of relativistic two-component Hamiltonians. Clearly, the DKH Hamiltonian is the most accessible one as we can find it in almost every code. We should emphasize that in most cases the DKH2 or DKH3 Hamiltonians are available. Only MOLPRO is an exception and supports higher-order corrections of the DKH Hamiltonian (Reiher and Wolf 2004a, b). The popularity of DKH is followed by the ZORA and X2C methods.

While the so-called scalar relativistic effects (one-component form) are usually accounted for on the same footing in all quantum chemistry software packages, the implementation of spin-orbit coupling varies from one software to another. In general, two strategies are used:

- treat spin-orbit coupling in a perturbative manner on top of a scalar relativistic calculation (e.g., MOLPRO)
- perform a two-component calculation with an explicit inclusion of spin-orbit coupling (e.g., ADF)

Yet, in the majority of applications, spin-orbit coupling is treated within an atomic mean field (AMF) (Hess et al. 1996; Schimmelpfennig et al. 1998a, b)

Table 1 Most popular two-component Hamiltonians available in quantum chemistry codes

Software package	Two-component Hamiltonian				
	ZORA	NESC	DKH	IOTC	X2C
ADF	✗				
CFour					✗
COLUMBUS			✗		
DALTON			✗		
DIRAC	✗		✗	✗	✗
GAMESS-US		✗	✗	✗	
GAUSSIAN			✗		
MOLCAS			✗		
MOLPRO			✗		
NWCHEM	✗		✗		
ORCA	✗		✗		
Q-CHEM					✗
RESPECT			✗		
TURBOMOLE			✗		✗

approximation. In AMF, the two-electron integrals are computed in an atomic mean field fashion. This procedure has considerable computational advantages, while the loss in accuracy is negligibly small (Tecmer et al. 2014).

Finally, we should note that the DIRAC and RESPECT software packages are fully four-component codes, and therefore their code structure is rather different from all the other quantum chemistry programs listed in Table 1.

Which Basis Sets to Use for Relativistic Calculations

Every relativistic calculation should be performed in a combination with a relativistic basis set. As mentioned at the beginning of this chapter, relativistic effects cause the shrinkage of the core s and p and the expansion of the valence d and f atomic orbitals. This dictates modification of the contraction coefficients compared to standard (nonrelativistic) basis sets (Hill 2013; Peterson and Dyal 2015). These differences are even larger the heavier the nuclei. Moreover, one should be aware of the fact that performing calculations with relativistic two-component methods, the basis set should, in principle, be optimized for a given two-component Hamiltonian. This general guideline is best understood using the following example:

Example 3. Let us assume that we have a basis set optimized for the DKH2 Hamiltonian. What will happen when we use this basis set in calculations with the IOTC Hamiltonian? To exemplify the interdependence of basis sets and Hamiltonians, we calculate total Hartree–Fock (HF) energies for the closed-shell atoms Zn ($Z = 30$), Cd ($Z = 48$), and Hg ($Z = 80$) using both the DKH2 and IOTC Hamiltonian. Furthermore, we optimize a global scaling parameter k , which scales all basis set

Table 2 HF energies for optimized values of the scaling parameter k using the DKH2 and IOTC Hamiltonians. $\Delta_{\text{IOTC-DKH2}}$ denotes the difference between the HF energies calculated with the IOTC Hamiltonian with k optimized for DKH2 and for IOTC, respectively

Atom	DKH2		IOTC		$\Delta_{\text{IOTC-DKH2}}$
	E_h	k	E_h	k	E_h (eV)
Zn	-1794.150 734	1.213	-1794.224 690	1.255	0.000 035 (0.001)
Cd	-5590.160 999	1.058	-5591.273 439	1.080	0.000 323 (0.009)
Hg	-19602.652 706	1.255	-19622.039 138	1.319	0.012 268 (0.334)

coefficients in such a way that the total energy is minimized. Note that, in general, all basis set coefficients are optimized when relativistic basis sets are “designed.” This is, however, not a trivial task to do and thus we will restrict our example to only one optimization parameter, which will be sufficient for a proof of principle analysis. The results of our example calculations are presented in Table 2. Most importantly, the basis set optimized for DKH2 is no longer suitable for IOTC when moving toward heavier nuclei. This basis set dependence constitutes an error of approximately 0.33 eV for Hg, which is unacceptable.

There is a large variety of relativistic basis sets available in the literature. We will briefly discuss those that are used in correlated and DFT calculations, i.e., the most important methods to accurately model electronic structures of heavy elements. One of the most popular relativistic basis set libraries is the ANO-RCC (atomic natural orbital-relativistic correlation consistent) basis set family, developed by Roos and coworkers (2004, 2005a, 2005b, 2008) and specifically optimized for the DKH2 Hamiltonian. This generally contracted Gaussian basis sets are available for the whole periodic table (Hill 2013). Another group of Gaussian basis sets optimized for the DKH Hamiltonian of different orders have been developed by Peterson and coworkers (Balabanov and Peterson 2005; Prascher et al. 2011; Bross and Peterson 2014; Peterson 2015). Their work represents an extension to the well-known correlation-consistent basis set collection. In analogy to its nonrelativistic counterpart, they are denoted as cc-pVnZ-DK, where n denotes the cardinal number. Similar basis sets were also developed by the Sapporo group (Sekiya et al. 2010; Noro et al. 2009, 2012, 2013). Pantazis and Neese introduced the so-called segmented all-electron relativistically contracted (SARC) Gaussian basis set library optimized for the DKH2 and the ZORA two-component Hamiltonians, which are denoted as SARC-DKH2 and SARC-ZORA, respectively (Pantazis et al. 2008; Pantazis and Neese 2009, 2011, 2012, 2014). Another basis set library optimized for the ZORA Hamiltonian comes in combination with the ADF software package (Van Lenthe and Baerends 2003). It covers all elements in the periodic table and is design for Slater-type orbitals (as ADF works with this type of orbital basis sets). If the ZORA Hamiltonian is activated in an ADF calculation, the program automatically selects the correct, relativistic basis set from its basis set library. To the best of our knowledge, there currently aren't any publicly available basis sets for the IOTC and X2C relativistic Hamiltonians.

In general, all basis sets are of comparable quality. The user should, however, ensure that the basis set was chosen in conjunction with the two-component Hamiltonian.

Which Two-Component Hamiltonians to Choose: The Good, the Bad, and the Ugly

Two-component methods are particularly useful for modeling electronic structures of heavy-element compounds and predicting their properties. Specifically, all of the two-component Hamiltonians described in this chapter allow for an accurate description of valence properties of atoms and molecules, such as excitation energies, vibrational spectroscopy, bond lengths, and potential energy surfaces. As an example, we will consider the spin-free (vertical) electronic spectra of the UO_2^{2+} molecule calculated by time-dependent density functional theory using the ZORA and DC Hamiltonians. The electronic spectra for each method are presented in Table 3. In general, the ZORA excitation energies match very well the DC reference data, with differences being between 0.01 and 0.2 eV. We should note that ZORA and DC have been calculated in different software packages, DIRAC and ADF, respectively, and hence different basis sets have been used. This, however, does not pose any problems because both basis sets are of similar quality (Tecmer et al. 2011). Another example of the good performance of two-component theories is the reproduction of the highly accurate, experimentally determined potential energy surface of the RbYb molecule using the DKH3 Hamiltonian, where an agreement with experiment was found within 1 cm^{-1} (Borkowski et al. 2013). Usually the accuracy of quantum chemistry methods spans an error bound of several tens of cm^{-1} . The remarkable agreement between theoretical chemistry and experiment represents one of the most impressive success stories of two-component theories.

Although two-component Hamiltonians are excellent in modeling valence properties, large differences between two-component approaches are observed for properties of core electrons. Table 4 lists selected eigenstates of the Rn^{85+} one-electron atom obtained from the IOTC, DKH2, scaled ZORA, and the ZORA Hamiltonians. Among those two-component Hamiltonians, IOTC is expected to be the most accurate one, i.e., yielding results closest to the four-component Dirac

Table 3 UO_2^{2+} vertical excitation energies obtained from the spin-free ZORA and DC Hamiltonians (in eV) (Data taken from Tecmer et al. (2011))

Electronic state	ZORA/B3LYP	DC/B3LYP
$1^3\Phi_g$	2.30	2.27
$1^1\Phi_g$	2.80	2.92
$1^3\Delta_g$	2.51	2.58
$1^1\Delta_g$	3.39	3.59
$1^3\Gamma_g$	4.01	4.04
$1^1\Gamma_g$	4.23	4.31
$1^3\Phi_u$	4.22	4.23
$1^1\Phi_u$	4.31	4.33

Table 4 Selected eigenstates of the Rn^{85+} one-electron system in E_h (Data taken from Iliáš and Saue 2007)

spinor	IOTC	DKH2	Scaled ZORA	ZORA
1 $s_{1/2}$	−4154.662 453	−4142.265 045	−4154.779 786	−4670.480 658
1 $p_{1/2}$	−1070.029 516	−1069.130 996	−1070.029 629	−1101.405 752
1 $p_{3/2}$	−948.451 386	−948.434 642	−948.451 386	−973.023 298
1 $d_{3/2}$	−425.136 259	−425.135 655	−425.136 259	−430.003 698
1 $d_{5/2}$	−415.485 189	−415.485 075	−415.485 190	−420.132 939
1 $f_{5/2}$	−234.522 726	−234.522 723	−234.522 727	−235.996 352
1 $f_{7/2}$	−234.564 803	−232.564 801	−232.564 803	−234.013 854

reference. It is evident from Table 4 that the ZORA eigenstates need to be scaled to match those of the other Hamiltonians. In the valence region, DKH2 and scaled ZORA give eigenstates that are almost indistinguishable from IOTC but considerably differ in the core region. Specifically, for the lowest eigenstate $s_{1/2}$, the DKH2 Hamiltonian is much less accurate than the scaled ZORA Hamiltonian. We should note that the quality of DKH can be increased by using higher-order corrections. In practical applications, the DKH10 Hamiltonian gives satisfactory results (Reiher and Wolf 2014), and higher-order corrections are usually not required (and usually not supported in quantum chemistry software). The erroneous behavior of the core region will have a significant effect on properties that depend on core electrons. Examples are X-ray spectra, Mössbauer spectra, and hyperfine interactions observed in electron paramagnetic resonance (EPR) and nuclear magnetic resonance (NMR) experiments.

Therefore, one should be particularly careful when calculating such hyperfine interactions. It is well known that these properties strongly depend on the applied nuclear model in heavy-element compounds. While for the lighter elements the point-charge-nucleus model is sufficiently accurate, finite-size-nucleus models are more appropriate for heavier elements, like the Gaussian charge distribution (Visser et al. 1987; Visscher and Dyaal 1997). For example, a recent study by Autschbach (2009) shows the reduction in the isotropic coupling between a sixth-row nuclei and a lighter nuclei of approximately 10 % when a finite-size-nucleus model was used instead of the point-charge model.

When Do We Need Spin–Orbit Coupling?

For the case of ground-state electronic structures of heavy-element compounds with *no* open-shell electrons, a scalar relativistic treatment is sufficiently accurate. Spin–orbit coupling becomes, however, important for open-shell systems with degenerate (ground or excited) states. An illustrative example to demonstrate the importance and effect of spin–orbit coupling is the CUO molecule diluted in a Ne and Ar noble gas matrix. Furthermore, the electronic structure of the CUO molecule embedded

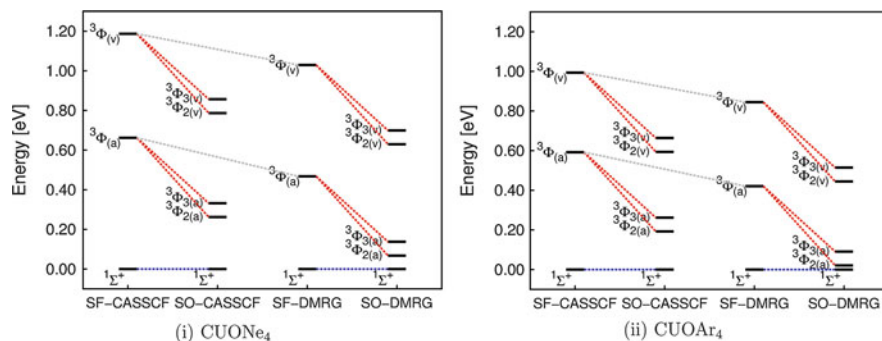


Fig. 1 Spin-state splittings for the CUO molecule diluted in different noble gas matrices within a spin-free (SF) treatment (DKH10 Hamiltonian). Spin-orbit (SO) coupling was added in a perturbative manner. ${}^3\Phi_{(v)}$: vertically excited ${}^3\Phi$ state. ${}^3\Phi_{(a)}$: adiabatically excited ${}^3\Phi$ state. CASSCF: complete active space self-consistent field. DMRG: density matrix renormalization group (Reproduced from Tecmer et al. (2014) with permission from the PCCP Owner Societies)

in a noble gas environment is a particularly challenging system for present-day quantum chemistry. Experimentalists anticipate that the CUO unit reverts its ground-state electronic configuration when the noble gas matrix is changed from Ne to Ar. This ground-state reversal could not be confirmed theoretically for a long time and the interaction between the CUO molecule and the noble gas matrix remained a “mystery.” It was only recently when quantum chemistry was able to predict a ground-state change in the CUO unit. Most importantly, this ground-state change is only observable after inclusion of spin-orbit coupling (Tecmer et al. 2014). Figure 1 shows the spin-state splittings between the lowest-lying electronic states of the CUO unit with (SO) and without (SF) spin-orbit coupling determined for two different quantum chemistry methods (CASSCF and DMRG). Only if spin-orbit coupling is included, the ${}^1\Sigma$ and ${}^3\Phi$ state of CUO diluted in Ar becomes equienergetic (see SO-DMRG data in Fig. 1) suggesting that a ground-state reversal can occur.

Therefore, if one is interested in reaching spectroscopic accuracy for open-shell systems to predict spectroscopic constants and electronic spectra with two-component Hamiltonians, spin-orbit coupling should be included. All two-component methods are equally good for this purpose and the user is free to choose the Hamiltonian of their choice. We should, however, emphasize that the largest spin-orbit splittings are observed for degenerate core orbitals, for instance, the $2p$ atomic orbitals of heavy elements. As a result, all properties that involve core orbitals (e.g., X-ray spectra, hyperfine interactions) will be considerably affected by spin-orbit coupling. This means that even for a qualitative analysis of core properties the inclusion of spin-orbit coupling is indispensable. If such properties are to be modeled with two-component Hamiltonians, special care should be taken, primarily because the chosen Hamiltonian should be accurate in the core region. Good candidates for such an analysis are, for instance, Scaled ZORA, X2C, IOTC, nARE, and higher-order DKH (e.g., DKH10) Hamiltonians.

Concluding Remarks

In this chapter we have scrutinized the most promising relativistic two-component Hamiltonians used in quantum chemistry. These Hamiltonians can be grouped into two categories: those based on the elimination of the small component (ZORA, FORA, and NESC) and those using a unitary transformation (DKH, IOTC, X2C, and nARE). All of them are equally good for valence electronic structures of atoms and molecules and can accurately model valence properties such as geometries, vibrational frequencies, and excitation energies of heavy-element molecules. The largest differences between the relativistic two-component Hamiltonians are found for the core region. Specifically, one should only use the most accurate two-component methods (Scaled ZORA, X2C, IOTC, and nARE) whenever properties of heavy-element compounds are to be modeled that directly depend on the core orbitals. Furthermore, a relativistic calculation should be performed in combination with a relativistic basis set specifically optimized for a given two-component Hamiltonian.

Finally, we would like to stress that in order to achieve spectroscopic accuracy with two-component Hamiltonians, spin-orbit coupling should be included. Exceptions to this rule are valence properties of closed-shell molecules. Furthermore, one should also keep in mind that for the heaviest elements explicit inclusion of spin-orbit coupling is preferred as an perturbative treatment might not be sufficiently accurate.

Acknowledgements We gratefully acknowledge the financial support from the National Science Center Poland, Grants No. DEC-2012/07/B/ST4/01347 and No. DEC-2013/11/B/ST4/00771. K.B. acknowledges financial support from a SONATA BIS grant of the National Science Centre, Poland (no. 2015/18/E/ST4/00584). We had many helpful discussions with Prof. Markus Reiher and Prof. Lucas Visscher.

Bibliography

- Autschbach, J. (2009). Magnitude of finite-nucleus-size effects in relativistic density functional computations of indirect NMR nuclear spin-spin coupling constants. *Chem Phys Chem*, *10*, 2274.
- Autschbach, J. (2012). Perspective: Relativistic effects. *Journal of Chemical Physics*, *136*, 150902.
- Autschbach, J., Peng, D., & Reiher, M. (2012). Two-component relativistic calculations of electric-field gradients using exact decoupling methods: Spin-orbit and picture-change effects. *Journal of Chemical Theory and Computation*, *8*, 4239.
- Balabanov, N. B., & Peterson, K. A. (2005). Systematically convergent basis sets for transition metals. I. All-electron correlation consistent basis sets for the 3d elements Sc–Zn. *Journal of Chemical Physics*, *123*, 064107.
- Barysz, M., & Sadlej, A. J. (2002). Infinite-order two-component theory for relativistic quantum chemistry. *Journal of Chemical Physics*, *116*, 2696.
- Barysz, M., Sadlej, A. J., & Snijders, J. G. (1997). Nonsingular two/one-component relativistic Hamiltonians accurate through arbitrary high order in alpha. *International Journal of Quantum Chemistry*, *65*, 225.

- Borkowski, M., Żuchowski, P. S., Ciuryło, R., Julienne, P. S., Kędziera, D., Mentel, L., Tecmer, P., Münchow, F., Bruni, C., & Görliitz, A. (2013). Scattering lengths in isotopologues of the RbYb system. *Physical Review A*, *88*, 052708.
- Bross, D. H., & Peterson, K. A. (2014). Correlation consistent, Douglas–Kroll–Hess relativistic basis sets for the 5p and 6p elements. *Theoretica Chimica Acta*, *133*, 1.
- Brown, G. E., & Ravenhall, D. G. (1951). On the interaction of two electrons. *Proceedings of the Royal Society of London Series A*, *208*, 552.
- Chang, C., Pelissier, M., & Durand, P. (1986). Regular two-component Pauli-like effective Hamiltonians in Dirac theory. *Physica Scripta*, *34*, 394.
- Dirac, P.A.M. (1928a). The quantum theory of the electron. *Proceedings of the Royal Society of London Series A*, *117*, 610.
- Dirac, P.A.M. (1928b). The quantum theory of the electron part II. *Proceedings of the Royal Society of London Series A*, *118*, 351.
- Douglas, N., & Kroll, N. M. (1974). Quantum electrodynamic corrections to fine-structure of helium. *Annals of Physics*, *82*, 89.
- Dyall, K. G. (1994). An exact separation of the spinfree and spindependent terms of the dirac–coulomb–breit Hamiltonian. *Journal of Chemical Physics*, *100*, 2118.
- Dyall, K. G. (1997). Interfacing relativistic and nonrelativistic methods. I. Normalized elimination of the small component in the modified Dirac equation. *Journal of Chemical Physics*, *106*, 9618.
- Dyall, K. G. (1998). Interfacing relativistic and nonrelativistic methods. II. Investigation of a low-order approximation. *Journal of Chemical Physics*, *109*, 4201.
- Dyall, K. G., & Faegri K., Jr. (2007). *Introduction to Relativistic Quantum Chemistry*. Oxford: Oxford University Press.
- Dyall, K. G., & van Lenthe, E. (1999). Relativistic regular approximations revisited: An infinite-order relativistic approximation. *Journal of Chemical Physics*, *111*, 1366.
- Filatov, M. (2002). *Relativistic Ab Initio Theory using the Regular Approximation*. Weinheim, Amsterdam, Wiley.
- Foldy, L. L., & Wouthuysen, S. A. (1950). On the Dirac theory of spin 1/2 particles and its non-relativistic limit. *Physical Review*, *78*, 29.
- Hess, B. A. (1985). Applicability of the no-pair equation with free-particle projection operators to atomic and molecular structure calculations. *Physical Review A*, *32*, 756.
- Hess, B. A., Marian, C. M., Wahlgren, U., & Gropen, O. (1996). A mean-field spin-orbit method applicable to correlated wavefunctions. *Chemical Physics Letters*, *251*, 365.
- Heully, J. L., Lindgren, I., Lindroth, E., & Lundqvist, S., Martensson-Pendrill, A. M. (1986). Diagonalisation of the Dirac Hamiltonian as a basis for a relativistic many-body procedure. *Journal of Physics B: Atomic and Molecular Physics*, *19*, 2799.
- Hill, J. G. (2013). Gaussian basis sets for molecular applications. *International Journal of Quantum Chemistry*, *113*, 21.
- Iliáš, M., & Saue, T. (2007). An infinite-order two-component relativistic Hamiltonian by a simple one-step transformation. *Journal of Chemical Physics*, *126*, 064102.
- Kędziera, D. (2005). Convergence of approximate two-component Hamiltonians: How far is the Dirac limit. *Journal of Chemical Physics*, *123*, 074109.
- Kędziera, D. (2006). Solving of the infinite-order two-component method equations. In *Recent Progress in Computational Sciences and Engineering* (VSP BV-C/O BRILL ACAD PUBL, Leiden, The Netherlands), (Lecture Series on Computer and Computational Sciences, vol. 7A–B, pp. 252–255)
- Kędziera, D., & Barysz, M. (2004). Two-component relativistic methods for the heaviest elements. *Journal of Chemical Physics*, *121*, 6719.
- Kędziera, D., & Barysz, M. (2007). Non-iterative approach to the infinite-order two-component (iotc) relativistic theory and the non-symmetric algebraic riccati equation. *Chemical Physics Letters*, *446*, 176.
- Kędziera, D., & Kaczmarek-Kędziera, A. (2012). Remarks on wave function theory and methods. In J. Leszczynski (Ed.), *Handbook of computational chemistry* (pp. 55–93). Netherlands: Springer.

- Kello, V., & Sadlej, A. J. (1997). Picture change and calculations of expectation values in approximate relativistic theories. *International Journal of Quantum Chemistry*, *68*, 159.
- Liu, W. (2010). Ideas of relativistic quantum chemistry. *Molecular Physics*, *108*, 1679.
- Liu, W., & Peng, D. (2009). Exact two-component Hamiltonians revisited. *Journal of Chemical Physics*, *131*, 031104.
- Mastalerz, R., Reiher, M. (2008). Relativistic electronic structure theory for molecular spectroscopy. In M. Quack & F. Merkt (Eds.), *Handbook of High-Resolution Spectroscopy* (pp. 405–442). Wiley
- Nakajima, T., & Hirao, K. (2000). The higher-order Douglas–Kroll transformation. *Journal of Chemical Physics*, *113*, 7786.
- Noro, T., Sekiya, M., Koga, T. (2012). Segmented contracted basis sets for atoms H through Xe: Sapporo-(DK)-nZP sets (n = D, T, Q). *Theoretica Chimica Acta*, *131*, 1124.
- Noro, T., Sekiya, M., & Koga, T. (2013). Sapporo-(DKH3)-nZP (n = D, T, Q) sets for the sixth period s-, d-, and p-block atoms. *Theoretica Chimica Acta*, *132*, 1363.
- Noro, T., Sekiya, M., Koga, T., & Saito, S. L. (2009). Relativistic contracted Gaussian-type basis functions for atoms K through Xe. *Chemical Physics Letters*, *481*, 229.
- Pantazis, D. A., & Neese, F. (2009). All-electron scalar relativistic basis sets for the lanthanides. *Journal of Chemical Theory and Computation*, *5*, 2229.
- Pantazis, D. A., & Neese, F. (2011). All-electron scalar relativistic basis sets for the actinides. *Journal of Chemical Theory and Computation*, *7*, 677.
- Pantazis, D. A., & Neese, F. (2012). All-electron scalar relativistic basis sets for the 6p elements. *Theoretica Chimica Acta*, *131*, 1292.
- Pantazis, D. A., & Neese, F. (2014). All-electron basis sets for heavy elements. *WIREs Computational Molecular Science*, *4*, 363.
- Pantazis, D. A., Chen, X. Y., Landis, C. R., & Neese, F. (2008). All-electron scalar relativistic basis sets for third-row transition metal atoms. *Journal of Chemical Theory and Computation* *4*, 908.
- Peng, D., Middendorff, N., Weigend, F., & Reiher, M. (2013). An efficient implementation of two-component relativistic exact-decoupling methods for large molecules. *Journal of Chemical Physics*, *138*, 184105.
- Peng, D., & Reiher, M. (2012). Exact decoupling of the relativistic Fock operator. *Theoretical Chemistry Accounts*, *131*, 1081.
- Peterson, K. A. (2015). Correlation consistent basis sets for actinides. I. the Th and U atoms. *Journal of Chemical Physics*, *142*, 074105.
- Peterson, K. A., & Dyal, K. G. (2015). Gaussian basis sets for lanthanide and actinide elements: Strategies for their development and use. In M. Dolg (Ed.), *Computational Methods in Lanthanide and Actinide Chemistry* (pp. 195–216). Wiley.
- Prascher, B. P., Woon, D. E., Peterson, K. A., Dunning, T. H., Jr., & Wilson, A. K. (2011). Gaussian basis sets for use in correlated molecular calculations. VII. Valence, core-valence, and scalar relativistic basis sets for Li, Be, Na, and Mg. *Theoretica Chimica Acta*, *128*, 69.
- Pyykkö, P. (1988). Relativistic effects in structural chemistry. *Chemical Reviews*, *88*, 563.
- Reiher, M. (2006). Douglas–Kroll–Hess theory: A relativistic electrons-only theory for chemistry. *Theoretical Chemistry Accounts*, *116*, 241.
- Reiher, M. (2012). Relativistic Douglas–Kroll–Hess theory. *WIREs Computational Molecular Science*, *2*, 139.
- Reiher, M., & Wolf, A. (2004a). Exact decoupling of the Dirac Hamiltonian. I. General theory. *Journal of Chemical Physics*, *121*, 2037.
- Reiher, M., & Wolf, A. (2004b). Exact decoupling of the Dirac Hamiltonian. II. The generalized Douglas–Kroll–Hess transformation up to arbitrary order. *Journal of Chemical Physics*, *121*, 10945.
- Reiher, M., & Wolf, A. (2014). *Relativistic Quantum Chemistry: The Fundamental Theory of Molecular Science*. Weinheim, Amsterdam. Wiley.
- Roos, B. O., Lindh, R., Malmqvist, P. A., Veryazov, V., Widmark, P. O. (2004). Main group atoms and dimers studied with a new relativistic ANO basis set. *Journal of Physical Chemistry A*, *108*, 2851.

- Roos, B. O., Lindh, R., Malmqvist, P. A., Veryazov, V., & Widmark, P. O. (2005a). New relativistic ANO basis sets for transition metal atoms. *Journal of Physical Chemistry A*, *109*, 6575.
- Roos, B. O., Lindh, R., Malmqvist, P. A., Veryazov, V., & Widmark, P. O. (2005b). New relativistic ANO basis sets for actinide atoms. *Chemical Physics Letters*, *409*, 295.
- Roos, B. O., Lindh, R., Malmqvist, P. A., Veryazov, V., Widmark, P. O., & Borin, A. C. (2008). New relativistic atomic natural orbital basis sets for lanthanide atoms with applications to the Ce diatom and LuF₃. *Journal of Physical Chemistry A*, *112*, 11431.
- Saue, T. (2011). Relativistic Hamiltonians for chemistry: A primer. *ChemPhysChem*, *3*, 3077.
- Schimmelpennig, B., Maron, L., Wahlgren, U., Teichteil, C., Fagerli, H., & Gropen, O. (1998a). On the combination of ECP-based CI calculations with all-electron spin-orbit mean-field integrals. *Chemical Physics Letters*, *286*, 261.
- Schimmelpennig, B., Maron, L., Wahlgren, U., Teichteil, C., Fagerli, H., & Gropen, O. (1998b). On the combination of ECP-based CI calculations with all-electron spin-orbit mean-field integrals. *Chemical Physics Letters*, *286*, 267.
- Schwerdtfeger, P. (2002). *Relativistic Electronic Structure Theory. Part I. Fundamentals*. Weinheim, Amsterdam. Elsevier.
- Sekiya, M., Noro, T., Koga, T., Saito, S. L. (2010). Relativistic correlating basis sets for 57Ia and 89Ac. *Journal of Computational Chemistry*, *31*, 497.
- Sikkema, J., Visscher, L., Saue, T., & Ilias, M. (2009). The molecular mean-field approach for correlated relativistic calculations. *Journal of Chemical Physics*, *131*, 124116.
- Tecmer, P., Gomes, A.S.P., Ekström, U., & Visscher, L. (2011). Electronic spectroscopy of UO²⁺₂, NUO⁺ and NUN: An evaluation of time-dependent density functional theory for actinides. *Physical Chemistry Chemical Physics*, *13*, 6249.
- Tecmer, P., Gomes, A.S.P., Knecht, S., & Visscher, L. (2014). Communication: Relativistic Fock-space coupled cluster study of small building blocks of larger uranium complexes. *Journal of Chemical Physics*, *141*, 041107.
- Tecmer, P., Boguslawski, K., Legeza, Ö., & Reiher, M. (2014). Unravelling the quantum-entanglement effect of noble gas coordination on the spin ground state of CUO. *Physical Chemistry Chemical Physics*, *16*, 719.
- Van Lenthe, E., & Baerends, E. J. (2003). Optimized Slater type basis sets for the elements 1-118. *Journal of Computational Chemistry*, *24*, 1142.
- van Lenthe, E., Baerends, E. J., & Snijders, J. G. (1993). Relativistic regular twocomponent Hamiltonians. *Journal of Chemical Physics*, *99*, 4597.
- van Lenthe, E., Baerends, E. J., & Snijders, J. G. (1994). Relativistic total energy using regular approximations. *Journal of Chemical Physics*, *101*, 9783.
- van Lenthe, E., van Leeuwen, R., Baerends, E. J., & Snijders, J. G. (1996). Relativistic regular two-component Hamiltonians. *International Journal of Quantum Chemistry*, *57*, 281.
- van Lenthe, E., Ehlers, A., & Baerends, E. J. (1999). Geometry optimizations in the zero order regular approximation for relativistic effects. *Journal of Chemical Physics*, *110*, 8943.
- van Wüllen, C. (2004). *Journal of Chemical Physics*, *120*, 7307.
- Visscher, L. (2013). In P. O. Widmark (Ed.), *Lecture Notes in Quantum Chemistry: European Summerschool in Quantum Chemistry* (pp. 5–43). Springer.
- Visscher, L., & Dyal, K. (1997). Atomic electronic structure calculations using different nuclear charge distributions. *Atomic Data and Nuclear Data Tables*, *67*, 207.
- Visser, O., Aerts, P., Hegarty, D., & Nieuwpoort, W. (1987). The use of Gaussian nuclear charge distributions for the calculation of relativistic electronic wavefunctions using basis set expansions. *Chemical Physics Letters*, *134*, 34.
- Wolf, A., Reiher, M., & Hess, B. A. (2002). In P. Schwerdtfeger (Eds.), *Relativistic Quantum Chemistry* (Theoretical and Computational Chemistry, pp. 622–663). Elsevier.
- Wolf, A., Reiher, M., & Hess, B. A. (2002). The generalized Douglas–Kroll transformation. *Journal of Chemical Physics*, *117*, 9215.
- Zou, W., Filatov, M., & Cremer, D. (2011). An improved algorithm for the normalized elimination of the small-component method. *Theoretica Chimica Acta*, *130*, 633.

Daniel Escudero, Adèle D. Laurent, and Denis Jacquemin

Contents

Introduction	928
Getting Started with TD-DFT: Vertical Absorption	930
Setting Up the Calculation	930
Analyzing the Results	933
Emission and Vibronic Effects	939
Fluorescence	940
0-0 Energies and Band Shapes	942
Photophysics of Transition Metal Complexes: Effect of SOCs, Phosphorescence, and TD-DFT Performance	946
Influence of SOCs on the UV/Vis Properties	948
Phosphorescence	949
TD-DFT Performance	951
Complex Environments and Biomolecules	953
Setting Up the Calculation	953
Toward Optical Properties	956
Conclusions	956
Bibliography	957

Abstract

The accurate description of electronically excited states remains a challenge for theoretical chemistry. Among the vast body of quantum mechanical methods available to perform this task, time-dependent density functional theory

D. Escudero (✉) • A.D. Laurent (✉)
CEISAM, UMR CNRS 6230, Université de Nantes, Nantes, France
e-mail: Daniel.Escudero@univ-nantes.fr; Adele.Laurent@univ-nantes.fr

D. Jacquemin (✉)
CEISAM, UMR CNRS 6230, Université de Nantes, Nantes, France
Institut universitaire de France, Paris, France
e-mail: Denis.Jacquemin@univ-nantes.fr

(TD-DFT) currently remains the most widely applied method, a success that one can explain not only by its very interesting accuracy/effort ratio but also by the ease to perform TD-DFT calculations for a large number of compounds and properties (absorption and emission spectra, band shapes, dipole moments, electron and proton transfers, etc.) in various environments. In the present chapter, we present TD-DFT as a tool for modeling such excited-state properties, with a focus on several practical aspects (choosing an exchange-correlation functional and an atomic basis set, analyzing the nature of the electronic transitions, comparing results with experiments, including environmental effects, etc.) that are useful to get a quick start. In that framework we rely on a series of examples of increasing complexity considering both organic and inorganic compounds as well as biomolecules.

Introduction

The chemistry occurring in electronically excited states (ES) tends to significantly differ from the one in the ground state (GS). This difference can be illustrated by considering photochemistry: the promotion of a molecule to one of its ES, induced by photon absorption, allows to perform chemical reactions that are impossible in the GS. However, the description and understanding of ES processes is very challenging for experimental methods because ES, which represent a less stable state of the matter than GS, are extremely short-lived, meaning that expensive experimental apparatus are often required to track the ES reactions and to extract relevant information. This is why theoretical chemistry is often used to complement experimental measures in this field: it offers insightful descriptions of ES properties at a reasonable cost. However, the quantum mechanical description of ES is not straightforward, as many of the theoretical “tricks” that are used for GS become inapplicable for ES, e.g., the total energy of an ES is not the minimal energy of the system, and such criterion cannot be used to help converging the calculations. Among the different *ab initio* and DFT theories that have been developed to model ES, time-dependent density functional theory (TD-DFT) is not only the most widely applied method (ca. 1000 papers/year during the last few years) but also one of the most handy. TD-DFT is the extension of DFT to ES and it originally relies on the theorem published by Runge and Gross in 1984. While DFT can provide, in principle, an exact solution of the time-independent Schrödinger equation, TD-DFT provides the same feat for its time-dependent counterpart. In a key contribution in 1995, Casida indeed provided an exact expression for calculating the excitation energies with TD-DFT (Casida 1995). However, to be applicable in practical calculations, TD-DFT generally requires two major approximations. First, one applies the adiabatic approximation (Casida 1995), that is, a frequency-independent exchange, and correlation kernel is considered, which induces a loss of memory

effects. It should be clear to the readers that all the results and advices given below are within this adiabatic approximation framework. Second, like in GS DFT, one has to select an approximate exchange–correlation functional (XCF) to perform the calculations. Over the years, there has been an impressive quantity of benchmarks performed in the TD-DFT framework, and it turned out that TD-DFT tends to be more sensitive than DFT to the chosen XCF (Laurent and Jacquemin 2013). As the present contribution is designed to be a practical guide rather than a general review, we will not describe in the following the underlying mathematical aspects of TD-DFT nor the latest fundamental developments in the field, and we refer the interested readers to other contributions for these aspects (Marques et al. 2012; Ullrich 2012; Barbatti and Crespo-Otero 2015; Huix-Rotllant et al. 2015).

The popularity of TD-DFT can be explained by several advantages compared to alternative theories: (i) it is a single-reference method that does not require to build complex input files contrary to most multi-reference theories; (ii) through the so-called Z-vector approach, one has access to the ES density at a reasonable computational cost, which offers an appropriate way to describe the changes of electron localization between the GS and ES; (iii) one can combine TD-DFT with a vast panel of models describing the environment so that one can model compounds in “real-life” situations, e.g., anchored on a semiconducting surface, near a metallic nanoparticle, embedded in a protein or solvated in an organic solvent; (iv) both analytic first (gradient) and second (Hessian) derivatives are available, making ES geometry optimizations and vibrational calculations efficient and giving access to a large panel of properties, e.g., fluorescence spectra and band shapes; (v) it is possible to efficiently determine spin–orbit couplings (SOC) with TD-DFT and hence to describe phenomena that imply changes of the spin multiplicity, e.g., triplet–singlet transitions; (vi) the scaling of TD-DFT with system size is formally $\mathcal{O}(N^4)$ which is more advantageous than most wave function approaches designed for ES; and (vii) the associated computational requirements are low enough to allow calculations on large molecules (ca. 300 atoms) that are more and more the focus of experimental studies. In short, TD-DFT is now a mature theory that can be routinely used to treat realistic chemical problems. Of course, as any approximated theory, adiabatic TD-DFT suffers from several problems and limitations, some of which will be discussed here.

In section “Getting Started with TD-DFT: Vertical Absorption,” we present the key elements necessary to perform the relatively simple vertical TD-DFT calculations. Which method to select? How to interpret the obtained results? This section can therefore be viewed as a “cookbook” for first-time users and is not intended to cover all possibilities nor problems. In section “Emission and Vibronic Effects,” we turn to the calculation of emission wavelengths, 0-0 energies and band shapes. Sections “Photophysics of Transition Metal Complexes: Effect of SOCs, Phosphorescence, and TD-DFT Performance” and “Complex Environments and Biomolecules” are devoted to the specific challenges of inorganic complexes and biomolecules, respectively.

Getting Started with TD-DFT: Vertical Absorption

As a generic rule, the modeling of ES properties with TD-DFT generally starts by performing a so-called vertical calculation. In such calculation, one uses a frozen GS geometry, typically obtained through a DFT optimization or from experimental inputs, and determines the transition energies to the first few low-lying ES using TD-DFT. These transition energies are referred to as vertical absorption energies, $E^{\text{vert-a}}$, in the following. They can be formalized as:

$$E^{\text{vert-a}} = E^{\text{ES}}(R^{\text{GS}}) - E^{\text{GS}}(R^{\text{GS}}), \quad (1)$$

where E^{ES} and E^{GS} are, respectively, the total energies of the ES and GS determined on the frozen GS geometry (R^{GS}). A large number of quantum chemistry packages allow to perform such calculations, and the number of parameters specific to TD-DFT that need to be set before running a vertical TD-DFT computation is often very limited: the total number of ES states (or *roots*) that should be computed and the multiplicity of the final states (e.g., singlet-singlet or singlet-triplet transitions). Of course, like in GS DFT calculations, one has also to select an atomic basis set (BS) and an XCF to perform the computations. At the end of the TD-DFT run, one obtains the $E^{\text{vert-a}}$ corresponding to each states as well as several associated information that can be analyzed.

Setting Up the Calculation

Selecting an Atomic Basis Set

As for GS calculations, the appropriate selection of a BS for TD-DFT can be a difficult task. It has been generally noticed that (i) TD-DFT is less sensitive to the size of the atomic BS than the correlated wave function theories, (ii) a basis set that is suited for a DFT calculation needs to be expanded with (at least) diffuse functions for TD-DFT, (iii) the nature of the ES has a dramatic impact on the BS requirements (while a double- ζ BS augmented with polarization and diffuse functions, e.g., 6-31+G(d), might be sufficient for low-lying ES of conjugated organic and inorganic molecules, much larger BS, e.g., *aug-cc-pVTZ*, might be necessary for describing higher-lying (Rydberg) ES), and (iv) as a consequence the ordering of the different ES might change with the size of the selected atomic BS. These different statements are well illustrated by considering the fifteen lowest-lying ES of benzene (Ciofini and Adamo 2007). For the three valence ES, Ciofini and Adamo reported that the mean absolute errors (MAE) with respect to experimental values attain 0.48, 0.25, and 0.23 eV with the 6-31G(d), 6-31+G(d,p), and 6-31++G(3df,3pd) BS, respectively, confirming the key importance of diffuse orbitals. In the same time, for the twelve first Rydberg states of benzene, the same authors obtained MAE of 2.27, 0.31, and 0.18 eV, with the same three BS, showing the difficulty to describe Rydberg states. Nevertheless, we also underline that very large BS tend

Table 1 Wavelength (in nm) corresponding to $E^{\text{vert}-a}$ of the first ES of indigo. All values obtained with the B3LYP XCF on a frozen geometry. See Jacquemin et al. (2006) for details

BS	Wavelength	BS	Wavelength
6-31G	564	6-311G(2d,2p)	581
6-31G(d,p)	566	6-311+G(2d,p)	591
6-31+G(d,p)	585	6-311++G(2d)	594
6-311G(d,p)	576	6-311++G(2d,2p)	591
6-311++G(d,p)	587	6-311+G(3d,p)	592

to yield stronger mixing between Rydberg and valence ES, and hence, TD-DFT results are harder to interpret (Isegawa et al. 2012; Baerends et al. 2013). When possible, it is always welcome to check the convergence of the transition energies of the considered ES using increasingly large atomic basis sets so to assess the magnitude of the residual error obtained with the selected BS. When treating a series of homologous molecules, such BS convergence test can be performed on a representative member of the series. This procedure is illustrated in Table 1 for the lowest ES of indigo, where the non-negligible impact of a second set of d polarization orbitals appears (Jacquemin et al. 2006). From that table, one would probably choose to go with 6-311+G(2d,p) to be very close from BS convergence.

Choosing an Adequate Exchange-Correlation Functional

As stated in the “Introduction”, the selection of an adequate XCF is extremely important to obtain accurate results. For vertical TD-DFT calculations, there have been numerous benchmarks which we have been reviewed in Laurent and Jacquemin (2013). Though the most suited XCF for a given ES in a given compound cannot be predicted beforehand, several general conclusions have been obtained: (i) “pure” XCF that do not include *exact* exchange often provide too small transition energies, and as they also tend to create unphysical (*spurious*) ES, their use should be discouraged; (ii) hybrid XCF that incorporate *exact* exchange give significantly more accurate results, increasing the amount of *exact* exchange generally inducing an increase of the transition energies; (iii) for localized valence ES, standard global hybrids, e.g., B3LYP, M06, and PBE0, are often well suited in the vertical approximation; (iv) for both charge transfer (CT) and Rydberg ES, range-separated hybrid (RSH) XCF, e.g., CAM-B3LYP, LC- ω PBE, or ω B97X-D, appear to be more adequate; (v) one can use optimally tuned (OT) range-separated hybrids, where the parameter(s) of the XCF is(are) nonempirically optimized to minimize the delocalization error (Baer et al. 2010; Srebro and Autschbach 2012), to avoid an arbitrary XCF choice, but the optimization process has to be carried out for each molecule; and (vi) when a suitable XCF is chosen, the typical theory–experiment deviations that are to be expected are in the 0.20–0.30 eV range, though significantly better (and worse) results can be obtained for specific chemical families.

The statistical errors obtained with a representative panel of functionals considering two training sets in the vertical TD-DFT approximation are given in

Table 2 Mean signed (MSE) and absolute (MAE) errors in eV obtained two sets of organic compounds (see text) in the vertical approximation. A positive MSE indicates an overestimation of the experimental transition energies. Val. and Ryd. stand for valence and Rydberg ES, respectively. The percentage of exact exchange (%XX) is given for all XCF. See Jacquemin et al. (2009a), Caricato et al. (2010), and Isegawa et al. (2012) for details about the computational protocols and additional results

XCF	%XX	614 ES – Jacquemin et al. (2009a)		69 ES – Caricato et al. (2010) and Isegawa et al. (2012)		
		MSE	MAE	MAE (Val.)	MAE (Ryd.)	MAE (All)
LDA	0	−0.32	0.45	0.45	1.20	0.88
BP86	0	−0.32	0.38	0.38	1.62	1.08
PBE	0	−0.29	0.39	0.40	1.70	1.13
TPSS	0	−0.20	0.34	0.26	1.63	1.03
TPSSh	10	−0.05	0.26	0.18	1.27	0.80
B3LYP	20	−0.01	0.23	0.20	1.03	0.67
PBE0	25	0.08	0.22	0.22	0.80	0.55
BHHLYP	50	0.36	0.40	0.56	0.36	0.44
M05-2X	52	0.29	0.38	0.37	0.35	0.36
CAM-B3LYP	19–65	0.25	0.30	0.31	0.35	0.33
LC- ω PBE	0–100	0.46	0.48	0.41	0.32	0.36

Table 2. The performances of many more XCF have been evaluated in the original works (Jacquemin et al. 2009a; Caricato et al. 2010; Isegawa et al. 2012). In the first set, 614 low-lying ES of small, medium, and large organic molecules have been considered (Jacquemin et al. 2009a), whereas in the second, 69 ES in small molecules, almost equally divided between valence (30) and Rydberg (39) states, have been used (Caricato et al. 2010; Isegawa et al. 2012). Consistently with the above statements, for valence states, the smallest MAE are obtained with three global hybrids, namely, TPSSh, B3LYP, and PBE0, a result illustrative of the large weight of localized $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transition in these training sets. For the Rydberg states, only range-separated hybrids or global hybrids including a large share of *exact* exchange deliver accurate results, the same holding for CT states (Peach et al. 2008). We underline that CAM-B3LYP, M05-2X, and M06-2X (not shown in Table 2) provide relatively similar errors irrespective of the nature of the ES.

Including Solvent Effects

Most experiments are carried out in condensed phase. As the impact of the environment is generally larger in the ES than in the GS, taking the medium into account is often necessary to reach chemically valuable conclusions. The total model is typically divided into two parts: on the one hand, the center of interest, where the electronic excitation takes place, which is described with TD-DFT, and, on the other hand, the surroundings that are modeled with a simpler model. To describe solvation effects, continuum models, such as the well-known polarizable continuum model (PCM) (Tomasi et al. 2005), have been combined with TD-DFT

and we limit ourselves to such combination at this stage. In PCM-TD-DFT, the solute undergoing the GS-to-ES transition is embedded in a cavity formed in a structureless material presenting the macroscopic properties of the actual solvent. In the GS, charges located on the surface of this cavity are determined self-consistently to account for the electrostatic interactions between the solute and the solvent (Tomasi et al. 2005). For ES calculations, two new difficulties, not present in GS, appear. First, as photon absorption is an extremely fast process, only the electrons of the environment have the time to adapt to the new electronic configuration of the solute, the nuclei of the solvent being frozen in their GS configuration. This corresponds to the so-called nonequilibrium (neq) regime, and the vertical transition energy now becomes

$$E^{\text{vert-a}}(\text{neq}) = E^{\text{ES}}(R^{\text{GS}}, \text{neq}) - E^{\text{GS}}(R^{\text{GS}}, \text{eq}). \quad (2)$$

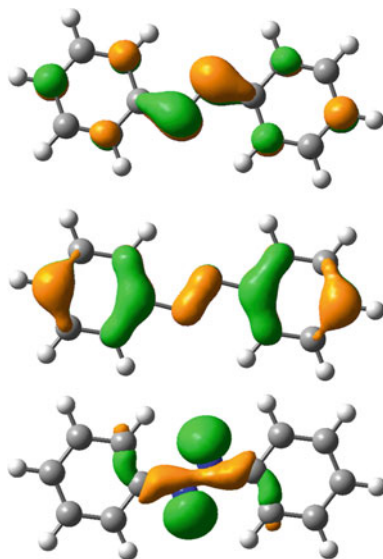
In this neq regime, the effective dielectric constant for the specific ES terms of the PCM-TD-DFT calculation is the optical permittivity (ϵ_∞), that is of the order of the square of the refractive index, i.e., around 2 for most solvents. This may induce significant differences for polar solvents compared to the equilibrium (eq) regime where the complete permittivity (ϵ_r) is used. Various default regimes (eq/neq) are applied by different computational chemistry packages, and we advice a careful check before starting the calculation. Secondly, one should determine how the charges located on the surface of the cavity are determined at the ES. In the linear-response (LR) PCM approximation (Cammi and Mennucci 1999; Cossi and Barone 2001), the variations of these charges from the GS to the ES are determined using the TD-DFT transition densities. As these densities are available in any TD-DFT calculation, this allows to rapidly include solvent effects for all ES considered in the calculation at a moderate computational cost, but this LR model is not suited when there is a strong variation of the electronic cloud between the two states, e.g., for CT ES. In those cases, more refined models that rely on the ES density to determine the polarization of the cavity in the ES could be used, an aspect discussed in more details in section “Emission and Vibronic Effects.”

Analyzing the Results

Determining the Nature of an Electronic Transition

As can be seen from the above considerations, it is useful to know beforehand the nature of the ES that is investigated as this might influence the selection of the atomic BS, XCF, and even solvent model. However, if this is not the case, one can use the results of a TD-DFT calculation to determine this nature. Indeed, besides the $E^{\text{vert-a}}$ values, a typical TD-DFT output contains (i) the oscillator strengths (f) associated to these ES, which are related to the transition probabilities from the GS to the ES (the larger the f , the more intense the absorption), (ii) the point group symmetry of each ES, and (iii) the contributions of transitions between molecular orbitals (MO) corresponding to each ES. Altogether, they can be used to assign

Fig. 1 Representation of the three key MO in *trans*-azobenzene as computed at PBE0/6-31+G(d) level. From *top* to *bottom*: LUMO, HOMO, and HOMO-1 of π^* , π , and n nature, respectively. Contour threshold: 0.05 a.u.



the nature to each ES. A simple example of such procedure can be given by *trans*-azobenzene that belongs to the C_{2h} point group. A TD-DFT calculation provides two low-lying ES. The first of B_g symmetry presents a $E^{\text{vert-a}}$ of 2.64 eV associated to f of 0.00 and is mainly ascribed to a HOMO-1 to LUMO transition, whereas the second ES of B_u symmetry presents a $E^{\text{vert-a}}$ of 3.78 eV associated to f of 0.78 and corresponds to a HOMO to LUMO transition. These orbitals are sketched in Fig. 1. The first transition is forbidden by symmetry (zero oscillator strength) and implies a transition from the nitrogen lone pairs toward an unoccupied orbital delocalized on the whole system. It is therefore a valence $n \rightarrow \pi^*$ transition. The second ES is strongly dipole allowed and corresponds to an increase of the number of nodes in the frontier π MO: it is $\pi \rightarrow \pi^*$ transition. Interestingly, there has been a reversal between the energetic ordering of the occupied MO and the corresponding ES, which illustrates the importance of the electronic reorganization accounted for in TD-DFT. These assignments are in line of experiment and are also consistent with the solvatochromic shifts that can be measured or calculated. Indeed, the same calculation considering a PCM(water) environment yields a small hypsochromic shift (from 2.64 to 2.66 eV) for the first ES, but a larger bathochromic shift (from 3.78 to 3.63 eV) for the second ES, which are typical signatures of $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions, respectively.

Unfortunately, it is generally not so straightforward to assign a clear nature to each ES mainly because the orbital compositions of the TD-DFT transitions often include very large numbers of MO pairs of different nature, especially in large molecules and/or higher-lying ES. In those situations, three strategies might be applied. First, one can compute natural transition orbitals (NTO) (Martin 2003), which is inexpensive. NTO are designed to provide a compact representation of the transition density (matrix) and allow to express the TD-DFT results for each ES

as a transition between (often a single pair of) an excited particle and an empty hole. Therefore NTO can provide a simple “MO-like” picture, irrespective of the complexity of the actual MO blend of the original TD-DFT calculation. A second approach, often used in inorganic chemistry, is to evaluate the contribution of each atom in all the starting and final MO implied in the electronic transition, so to obtain an averaged contribution of each atom in the electronic process. This can be useful to rapidly distinguish metal-to-ligand from ligand-to-ligand transitions in transition metal complexes (see section “Photophysics of Transition Metal Complexes: Effect of SOCs, Phosphorescence, and TD-DFT Performance” for specific challenges for these compounds). Finally, one can compute the ES density (ρ^{ES}) of the state(s) of interest. By making the difference with the GS density, one obtains a single picture of the “electron displacement” irrespective of the complexity of the considered ES. The drawback of this approach is that it implies the calculation of the ρ^{ES} for each ES of interest, which requires a significant additional computational effort. We illustrate this approach in the section “Going Further: ES Dipoles, Charges, and Densities.”

Comparing with Experiments

In many cases, and in particular in the high-energy range of the spectrum and/or for inorganic compounds, there are much more ES returned by the calculation than experimentally observed peaks in the optical spectrum. This is originating in the overlapping of contributions of several ES with different transition probabilities that appear as a single band, e.g., forbidden $n \rightarrow \pi^*$ transitions are often hidden in the experimental spectra by nearby strongly allowed $\pi \rightarrow \pi^*$ transitions. Consequently, to compare the TD-DFT output values ($E^{\text{vert-a}}$ and f) to experimental UV/Vis spectra, a common initial step is to convolute the computed “stick” values with a broadening Gaussian (typically with a width at half maximum in the 0.2–0.8 eV range). Many codes can be used to perform such task and refined procedures have been designed to reach an optimal fit (Brémond et al. 2010). The result of such process is illustrated in Fig. 2 where the TD-DFT sticks and convoluted spectra are compared to experiment for a Pt(II) complex (Latouche et al. 2015). The fact that numerous ES contribute to a single band in the 200–300 nm range is obvious.

As will become evident in section “Emission and Vibronic Effects,” there are often no experimental data that can be directly compared to the computed $E^{\text{vert-a}}$, so that the comparisons between experimental UV/Vis spectra and (convoluted) TD-DFT spectra remains only qualitative. Indeed, while it is usual to compare TD-DFT $E^{\text{vert-a}}$ (or the spectra obtained with a broadening Gaussian) to experimental λ_{max} (Laurent et al. 2014), it should be made clear that this approach that completely disregard geometrical and vibrational relaxation effects is not physically well grounded and is not to be favored, despite its popularity. In a homologous series of compounds, in which the vibronic couplings are rather weak, such approximation can nevertheless be useful to extract some trends regarding the impact of the substituents (Laurent et al. 2014). For instance, for 169 compounds presenting an easily identifiable low-lying $n \rightarrow \pi^*$ band, Fig. 3 provides a comparison between PCM-TD-PBE0 $E^{\text{vert-a}}$ and measured λ_{max} , and the general correlation is obvious

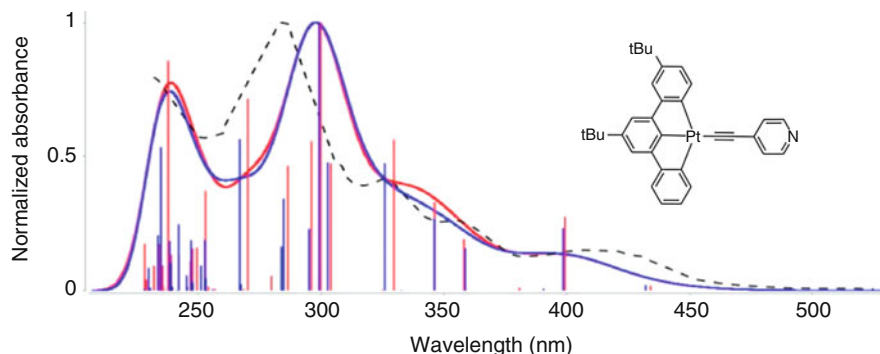
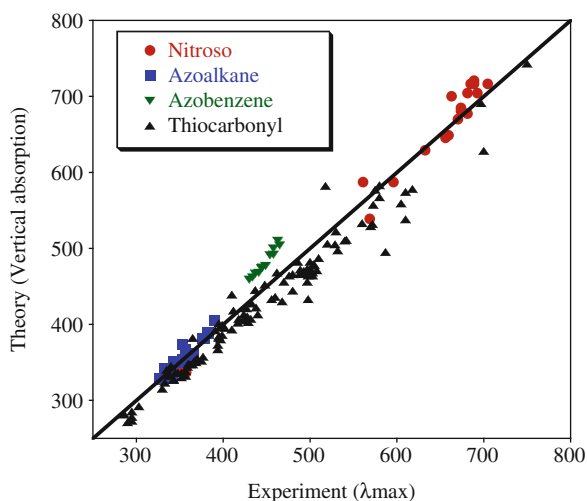


Fig. 2 Comparison between TD-DFT (two different atomic basis sets, corresponding to *blue* and *red* sticks and *curves*) and experimental (*black dashed line*) for the shown Pt complex. The position of each stick is given by $E^{\text{vert}-a}$, its height being proportional to the corresponding f (Adapted with permission from Latouche et al. (2015). Copyright 2015 American Chemical Society)

Fig. 3 Comparison between PCM-TD-PBE0 $E^{\text{vert}-a}$ and measured λ_{max} for four families of $n \rightarrow \pi^*$ dyes. All values are in nm. The *central line* indicates a perfect match (Reproduced from Laurent et al. (2014) with permission from the PCCP Owner Societies)



despite the presence of several outliers. For the full set of compounds, the obtained MAE is limited to 19 nm or 0.12 eV only, a success probably related, on the one hand, to the localized nature of the considered transitions and, on the other hand, to error compensation mechanisms.

Going Further: ES Dipoles, Charges, and Densities

One can also compute ES dipole moments, partial atomic charges, and total electron densities with TD-DFT. This provides information about the variations of polarity and charge localizations between the GS and the ES. However, determining these data require an additional effort, similar to a determination of TD-DFT forces for

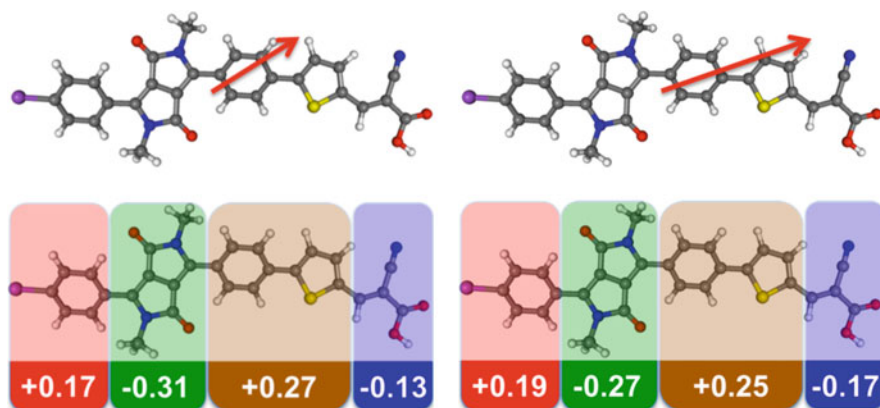


Fig. 4 *Top:* GS (*left*) and ES (*right*) dipole moments for a typical push–pull dye. *Bottom:* partial Merz-Kollman atomic charges borne by several moieties of the molecule in the GS (*left*) and ES (*right*). The first ES corresponding to a CT process is considered here (Adapted from Adamo and Jacquemin (2013) with permission from the Royal Society of Chemistry)

each considered ES. Indeed, contrary to GS, the ES density is not a “free” by-product of the calculation. In other words, if one wishes to determine ρ^{ES} for the 10 lowest-lying ES, a substantial increase of computational effort is necessary compared to standard TD-DFT calculations. Figure 4 illustrates the result of such calculation for a dye designed for dye-sensitized solar cell applications (Adamo and Jacquemin 2013). First one notes that while the overall direction of the dipole moment is similar in the GS and the ES, its magnitude has been increased by 2.93 D (from 5.31 to 8.24 D) which is typical of a CT ES. This is confirmed by investigating the partial atomic charges: the electronic charge borne by the cyano-acetic moiety (in blue) increases, while the electronic charge borne by the diketopyrrolopyrrole group (in green) decreases. It is possible to correlate these variations to experimental solvatochromic effects, that is, to investigations of the evolutions of experimental spectra with solvent polarity.

Having determined the ES density, one can easily compute the difference of electron density compared to the GS,

$$\Delta\rho(\mathbf{r}) = \rho^{\text{ES}}(\mathbf{r}) - \rho^{\text{GS}}(\mathbf{r}), \quad (3)$$

for each relevant ES, and the corresponding plot provides a one-picture representation of each electronic transition, sometimes called electron density difference map (EDDM). Examples of such maps are given in Fig. 5 for two large BODIPY dyes in their neutral and diprotonated forms (Chibani et al. 2013). In very simplified terms, the electron density moves from blue to red regions of these plots upon photon absorption. Both the left and right compounds of Fig. 5 are very sensitive to pH experimentally: the former undergoes a strong hypsochromic shift upon

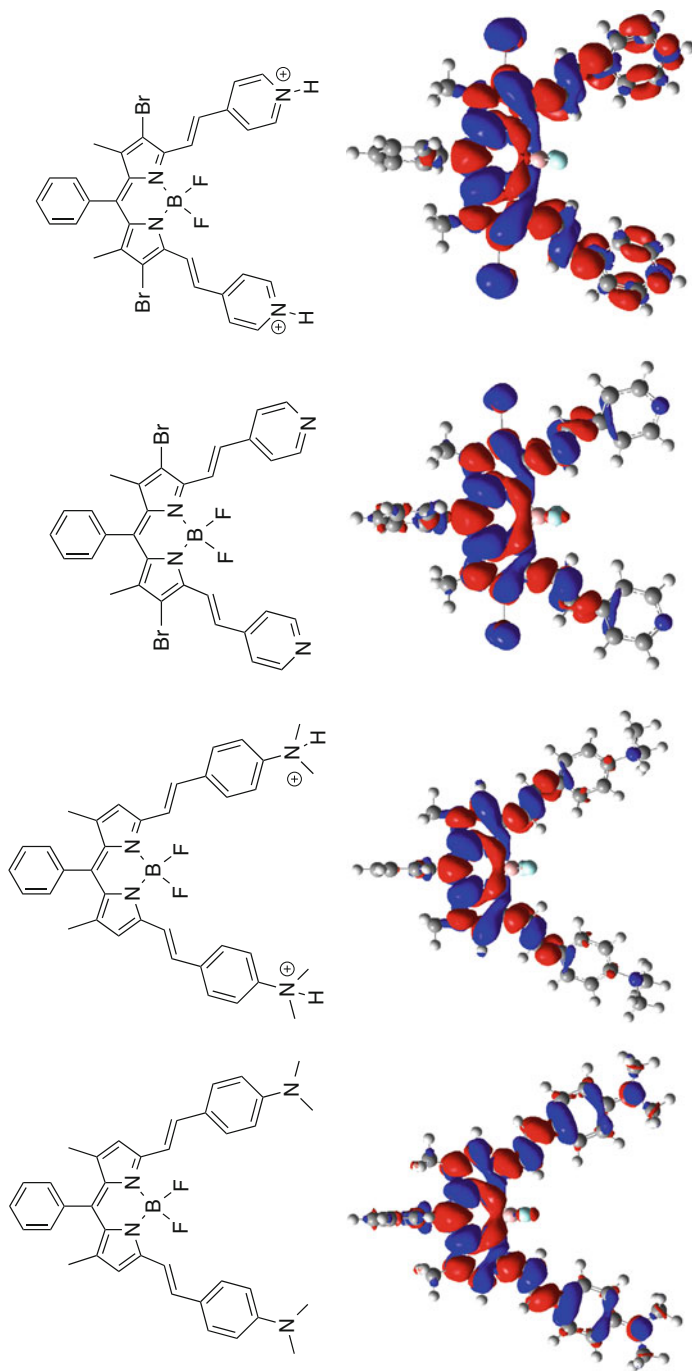


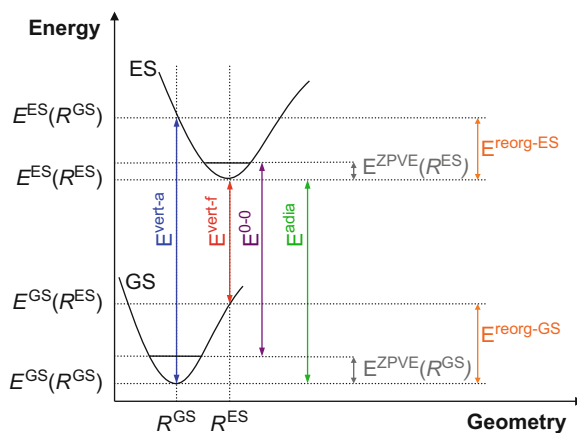
Fig. 5 $\Delta\rho(\mathbf{r})$ plots (EDDM) obtained for two BODIPY dyes in their neutral and di-protonated forms at the PCM-TD-M06-2X/6-311+G(2d,p) level. The red (blue) zones indicate increase (decrease) of density upon electronic transition (Adapted from Chibani et al. (2013) with permission from the Royal Society of Chemistry)

protonation of the two dimethylamino groups, whereas the second undergoes a significant bathochromic shift when the two pyridine rings are protonated. This can be qualitatively explained from the change of the $\Delta\rho(\mathbf{r})$ plots computed for the neutral and diprotonated derivatives. Indeed, in its neutral form, the ES of the first dye (left) is characterized by a strong CT from the dimethylamino moieties (mostly in blue, acting as donor groups) to the BODIPY core (mostly in red, acting as the accepting center), but this CT is annihilated at low pH as the lone pairs of the donor group are no longer available. In turn, this decrease of CT explains the hypsochromic displacement. The scenario is reversed for the second BODIPY (right): the pyridine rings are passive elements in the neutral form but act as strong acceptors after protonation. Therefore, the CT is enhanced at low pH, explaining the concomitant bathochromic evolution of the experimental spectrum. Eventually, we note that $\Delta\rho(\mathbf{r})$ can also be used to obtain qualitative indices quantifying the magnitude of the GS-to-ES CT (Le Bahers et al. 2011; Jacquemin et al. 2012b), a topic that we will not cover here.

Emission and Vibronic Effects

In this section, we show how going beyond the vertical absorption allows to determine vertical fluorescence energies (and hence estimate Stokes shifts) as well as 0-0 energies and band shapes. The two latter allow physically well-grounded comparisons with experimental optical spectra and are therefore extremely useful. All these data come at an increase computational cost compared to the properties listed in section “Getting Started with TD-DFT: Vertical Absorption” but this cost is mitigated by the existence of analytic first and second TD-DFT derivatives (Furche and Ahlrichs 2002; Scalmani et al. 2006; Liu and Liang 2011). The main parameters that are used in the following are represented in Fig. 6.

Fig. 6 Simplified energy diagram representing only two singlet states without intersections and describing the key theoretical parameters used in section “Emission and Vibronic Effects” (Reproduced with permissions from Jacquemin et al. (2012a). Copyright 2012, American Chemical Society)



Fluorescence

Vertical fluorescence is the transition energy determined at the optimal geometry of the relevant excited state (R^{ES}). In gas phase, it reads:

$$E^{\text{vert-f}} = E^{\text{ES}}(R^{\text{ES}}) - E^{\text{GS}}(R^{\text{ES}}), \quad (4)$$

which obviously parallels Eq. 1. One therefore needs to optimize the geometry of the ES, a task that can be performed with numerous computational chemistry packages. In many cases, one is interested in the first (lowest-energy) ES that is often responsible for fluorescence according to Kasha's empirical rule. Optimizing the geometry of higher-lying ES is technically possible but is often more challenging due to the possible crossings between the potential energy surfaces (PES) of different ES.

Setting Up the Calculation

For computing the vertical emission energy, that is, the $E^{\text{vert-f}}$ in Eq. 4, the indications given in section "Selecting an Atomic Basis Set" can be followed to select an atomic BS. However, using a more compact BS to perform the ES geometry optimization, that is, to determine R^{ES} , can be an interesting option to fasten the calculations. Indeed, one needs a BS that is typically smaller for R^{ES} than for transition energies, but typically larger than for GS structural optimizations, as the PES of the ES tend to be flatter than their GS counterparts. In a study of organic compounds, it was found that 6-31+G(d) was providing accurate ES geometries with TD-DFT (Jacquemin et al. 2012a), but significantly more extended BS may be necessary for flexible molecules.

The choice of the XCF to perform the TD-DFT geometry optimizations is crucial. For rigid molecules and for states in which the density reorganization between the GS and ES remains relatively limited (localized $\pi \rightarrow \pi^*$ transitions), global hybrids containing a small fraction of *exact* exchange, e.g., B3LYP and PBE0, perform adequately. However, when stronger variations of the density are noticed, e.g., $n \rightarrow \pi^*$ transitions, a larger share of *exact* exchange should be included in the functional to obtain accurate structures (Guido et al. 2010). This also holds for CT ES. Indeed, for these states, low-*exact* exchange hybrid functionals might lead to qualitatively incorrect ES PES. For instance, for the simple 4-(dimethylamino)-benzonitrile, B3LYP incorrectly predicts a twisted ES geometry with the dimethylamino group orthogonal to the aromatic ring (Wiggins et al. 2009). Such overtwisted PES is not uncommon with amino groups (Jacquemin et al. 2012a) and remains a common pitfall of fluorescence calculations performed with TD-DFT. Our experience is that the M06-2X XCF is a very valuable compromise as it tends to deliver chemically meaningful GS and ES structures for a vast majority of fluorophores (see also section "Adiabatic and 0-0 Energies").

In solution, the geometrical relaxation of the molecule is a relatively slow process, and the R^{ES} should be determined in the equilibrium regime. However,

the emission of a photon is a fast process and the nonequilibrium effects cannot be neglected. Indeed, in solution Eq. 4 becomes

$$E^{\text{vert-f}}(\text{neq}) = E^{\text{ES}}(R^{\text{ES}}, \text{eq}) - E^{\text{GS}}(R^{\text{ES}}, \text{neq}), \quad (5)$$

and it can be noticed that the neq regime is now applied to the GS energy, contrasting with Eq. 2. For fluorescence calculations, the charges on the PCM cavity should be equilibrated with the ES density. In practice, applying Eq. 5 requires the selection of a quite advanced PCM approach. Three models have been developed to carry out this task: the corrected linear response (cLR) (Caricato et al. 2006), the state-specific (SS) (Improta et al. 2007), and the vertical excitation models (VEM) (Marenich et al. 2011). All rely on the one-particle TD-DFT density matrix [basically on $\rho^{\text{ES}}(\mathbf{r})$] to compute the charges of the cavity in the ES but with significantly different procedures: cLR is a fast perturbative approach, whereas both SS and VEM are two more computationally demanding self-consistent schemes, the former (latter) implying a modification (no modification) of the GS density. Of course, cLR, SS, and VEM models can also be used to determine vertical absorption energies. However, contrary to absorption, the LR-PCM scheme is not well suited for emission, as $E^{\text{GS}}(R^{\text{ES}}, \text{neq})$ is not accessible with LR-PCM. This calls for two additional remarks: (i) while the transition energy obtained at the end of a gas-phase TD-DFT ES optimization corresponds to the vertical fluorescence energy, it is not the case in solution, and a separate (two-step) calculation has to be performed to determine $E^{\text{vert-f}}(\text{neq})$; (ii) to date analytic TD-DFT optimizations are only available with LR-PCM, and the cLR-PCM, SS-PCM, or VEM-PCM fluorescence energies are typically computed on geometries obtained with LR-PCM. We recall that different programs have different default eq/neq regime and that a careful check is necessary prior to starting the computations.

Analyzing the Results

Similar to absorption, one obtains at the end of a fluorescence calculation the vertical transition (emission) energy, the corresponding oscillator strength, as well as MO composition of the transition. All can be analyzed using the approaches described above. Of course, one advantage of TD-DFT compared to most experimental approaches is that it provides the ES geometry and hence allows to quantify the structural relaxation between the GS and ES. From the computed vertical absorption and fluorescence energies, one can also directly estimate this Stokes shift, although such estimation would neglect vibronic couplings. In first approximation, the Stokes shift is proportional to the structural relaxation between the two states. To avoid misunderstandings, we also underline that: (i) performing a TD-DFT geometry optimization is not equivalent to investigating photochemistry, i.e., the optimization steps connecting R^{GS} to R^{ES} are not providing a proper photochemical reaction path, and (ii) the oscillator strengths corresponding to $E^{\text{vert-f}}$ are not directly comparable to experimental emission quantum yields (Φ_f), because the latter are also affected by the (often numerous) non-radiative deactivating processes that are

not accounted for in Eq. 4, whereas the oscillator strengths are related to the radiative rate constant only.

0-0 Energies and Band Shapes

Let us now turn to properties that can be rightfully compared to experimental values: (i) 0-0 energies that correspond to the measured absorption-fluorescence crossing point (AFCP) and (ii) band shapes that (mainly) result from vibronic couplings and can also be compared to experiment. We underline that for both properties, it is nevertheless necessary to transform the experimentally reported spectra into normalized line shapes (Avila Ferrer et al. 2013). Additionally, the calculation of vibronic couplings and 0-0 energies allows one to determine molar absorptivities that also offer a direct confrontation with measurements. All these properties require the determination (or at least the estimation) of the (harmonic) ES Hessian, that is, of the second-order geometrical derivatives of the TD-DFT energy. Consequently, a significant computational effort is required but molecules with ca. 100 atoms can now be routinely modeled.

Adiabatic and 0-0 Energies

We directly discuss the solvated case here, as the simplification to the gas-phase limit is trivial. As can be seen in Fig. 6, the adiabatic energy, an equilibrium property, can be obtained as a simple by-product of absorption and emission calculations:

$$E^{\text{adia}}(\text{eq}) = E^{\text{ES}}(R^{\text{ES}}, \text{eq}) - E^{\text{GS}}(R^{\text{GS}}, \text{eq}). \quad (6)$$

To determine the 0-0 energies,

$$E^{0-0}(\text{eq}) = E^{\text{adia}}(\text{eq}) + \Delta E^{\text{ZPVE}}(\text{eq}), \quad (7)$$

one needs to determine the difference of zero-point vibrational energy (ZPVE) between the ES and GS,

$$\Delta E^{\text{ZPVE}}(\text{eq}) = E^{\text{ZPVE}}(R^{\text{ES}}, \text{eq}) - E^{\text{ZPVE}}(R^{\text{GS}}, \text{eq}). \quad (8)$$

In practice $\Delta E^{\text{ZPVE}}(\text{eq})$ can only be computed in the harmonic approximation and remains the time-limiting step of 0-0 calculations. As explained above, the PES tends to be flatter in the ES than in the GS, and $\Delta E^{\text{ZPVE}}(\text{eq})$ is negative for the vast majority of molecules and states. Consequently, $E^{0-0}(\text{eq})$ are smaller than $E^{\text{adia}}(\text{eq})$, and the difference between the two is typically in the -0.04 to -0.12 eV range (Goerigk and Grimme 2010; Send et al. 2011; Jacquemin et al. 2012a; Laurent and Jacquemin 2013) and is therefore not negligible. As the experimental AFCP energy corresponds to the crossing point between two nonequilibrium curves, the results of Eq. 7 should be further corrected using

$$E^{\text{AFCP}}(\text{neq}) = E^{0-0}(\text{eq}) + \frac{1}{2} [\Delta E^{\text{vert-a}} + \Delta E^{\text{vert-f}}], \quad (9)$$

where the two correcting terms read

$$\Delta E^{\text{vert-a}} = E^{\text{vert-a}}(\text{neq}) - E^{\text{vert-a}}(\text{eq}), \quad (10)$$

$$\Delta E^{\text{vert-f}} = E^{\text{vert-f}}(\text{neq}) - E^{\text{vert-f}}(\text{eq}). \quad (11)$$

The rationale explaining the form of these correction terms, which are strictly zero in gas phase, can be found in Jacquemin et al. (2012a).

To determine $E^{\text{AFCP}}(\text{neq})$, one can follow the advices given above, that is, use a diffuse containing atomic basis set, select a functional adapted to the nature of the investigated ES (see also below), and apply a refined environmental model. Nevertheless, we underline that $\Delta E^{\text{ZPVE}}(\text{eq})$ is less sensitive to BS and XCF effects than most other ES properties but depends significantly on the considered molecule and ES. In other words, one can generally use a less demanding approach to determine this term. Nevertheless, the calculation of $E^{\text{AFCP}}(\text{neq})$ will require a substantial number of steps, typically ordered as (i) optimization of the GS geometry, (ii) calculation of the vibrational frequencies of the GS, (iii) determination of vertical absorption energies in both equilibrium and nonequilibrium regimes, (iv) optimization of the ES geometry, (v) calculation of the vibrational frequencies of the ES, and (vi) determination of vertical emission energies in both equilibrium and nonequilibrium regimes. Detailed protocols can be found in Chibani et al. (2013) and Jacquemin et al. (2012a).

As AFCP energies allow comparisons with experiment, let us now briefly comment on the expected results and the XCF dependency. For a set of 40 representative medium-sized dyes, the $E^{\text{AFCP}}(\text{neq})$ energies have been determined using: (i) the 6-31+G(d) BS for structures (R) and the 6-311++G(2df,2p) ES for total and transition energies (E); (ii) the LR-PCM model for structures and the cLR-PCM model for energies; and (iii) 12 hybrid XCF (Jacquemin et al. 2012a, 2014; Moore et al. 2014). Selected results are given in Table 3. It turns out that the smallest MSE and MAE are obtained with PBE0 and M06, but these XCF significantly fail in a few specific compounds presenting strong CT, yielding a moderate correlation with experiment. By contrast, M06-2X tends to systematically overestimate the experimental $E^{\text{AFCP}}(\text{neq})$, but the correlation with experimental trends is significantly improved. In choosing the XCF, there is thus a tradeoff between absolute and relative accuracies. The most satisfying results are reached with the OT LC-PBE* but at the price of an additional preliminary step (determination of the range-separation parameter) for each molecule.

The valuable correlation between the experimental and theoretical evolution of the $E^{\text{AFCP}}(\text{neq})$ energies in a homologous series of compounds can be illustrated by considering BODIPY derivatives (Chibani et al. 2013). This class of molecules can be viewed as a subgroup of the cyanine family. Although TD-DFT has a strong tendency to overestimate the absolute transition energies in cyanines (Le Guennic

Table 3 Statistical analysis (w.r.t. experiment) obtained for cLR-PCM $E^{\text{AFCP}}(\text{neq})$ of 40 compounds. The MSE and MAE are given in eV; R_{LC} is the linear correlation coefficient. %XX is the (rounded) percentage of *exact* exchange in each functional. See Jacquemin et al. (2012a, 2014) and Moore et al. (2014) for more details

XCF	%XX	MSE	MAE	R_{LC}
B3LYP	20	−0.14	0.27	0.86
APF-D	23	−0.06	0.27	0.79
PBE0	25	−0.03	0.22	0.89
M06	27	0.05	0.23	0.89
PBE0-1/3	33	0.14	0.22	0.91
SOGGA11-X	40	0.21	0.24	0.92
M06-2X	54	0.25	0.26	0.95
CAM-B3LYP	19–65	0.24	0.25	0.94
LC-PBE	0–100	0.56	0.57	0.94
LC-PBE* ^a	0–100	0.12	0.20	0.93
ω B97X-D	22–100	0.30	0.30	0.94
LC-PBE0* ^a	25–100	0.25	0.26	0.93

^aOptimally tuned range-separated functionals

and Jacquemin 2015), the relative trends, that is, the auxochromic shifts, are restored very satisfactorily. This is illustrated in Fig. 7 where, starting from the central BODIPY core, a diverse panel of chemical modifications (adding substituents, extending the π -delocalization, stiffening the core, etc.) has been considered. As can be seen, for most variations, TD-DFT (in red) either accurately reproduces the measured (in blue) auxochromic shifts or slightly overshoots them. This figure illustrates that TD-DFT can be effectively used to screen dyes belonging to a given family (Laurent et al. 2014).

Band Topologies

Besides AFCP energies, the band topologies offer another option allowing direct comparisons between experimental and theoretical optical spectra. As stated above, it is nevertheless necessary to renormalize the measured spectra to obtain data directly comparable to the calculation (Avila Ferrer et al. 2013). In addition to the GS and (at least estimates of the) ES vibrational levels, band shapes require the determination of the vibronic couplings, that is, the couplings between the different vibrational states of the GS and the ES. This may be a very demanding task and many different approaches and programs have been designed to tackle this goal (Dierksen and Grimme 2004; Santoro et al. 2008; Avila Ferrer et al. 2011; Baiardi et al. 2013). To determine these couplings, a Taylor series of the transition moment is considered: the two first terms give rise to the so-called Franck–Condon (FC) and Herzberg–Teller (HT) contributions. As a rule of thumb, the former dominates the total response for strongly dipole allowed transitions, whereas the latter is crucial for forbidden or weakly allowed transitions. In practice, besides the ingredients necessary to compute AFCP energies, one should pay notably attention to (i) the choice of the FC, HT, or FC-HT approximation, (ii) the importance of temperature effects, and (iii) the adequacy of the harmonic approximation. Indeed, when large

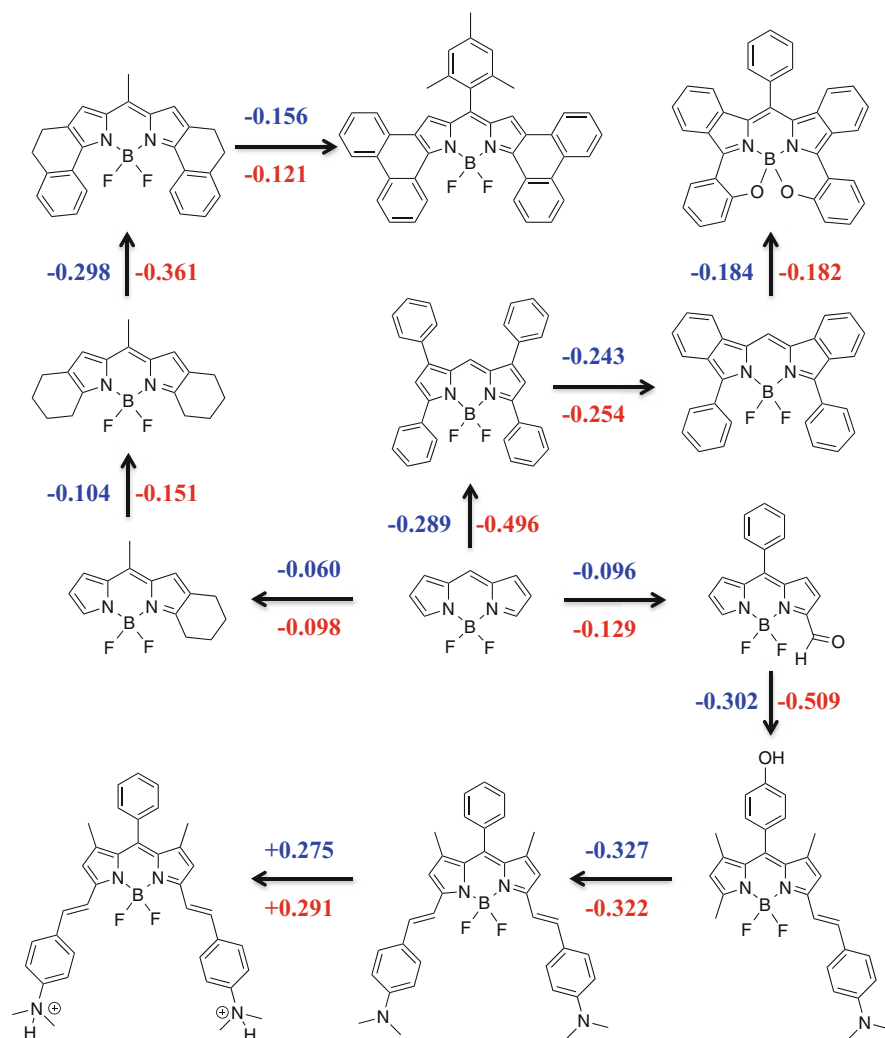


Fig. 7 Comparison between the experimental (in blue) and theoretical (in red) evolution of the AFCP energies determined for 12 BODIPY derivatives. The theoretical data have been determined at SS-PCM-TD-M06-2X/6-311+G(2d,p)//LR-PCM-TD-M06-2X/6-31G(d) level of theory. See Chibani et al. (2013) for solvent used, experimental references, and additional examples

deformations between the GS and ES are at play, anharmonic contributions can become crucial. An example of a vibronic calculation is given in Fig. 8 (Barone et al. 2014). As can be seen, several ES with significant vibronic contributions overlap to generate the experimentally observed spectrum that could be restored with a good accuracy.

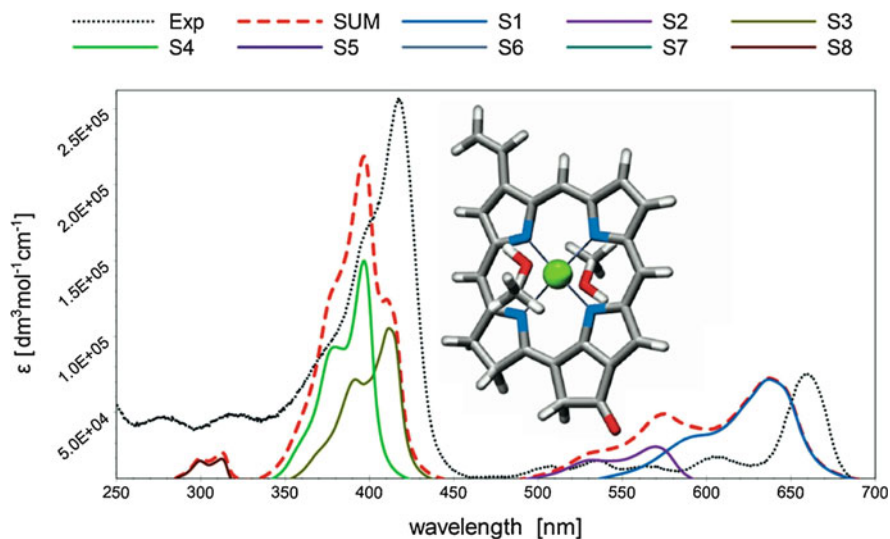


Fig. 8 Absorption spectrum of chlorophyll *a* (with two explicit methanol molecules) computed with vibronic couplings. The individual contribution of the different ES (*full lines*) and the total (*dashed lines*) are compared to experiment (Reproduced from Barone et al. (2014) with permission from Wiley)

Photophysics of Transition Metal Complexes: Effect of SOCs, Phosphorescence, and TD-DFT Performance

As discussed above the photophysical properties of systems containing metals are in most cases more intricate than purely organic systems. This is not only due to the high density of ES contributing to a single band (see Fig. 2) but also to the fact that their ES can be of very different character. Mainly five types of electronic transitions are found in transition metal (TM) complexes (González et al. 2012):

- (i) Ligand field (LF) excitations, also called d-d or metal-centered (MC) electronic excitations. In centrosymmetric environments, they are strictly not allowed by Laporte rules, but they gain intensity due to vibronic and/or spin-orbit couplings. Consequently, they generally possess lower extinction coefficients than CT transitions.
- (ii) Ligand-to-metal CT (LMCT) excitations. These are electronic excitations from occupied ligand orbitals to the partially empty d-shell of the metal.
- (iii) Metal-to-ligand CT (MLCT) excitations. These are excitations from metal-based orbitals to low-lying empty orbitals located on the ligand, typically of π^* character.

- (iv) Ligand-centered (LC) transitions. These are excitations within the ligand-based orbitals that are also present in the UV/Vis spectrum of the uncomplexed ligands.
- (v) Ligand-to-ligand CT (LLCT) states. These excitations involve ligand-based orbitals, but located in different ligands.

The characterization of the ES of TM complexes may already lead to some problematic situations, since in some cases there is a certain degree of admixture between different ES (leading, e.g., to MLCT/LC or MLCT/LLCT mixed states). Toward their characterization, as discussed above, the transformation of canonical MOs into more “compact” NTOs and the evaluation of atomic contributions in the involved molecular orbitals are useful strategies. In very problematic cases, the best option is to provide the EDDM plots. Exemplarily, in Fig. 9, are shown the different orbital representations of the first excited state of a Ru(II) complex (Jäger et al. 2015). This ES is described by two main MLCT contributions in the canonical representation. The NTOs representation allows one to clearly identify the $d_{xz} \rightarrow \pi^*$ as the main contribution to the ES. Finally, the EDDM plot provides the most compact representation.

Analogously to organic compounds, vibronic and solvent effects are also crucial for the ES of TM complexes, and they will not be further discussed here. Nevertheless, having in mind that metal complexes are often highly charged species, the comparison with experiments should always be done carefully, since solvent effects may induce very important shifts of the excitation energies. In this regard, in some cases, a combination of both explicit and implicit solvation approaches is needed to reach spectroscopic accuracy. Furthermore, and more specifically, inorganic compounds present several inherent complexities. Among them, we highlight the importance of relativistic effects (especially SOCs, see below), near degeneracies and double or multiply ES (not exclusively for these systems). Therefore, having in mind these challenges, it is necessary in many cases to play with various ab initio- and DFT-based methods in a proper relativistic framework to provide an appropriate description of the diversity of electronic structure situations and problematics. Spin-orbit effects are needed to interpret many spectroscopic phenomena, e.g., from zero-field splittings (ZFS) to intersystem crossings (ISC). To reach affordable relativistic calculations including SOCs, 2-component approaches are often used, such as the zero-order regular approximation (ZORA) or the Douglas–Kroll–Hess (DKH) Hamiltonians. These approaches can be adapted to DFT, perturbation, and wave function theories. From a technical viewpoint, SOC can be incorporated in the orbital optimization step, such as in the self-consistent SOC-TD-DFT method (Wang et al. 2005), or as a perturbation based on the scalar relativistic orbitals, such as in the quadratic response (QR) TD-DFT method (Minaev et al. 2014).

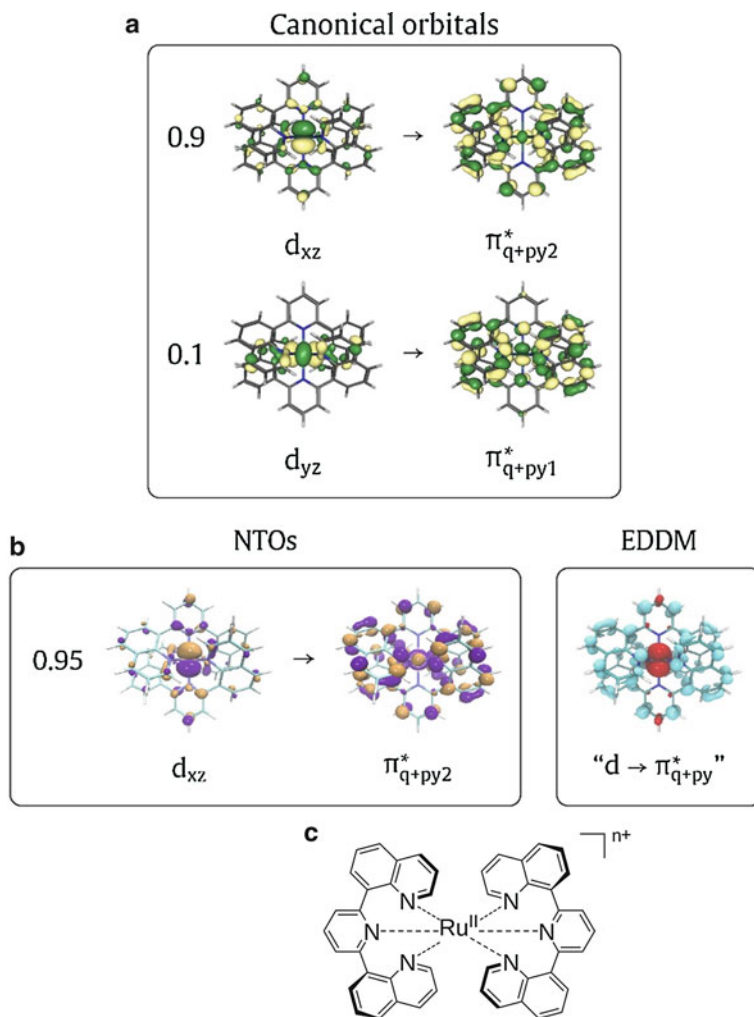


Fig. 9 Orbital representations of the character of the first excited state for the shown Ru(II) complex (Adapted from Jäger et al. (2015), Copyright (2015), with permission from Elsevier)

Influence of SOCs on the UV/Vis Properties

As in other playgrounds, TD-DFT is widely employed for TM spectroscopy, because of its simplicity and apparent black-box nature, in particular for medium/large TM complexes, which are still out of reach for more accurate ab initio methods. In many scenarios, TD-DFT can provide a reasonable agreement with the experiment, provided, on the one hand, that a judicious selection of XCF and atomic basis set is made and, on the other hand, that a proper treatment of

environmental effects is considered (see also below). As discussed above, from TD-DFT calculations, we can access to the electronic excitation and emission energies, oscillator strengths (related to the intensities), ES characters, and optimal geometries of the relevant ES. These properties are sufficient in principle to reproduce the UV/Vis absorption and emission spectra. However, for TM, due to the high density of states in the UV/Vis region and the large nuclear dimensionality of the problem, a deeper understanding of the actual photodeactivation mechanisms taking place after light irradiation is often needed to interpret the experimental evidences. This requires the assignment of (i) the photoactive excited states and (ii) the relevant nuclear deformations (which are affected by spin-orbit and vibronic interactions) along the photodeactivation pathways.

Importantly, SOCs impact the ES characters, the UV/Vis absorption and emission properties, the relaxation dynamics, as well as the radiative and non-radiative decay pathways. To illustrate the effect of SOCs in the UV/Vis absorption properties, we show in Fig. 10a the experimental UV/Vis spectra superimposed to the PCM-TD-B3LYP vertical excitations for two Ru(II) complexes (Santos et al. 2015). The long tail between 470 and 520 nm (highlighted in red in Fig. 10a), where singlet-to-singlet excitations are dark or do not exist, is attributed to singlet-to-triplet excitations. This is a common scenario for many other TM complexes. In fact, there is often an important grade of admixture between pure spin-free states *via* SOCs, which leads to mixed spin-orbit (SO) states. Not only the spectroscopic properties are modulated by the grade of admixture, but also the classical picture of pure (triplet, singlet, or any other possible multiplicity) states may become meaningless. This is illustrated in Fig. 10b, where the correlation of the calculated TD-PBE0 singlets (left) and triplets (right) spin-free states with respect to the SO-TD-PBE0 states (middle) is shown for a Re(I) complex (Baková et al. 2011). First, there are more SO states than spin-free states, since the three triplet substates split after the inclusion of the SOCs. Second, some of the SO states are energetically shifted with respect to the spin-free states. Third, some of the SO states contain significant triplet contributions.

Two main factors, that is, (i) the amplitude of the SOCs and (ii) the singlet-triplet splittings, control the final mixing at the SO states. Numerous investigations of TM complexes demonstrated that the photophysical properties are often uncorrelated to the strength of the SOC constant of the metal atom, contrary to the expectations based on the so-called heavy atom effect in organic photochemistry (Chergui 2015). As a consequence, to obtain some qualitative description of the photophysical properties, it is very often more suited to analyze the SOC contributions on the involved photoactive states, since typically $^1,^3\text{MC}$ and $^1,^3\text{MLCT}$ ES are more prone to couple *via* SOCs than $^1,^3\text{LC}$ and $^1,^3\text{LLCT}$ ES.

Phosphorescence

As discussed in the previous sections, modeling fluorescence and/or phosphorescence implies finding the optimal geometry of the relevant singlet and/or triplet

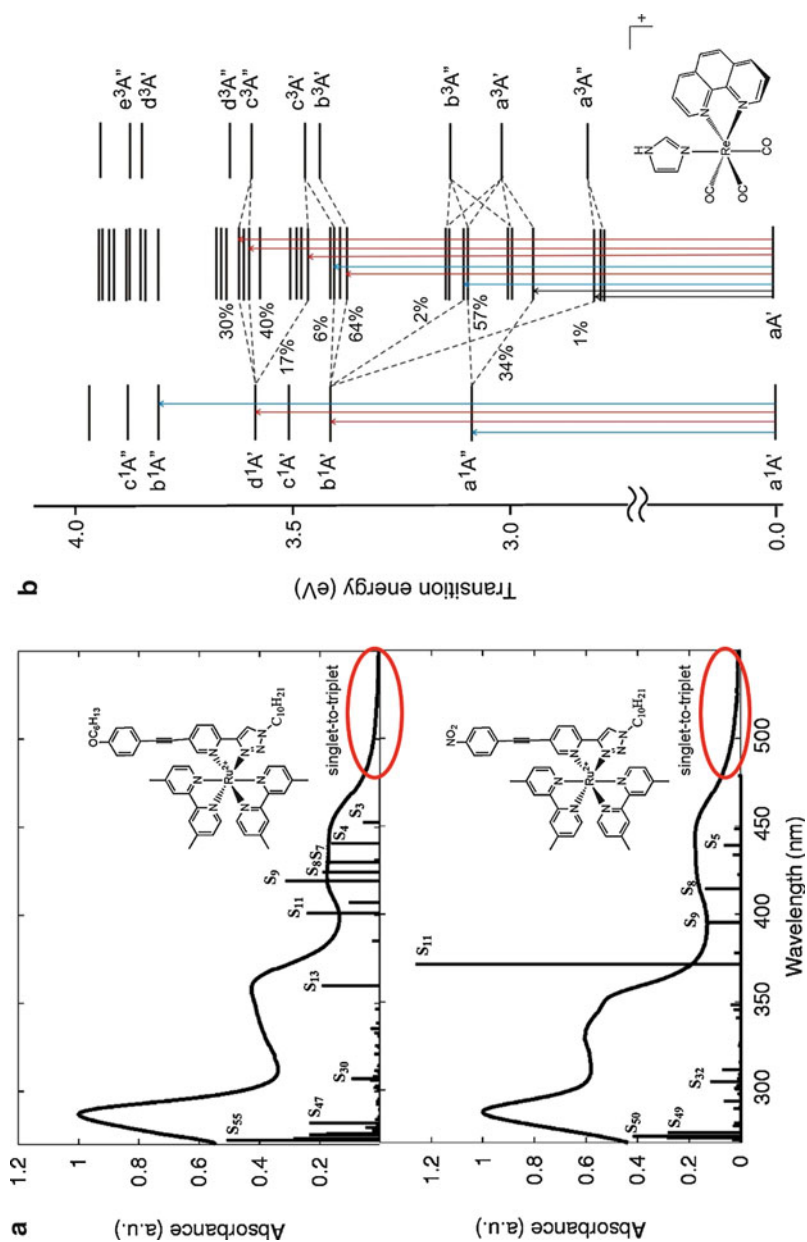


Fig. 10 (a) Experimental spectra in *solid lines* superimposed to the TD-DFT (B3LYP/6-31G(d) in CH_2Cl_2) vertical excitations for the shown Ru(II) complexes (Adapted with permissions from Happ et al. (2010). Copyright 2010, American Chemical Society). (b) Correlation between the calculated TD-PBE0 singlet (*left*) and triplet (*right*) spin-free states with the SO-TD-PBE0 states (*middle*) for the shown Re(I) complex in DMF solution (Adapted from Baková et al. (2011), Copyright (2011), with permission from Elsevier)

excited state, followed by the computation of the vertical transition energy (see Eq. 4). In TM complexes photoluminescence usually originates from the lowest triplet ES, but violations of the Kasha rule and dual emissions have been reported (Escudero and Jacquemin 2015). Recent progresses in *ab initio* and DFT-based methods in quasi-relativistic frameworks have expanded the range of applications in TM spectroscopy from a quantitative assignment of absorption and emission colors to a semiquantitative interpretation of emission spectroscopy and photochemical reactivity (Minaev et al. 2014). Today the calculation of ISC and radiative rates has become feasible. We prognose that these recent improvements will permit in the next years the estimation of more intricate photophysical properties, such as the photoluminescence quantum yields Escudero, D. (2016). *Chemical Science*, 7, 1262.

TD-DFT Performance

As described above, the results of TD-DFT are not extremely sensitive to the size of the atomic basis set when low-lying ES are considered. Therefore, a relatively compact atomic basis set, such as 6-31G(d), was often sufficient for TM complexes. However, as in other playgrounds, the selection of an adequate XCF in the TD-DFT calculations is extremely important to reach accurate results. In the following discussion, we will not mention states with a strong multi-reference character, as they cannot be handled with TD-DFT in its adiabatic approximation. Numerous investigations show that obtaining an accurate and balanced description of all the different ES is particularly challenging for inorganic systems. Consequently, the development of general rules of thumb for TM spectroscopy still remains difficult. Among the different XCF flavors, it is generally observed that hybrid functionals with intermediate amounts of exact exchange, such as PBE0 (25%) or B3LYP (20%), outperform the rest of functionals and provide a well-balanced description of several types of singlet and triplet excitations, especially for overall positively charged Ru(II), Os(II), Pt(II), Ir(III), Re(I), and Au(I) complexes (to mention some of the most widely spread TM and organometallic complexes) (González et al. 2012; Zálaiš 2007; Niehaus et al. 2015; Daniel 2015). In contrast, pure functionals, e.g., BP86 and PBE, are more suited for negatively charged TM complexes (e.g., bearing Fe(II) and Cr metal centers). Hence, despite the systematic overestimation of the MC states by pure XCF and their inadequacy for organic compounds, they outperform the rest of TD-DFT flavors for these systems, especially in the description of CT excitations involving the metal (Escudero and Thiel 2014). More surprising is the fact that RSH XCF, such as CAM-B3LYP, do not provide an accurate description of these CT states, regardless of the nature of the TM complex. This is illustrated in Fig. 11, where we show the experimental UV/Vis spectrum of a Ru(II) complex superimposed to the convoluted PCM-TD-B3LYP and PCM-TD-CAM-B3LYP predictions (Santos et al. 2015). Both the intense absorption band peaking at ca. 260 nm, which is mainly due to LC excitations and the low-energy band centered at ca. 460 nm, which results from MLCT excitations, are clearly better described at the PCM-TD-B3LYP level of theory. Indeed, for these positively

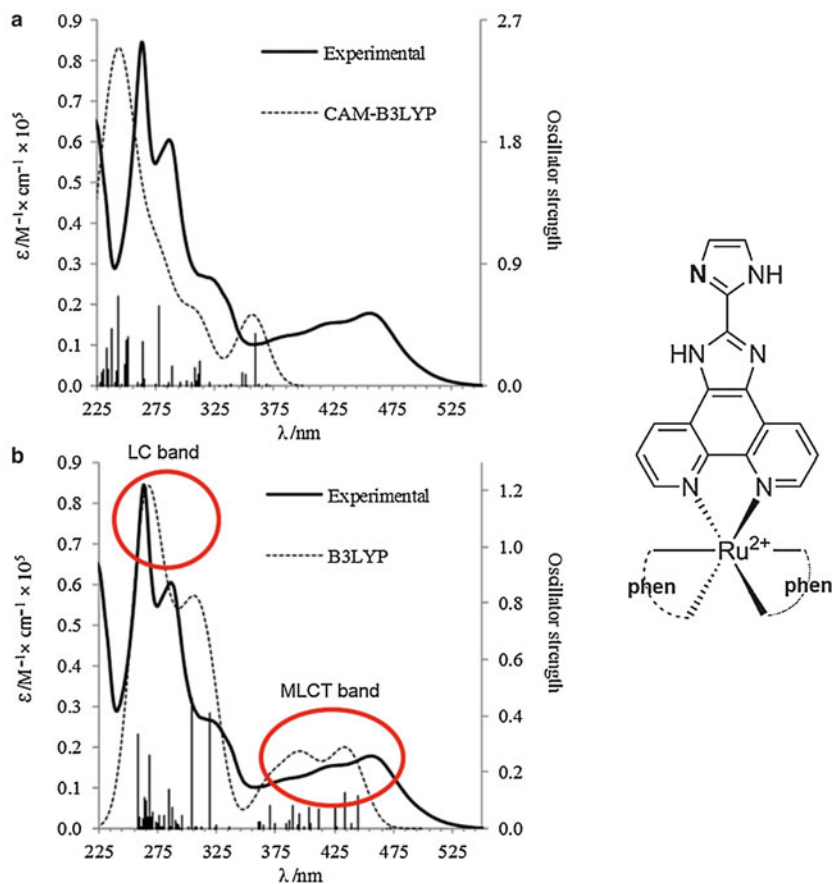


Fig. 11 UV/Vis absorption spectrum for the shown complex in methanol (*solid line*) and computed PCM-TD-DFT vertical excitation energies using the CAM-B3LYP (**a**) and B3LYP (**b**) functionals the *dashed line* represents a Gaussian convolution of the vertical excitation energies. Phen stands for phenanthroline (Adapted from Santos et al. (2015) with permission from Wiley)

charged TM complexes, PCM-TD-CAM-B3LYP substantially overestimates the excitation energies of both local and CT excitations. Therefore, these RSH XCF are generally not ideal for TM complexes. Finally, as the inclusion of exact exchange exacerbates the triplet instability problems, it is often recommended to use the Tamm–Dancoff (TDA) approximation in TD-DFT to determine triplet excitation energies, though extensive benchmarks on TM complexes have yet to appear.

In a nutshell, the assignment of the UV/Vis spectrum of very well-behaved systems is nowadays possible with TD-DFT. In some problematic cases, the assignment of the experimental data exclusively based on TD-DFT can become dangerous without a careful validation. Indeed, a simple matching of theoretical and experimental bands might be misleading, because there may be fortuitous

matches due to the large number of low-lying transitions, the inherent inaccuracies of TD-DFT calculations, the possible error cancelations, and the interplay with other effects (solvent effects, spin-orbit, and vibronic couplings).

Complex Environments and Biomolecules

In this last section we aim to show how to deal with the photophysical properties of large biomolecules rather than “isolated” organic and inorganic molecules. Indeed, during the last decades, huge experimental and theoretical efforts have been made to mimic and improve natural processes allowing to capture, transfer, and store solar energy. Such photobiological reactions are possible only if the organisms contain one or more chromogens. These light-sensitive compounds are not only essential for light harvesting plants that produce energy but also play a central role in new technological applications (artificial molecular motors, biosensors, etc.). For such photobiological systems, it is usual to combine quantum mechanics (QM) to a lower level of theory, very often molecular mechanics (MM) or fragment-based methods such as effective fragment potential (EFP), to model both the chromophore and its full environment. A detailed description of such hybrid approaches is beyond our scope and we redirect the interested readers to Senn and Thiel (2009) and Gordon et al. (2012).

Setting Up the Calculation

In practice, the site responsible for the photophysical properties is treated with QM, typically with TD-DFT, though highly correlated wave function method has also been used, while the remaining of the biomolecules (e.g., protein) are described thanks to a force field.

Initial Set of Biomolecule Coordinates

In most cases the structures (and hence the coordinates) of the biomolecule of interest have been published in the Protein Data Base (PDB). These coordinates can be directly used and in first approximation, considered as an averaged structure of the dynamic of the system. However, caution is necessary, especially when two or more conformations of the systems coexist. It is therefore valuable to perform molecular dynamic (MD) simulation of the systems before running QM/MM computations. Indeed the MD simulations allow to retrieve several possible conformations that might play an important role in the final optical signatures, especially when significant geometrical changes are made close to the chromophore. After such MD simulations, ES energies can be computed either: (i) on an averaged structure meaning that the information on the dynamics is lost or (ii) on several snapshots extracted from the MD simulation, if one aims to get more accurate insights and/or define an uncertainty on the theoretical values. For the evaluation of ES energies, the convergence is usually obtained rapidly, e.g., single-point computations on

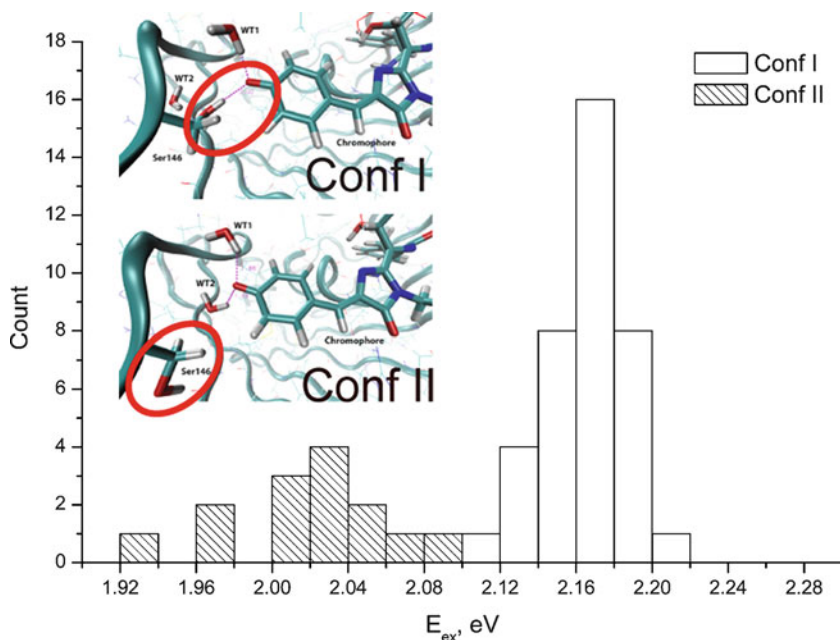


Fig. 12 Histograms showing the distribution of the excitation energies depending of the protein conformation (one bin is equal to 0.02 eV) (Adapted with permission from Laurent et al. (2012) from Wiley. Copyright 2012 American Chemical Society)

ca. 20 snapshots typically yield a standard deviation of ca. 0.04 eV – ionization energies and electronic affinities require much more calculations. This is illustrated in Fig. 12, where the numerous computed ES energies are represented through histograms for two conformations of the mStrawberry fluorescent protein. In this system both polar (serine) and a charged (lysine) residues are playing a pivotal role on the absorption spectra. We underline that the experimental spectrum is only reproduced when weighting the excitation energies by the population of each conformation (Laurent et al. 2012).

Another choice to obtain initial coordinates is to use QM/MM MD, e.g., Car–Parrinello or Born–Oppenheimer dynamics. Though these approaches have become increasingly popular, they remain rather limited due to: (i) the accessible simulation time (maximum 1 ns) and (ii) the plane-wave description of the systems that facilitates the use periodic condition but renders the applications of TD-DFT with hybrid XCF very challenging. When no precise structure of the system is available, homology modeling followed by MD simulations might also be way to obtain consistent coordinates.

Determination of the Size of the QM Region

In most of QM/MM computations, a preliminary study needs to be performed to determine the adequate size of the QM part where the TD-DFT calculations will be

performed. The sensitivity of QM/MM results to the extension of the QM part has been assessed for several properties including UV/visible signatures (Altun et al. 2009; Sanchez-Garcia et al. 2010). In photobiological systems, the chromophore was indeed historically the sole component included in the QM region. However, recently it has been highlighted that the residues surrounding the (photo)active site should also be included within the QM region (Hu et al. 2011; Liao and Thiel 2012). The number of residues rapidly becomes very large (>200 atoms), but typically the treatment of such enlarged QM part is manageable with TD-DFT. Despite these works, no general rule might be given regarding the size of the QM region, as this size is highly dependent upon the details of the considered systems (solvent exposed, polar binding pocket, van de Waals interactions, salt bridge, etc.). Comparisons of the structural and optical properties with experimental data remain generally necessary.

In addition to the size of the QM part, an additional difficulty appears in QM/MM computations in which the QM part is chemically linked to the MM part. Dealing with a covalent bond between the QM and the MM parts is always delicate, but is rather common as such bond is present in the most widely studied biological systems, i.e., fluorescent proteins, retinal, and rhodopsins. The three major boundary schemes are: the very popular link atom (LA) scheme, the boundary atom approach, and the localized-orbitals method. Once more, it is difficult to give general advices, but one should be aware that the selection of boundary model might affect the simulated spectral signatures.

Importance of the Embedding

Within TD-DFT/MM approaches, the density of the QM part should be influenced by the MM surrounding. One generally splits this total interaction into three components: mechanical (ME), electrostatic (EE), and polarization (PE) embeddings, which should ideally be all simultaneously accounted for. Briefly, in the ME only structural effects induced by the environments are taken into account, the QM density remaining non-polarized by the electrostatic MM charges. In contrast, in the EE scheme, this polarization effect is accounted for. When both the QM and the MM parts are polarizing each other, the PE is also taken into account. In the computation of optical properties with TD-DFT, PE has been often neglected because: (i) consistent polarizable force fields are not widely available and (ii) the associate computational cost is significant. However, polarizable force fields and the QM-based EFP methods are now well developed (Gordon et al. 2001; Ghosh et al. 2013; Kaliman and Slipchenko 2013). It is therefore not surprising that several approaches have been applied in the recent years to account for PE (Jacquemin et al. 2009b; Labat et al. 2009; Laurent and Assfeld 2010; Neugebauer et al. 2010; Biancardi et al. 2013; Presti et al. 2014). One of these schemes is the so-called ERS (electronic response of the surroundings) (Laurent and Assfeld 2010), which takes into account for the PE in an implicit way by making a usual QM/MM computation combined to a continuum surrounding (with the adequate dielectric constant) for the QM part. For instance, for a squaraine/tetralactam complex, the experimental bathochromic shift due to the caging effect of the squaraine cannot be reproduced by the ME+EE

scheme (which predicts a hypsochromic displacement) but is correctly captured by TD-DFT/MM when PE is accounted for (Jacquemin et al. (2009b)). Likewise, a complete study has been performed on rhodopsin and its mutants comparing several force fields as well as the LoProp approaches in an important contribution (Söderhjelm et al. 2009), which cannot be briefly summarized. In the LoProp scheme, the electrostatic contribution is described not only with point charges (as usual for classical force field) but also through dipole and quadrupole contributions. In the same time, the PE is computed thanks to anisotropic polarizabilities.

Toward Optical Properties

Most of the TD-DFT/MM approaches are currently employed to simulate the absorption spectra of chromophore embedded into biomolecules within the vertical approximation. Indeed the computations of the excited state for the whole system remain challenging as the force fields have been mainly developed to characterize ground state properties. However, by assuming that the surrounding relaxation in the ES is rather weak, one can account for the chromophore relaxation only and hence obtain first insights into the emission properties of biomolecules. Further developments in this direction are necessary to reach the complete ES properties for a protein.

The usual approach to analyze the QM/MM computations is to apply a bottom-up strategy going from the description of the isolated or solvated chromophore using methods described in the previous sections and next considering the chromophore in its real environment. As clear from the above statements, the intrinsic features of QM/MM computations allow one to decompose the characteristic λ_{\max} into four contributions compared to the isolated structure ($\lambda_{\max}^{\text{isolated}}$):

$$\lambda_{\max} = \lambda_{\max}^{\text{isolated}} + \Delta\lambda_{\max}^{\text{ME}} + \Delta\lambda_{\max}^{\text{EE}} + \Delta\lambda_{\max}^{\text{PE}}, \quad (12)$$

where $\lambda_{\max}^{\text{ME}}$, $\lambda_{\max}^{\text{EE}}$, $\lambda_{\max}^{\text{PE}}$ account, respectively, for the geometrical modifications due to the presence of the biomolecule, for the polarization of the QM density by the force field mono- and multipoles, and for the mutual polarization between the QM and the MM parts. None of these contributions can be neglected as they typically present similar magnitudes.

Conclusions

Time-dependent density functional theory now stands as a mature tool to interpret and predict the optical spectra of organic, inorganic, and biological molecules. One can indeed determine a large number of properties (absorption, fluorescence, phosphorescence, dipoles, etc.) with a large variety of flavors of TD-DFT (different families of functionals adequate for specific excited states, accounting or not for vibronic and spin-orbit couplings, etc.) in a wide range of environments (gas phase,

solution, embedding in proteins, etc.). For the simplest cases, TD-DFT can even be applied as an almost black-box approach and nevertheless yield important insights into the excited states (geometrical relaxations, nature of the states, importance of the vibronic couplings, possibility to have intersystem-crossing, etc.). This is a nice and efficient tool; use it to model your excited-state properties!

However, TD-DFT in its adiabatic form suffers from several flaws and the theoretical community is very active in trying to circumvent them. Although this is beyond the scope of the present contribution, it is certainly worth to mention that schemes have been developed to model conical intersections (Huix-Rotllant et al. 2015) as well as to go beyond the adiabatic approximation and treat more consistently and more accurately all families of excited states (Casida and Huix-Rotllant 2012; Ziegler et al. 2015).

Acknowledgements The authors are indebted to many colleagues and collaborators for exciting discussions and joint works in the field. D.E. and D.J. acknowledges the European Research Council (ERC) for financial support in the framework of a Starting Grant (Marches – 278845).

Bibliography

- Adamo, C., & Jacquemin, D. (2013). The calculations of excited-state properties with Time-Dependent Density Functional Theory. *Chemical Society Reviews*, 42, 845.
- Altun, A., Yokoyama, S., & Morokuma, K. (2009). Mechanism of Spectral Tuning Going from Retinal in Vacuo to Bovine Rhodopsin and its Mutants: Multireference ab Initio Quantum Mechanics/Molecular Mechanics Studies. *Journal of Physical Chemistry B*, 112, 16883.
- Avila Ferrer, F. J., Improta, R., Santoro, F., & Barone, V. (2011). Computing the inhomogeneous broadening of electronic transitions in solution: a first-principle quantum mechanical approach. *Physical Chemistry Chemical Physics*, 13(38), 17007.
- Avila Ferrer, F. J., Cerezo, J., Stendardo, E., Improta, R., & Santoro, F. (2013). Insights for an Accurate Comparison of Computational Data to Experimental Absorption and Emission Spectra: Beyond the Vertical Transition Approximation. *Journal of Chemical Theory and Computation*, 9, 2072.
- Baer, R., Livshits, E., & Salzner, U. (2010). Tuned range-separated hybrids in density functional theory. *Annual Review of Physical Chemistry*, 61, 85.
- Baerends, E. J., Gritsenko, O. V., & van Meer, R. (2013). The Kohn–Sham gap, the fundamental gap and the optical gap: the physical meaning of occupied and virtual Kohn–Sham orbital energies. *Physical Chemistry Chemical Physics*, 15, 16408.
- Baiardi, A., Bloino, J., & Barone, V. (2013). General Time Dependent Approach to Vibronic Spectroscopy Including Franck–Condon, Herzberg–Teller, and Duschinsky Effects. *Journal of Chemical Theory and Computation*, 9(9), 4097.
- Baková, R., Chergui, M., Daniel, C., Vlček, A., Jr., & Zálaiš, S. (2011). Relativistic effects in spectroscopy and photophysics of heavy-metal complexes illustrated by spin–orbit calculations of [Re(imidazole)(CO)₃(phen)]⁺. *Coordination Chemistry Reviews*, 255(7–8), 975. A Celebration of Harry B. Gray’s 75th Birthday.
- Barbatti, M., & Crespo-Otero, R. (2015). Surface Hopping Dynamics with DFT Excited States. In N. Ferré, M. Filatov, & M. Huix-Rotllant (Eds.), *Density-functional methods for excited states* (Topics in current chemistry, pp. 1–30). Berlin/Heidelberg: Springer.
- Barone, V., & Biczysko, M., Borkowska-Panek, M., & Bloino, J. (2014). A multifrequency virtual spectrometer for complex bio-organic systems: vibronic and environmental effects on the UV/Vis spectrum of chlorophyll a. *ChemPhysChem*, 15(15), 3355.

- Biancardi, A., Biver, T., Secco, F., & Mennucci, B. (2013). An investigation of the photophysical properties of minor groove bound and intercalated DAPI through quantum-mechanical and spectroscopic tools, *Physical Chemistry Chemical Physics*, *15*(13), 4596.
- Brémond, E. A., Kieffer, J., & Adamo, C. (2010). A Reliable Method for Fitting TD-DFT Transitions to Experimental UV-Visible Spectra. *Journal of Molecular Structure: THEOCHEM*, *954*(1–3), 52.
- Cammi, R., & Mennucci, B. (1999). Linear response theory for the polarizable continuum model. *Journal of Chemical Physics*, *110*, 9877.
- Caricato, M., Mennucci, B., Tomasi, J., Ingrosso, F., Cammi, R., Corni, S., & Scalmani, G. (2006). Formation and relaxation of excited states in solution: A new time dependent polarizable continuum model based on time dependent density functional theory. *Journal of Chemical Physics*, *124*, 124520.
- Caricato, M., Trucks, G. W., Frisch, M. J., & Wiberg, K. B. (2010). Electronic Transition Energies: A Study of the Performance of a Large Range of Single Reference Density Functional and Wave Function Methods on Valence and Rydberg States Compared to Experiment. *Journal of Chemical Theory and Computation*, *6*, 370.
- Casida, M. E. (1995). *Time-dependent density-functional response theory for molecules* (Recent advances in density functional methods, Vol. 1, pp. 155–192). Singapore: World Scientific.
- Casida, M. E., & Huix-Rotllant, M. (2012). Progress in Time-Dependent Density-Functional Theory. *Annual Review of Physical Chemistry*, *63*, 287.
- Chergui, M. (2015). Ultrafast Photophysics of Transition Metal Complexes. *Accounts of Chemical Research*, *48*(3), 801. PMID: 25646968
- Chibani, S., Le Guennic, B., Charaf-Eddin, A., Laurent, A. D., & Jacquemin, D. (2013). Revisiting the optical signatures of BODIPY with ab initio tools. *Chemical Sciences*, *4*, 1950.
- Ciofini, I., & Adamo, C. (2007). Accurate Evaluation of Valence and Low-Lying Rydberg States with Standard Time-Dependent Density Functional Theory. *Journal of Physical Chemistry A*, *111*, 5549.
- Cossi, M., & Barone, V. (2001). Time-dependent density functional theory for molecules in liquid solutions. *Journal of Chemical Physics*, *115*, 4708.
- Daniel, C. (2015). Photochemistry and photophysics of transition metal complexes: Quantum chemistry. *Coordination Chemistry Reviews*, *282–283*, 19. Proceedings from the 20th International Symposium on the Photophysics and Photochemistry of Coordination Compounds.
- Dierksen, M., & Grimme, S. (2004). The Vibronic Structure of Electronic Absorption Spectra of Large Molecules: A Time-Dependent Density Functional Study on the Influence of “Exact” Hartree–Fock Exchange. *Journal of Physical Chemistry A*, *108*, 10225.
- Escudero, D., & Jacquemin, D. (2015). Computational insights into the photodeactivation dynamics of phosphors for OLEDs: a perspective. *Dalton Transactions*, *44*, 8346.
- Escudero, D., & Thiel, W. (2014). Assessing the density functional theory-based multireference configuration interaction (DFT/MRCI) method for transition metal complexes. *The Journal of Chemical Physics*, *140*(19), 194105.
- Furche, F., & Ahlrichs, R. (2002). Adiabatic time-dependent density functional methods for excited state properties. *Journal of Chemical Physics*, *117*, 7433.
- Ghosh, D., Kosenkov, D., Vanovschi, V., Flick, J., Kaliman, I., Shao, Y., Gilbert, A. T. B., Krylov, A. I., & Slipchenko, L. V. (2013). Effective fragment potential method in Q-CHEM: A guide for users and developers. *Journal of Computational Chemistry*, *34*, 1060.
- Goerigk, L., & Grimme, S. (2010). Assessment of TD-DFT methods and of various spin scaled CIS(D) and CC2 versions for the treatment of low-lying valence excitations of large organic dyes. *Journal of Chemical Physics*, *132*, 184103.
- González, L., Escudero, D., & Serrano-Andrés, L. (2012). Progress and Challenges in the Calculation of Electronic Excited States. *ChemPhysChem*, *13*(1), 28.
- Gordon, M. S., Freitag, M., Bandyopadhyay, P., Jensen, J., Kairys, V., & Stevens, W. J. (2001). The Effective Fragment Potential Method: A QM-Based MM Approach to Modeling Environmental Effects in Chemistry. *Journal of Physical Chemistry A*, *105*, 293.

- Gordon, M. S., Fedorov, D. G., Pruitt, S. R., & Slipchenko, L. V. (2012). Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chemistry Reviews*, *112*, 632.
- Guido, C. A., Jacquemin, D., Adamo, C., & Mennucci, B. (2010). On the TD-DFT Accuracy in Determining Single and Double Bonds in Excited-State Structures of Organic Molecules. *Journal of Physical Chemistry A*, *114*, 13402.
- Happ, B., Escudero, D., Hager, M. D., Friebe, C., Winter, A., Gorls, H., Altuntas, E., Gonzalez, L., & Schubert, U. S. (2010). N-Heterocyclic Donor- and Acceptor-Type Ligands Based on 2-(1H-[1,2,3]Triazol-4-yl)pyridines and Their Ruthenium(II) Complexes. *Journal of Organic Chemistry*, *75*, 4025–4038.
- Hu, L. H., Söderhjelm, P., & Ryde, U. (2011). On the Convergence of QM/MM Energies. *Journal of Chemical Theory and Computation*, *7*(3), 761.
- Huix-Rotllant, M., Nikiforov, A., Thiel, W., & Filatov, M. (2015). Description of conical intersections with density functional methods. In *Density-functional methods for excited states* (Topics in current chemistry, vol 368, pp. 445–476). Berlin/Heidelberg: Springer.
- Improta, R., Scalmani, G., Frisch, M. J., & Barone, V. (2007). Toward effective and reliable fluorescence energies in solution by a new state specific polarizable continuum model time dependent density functional theory approach. *Journal of Chemical Physics*, *127*, 074504.
- Isegawa, M., Peverati, R., & Truhlar, D. G. (2012). Performance of recent and high-performance approximate density functionals for time-dependent density functional theory calculations of valence and Rydberg electronic transition energies. *Journal of Chemical Physics*, *137*, 244104.
- Jacquemin, D., Preat, J., Wathelet, V., & Perpète, E. A. (2006). Substitution and chemical environment effects on the absorption spectrum of indigo. *Journal of Chemical Physics*, *124*, 074104.
- Jacquemin, D., Wathelet, V., Perpète, E. A., & Adamo, C. (2009a). Extensive TD-DFT Benchmark: Singlet-Excited States of Organic Molecules. *Journal of Chemical Theory and Computation*, *5*, 2420.
- Jacquemin, D., Perpète, E. A., Laurent, A. D., Assfeld, X., & Adamo, C. (2009b). Spectral properties of self-assembled squaraine–tetralactam: a theoretical assessment. *Physical Chemistry Chemical Physics*, *11*, 1258.
- Jacquemin, D., Planchat, A., Adamo, C., & Mennucci, B. (2012a). TD-DFT Assessment of Functionals for Optical 0–0 Transitions in Solvated Dyes. *Journal of Chemical Theory and Computation*, *8*, 2359.
- Jacquemin, D., Le Bahers, T., Adamo, C., & Ciofini, I. (2012b). What is the “best” atomic charge model to describe through-space charge-transfer excitations? *Physical Chemistry Chemical Physics*, *14*, 5383. Code available at Université de Nantes, <http://www.sciences.univ-nantes.fr/CEISAM/erc/marches/> (Accessed 1 May 2014).
- Jacquemin, D., Moore, B., Planchat, A., Adamo, C., & Autschbach, J. (2014). Performance of an Optimally Tuned Range-Separated Hybrid Functional for 0–0 Electronic Excitation Energies. *Journal of Chemical Theory and Computation*, *10*(4), 1677.
- Jäger, M., Freitag, L., & González, L. (2015). Using computational chemistry to design Ru photosensitizers with directional charge transfer. *Coordination Chemistry Reviews*, *304–305*, 146. cOST: European Cooperation in Science and Technology Current Challenges in Supramolecular Artificial Photosynthesis.
- Kaliman, I. A., & Slipchenko, L. V. (2013). A new parallel implementation of the effective fragment potential method as a portable software library. *Journal of Computational Chemistry*, *34*, 2284.
- Labat, F., Ciofini, I., Hratchian, H. P., Frisch, M. J., Raghavachari, K., & Adamo, C. (2009). First Principles Modeling of Eosin-Loaded ZnO Films: A Step toward the Understanding of Dye-Sensitized Solar Cell Performances. *Journal of the American Chemical Society*, *131*(40), 14290.
- Latouche, C., Skouteris, D., Palazzetti, F., & Barone, V. (2015). TD-DFT Benchmark on Inorganic Pt(II) and Ir(III) Complexes. *Journal of Chemical Theory and Computation*, *11*(7), 3281–3289.

- Laurent, A., & Assfeld, X. (2010). Effect of the enhanced cyan fluorescent protein framework on the UV/visible absorption spectra of some chromophores. *The Journal Interdisciplinary Sciences—Computational Life Sciences*, 2(1), 38.
- Laurent, A. D., & Jacquemin, D. (2013). TD-DFT benchmarks: A review. *International Journal of Quantum Chemistry*, 113, 2019.
- Laurent, A. D., Mironov, V. A., Chapagain, P. P., Nemukhin, A. V., & Krylov, A. I. (2012). Conformational Exploration of Two Peptides and Their Hybrid Polymer Conjugates: Potentialities As Self-Aggregating Materials. *Journal of Physical Chemistry B*, 116, 12426.
- Laurent, A. D., Adamo, C., & Jacquemin, D. (2014). Dye chemistry with time-dependent density functional theory. *Physical Chemistry Chemical Physics*, 16(28), 14334.
- Le Bahers, T., Adamo, C., & Ciofini, I. (2011). A qualitative index of spatial extent in charge-transfer excitations. *Journal of Chemical Theory and Computation*, 7, 2498. Code available at Chimie Paristech, www.chimie-paristech.fr/labos/LECA/Research/site_msc/
- Le Guennic, B., & Jacquemin, D. (2015). Taking Up the Cyanine Challenge with Quantum Tools. *Accounts of Chemical Research*, 48, 530.
- Liao, R. Z., & Thiel, W. (2012). Comparison of QM-Only and QM/MM Models for the Mechanism of Tungsten-Dependent Acetylene Hydratase. *Journal of Chemical Theory and Computation*, 8(10), 3793.
- Liu, J., & Liang, W. Z. (2011). Analytical approach for the excited-state Hessian in time-dependent density functional theory: Formalism, implementation, and performance. *Journal of Chemical Physics*, 135(18), 184111.
- Marenich, A. V., Cramer, C. J., Truhlar, D. G., Guido, C. G., Mennucci, B., Scalmani, G., & Frisch, M. J. (2011). Practical computation of electronic excitation in solution: vertical excitation model. *Chemical Sciences*, 2, 2143.
- Marques, M. A. L., Nogueira, F. M. S., Gross, E. K. U., & Rubio, A. (Eds.). (2012). *Fundamentals of time-dependent density functional theory* (Lecture notes in physics, Vol. 837). Heidelberg: Springer.
- Martin, R. L. (2003). Natural transition orbitals. *Journal of Chemical Physics*, 118(11), 4775.
- Minaev, B., Baryshnikov, G., & Agren, H. (2014). Principles of phosphorescent organic light emitting devices. *Physical Chemistry Chemical Physics*, 16, 1719.
- Moore, B., Charaf-Eddin, A., Planchat, A., Adamo, C., Autschbach, J., & Jacquemin, D. (2014). Electronic Band Shapes Calculated with Optimally Tuned Range-Separated Hybrid Functionals. *Journal of Chemical Theory and Computation*, 10(10), 4599.
- Neugebauer, J., Curutchet, C., Munoz-Losa, A., & Mennucci, B. (2010). A Subsystem TDDFT Approach for Solvent Screening Effects on Excitation Energy Transfer Couplings. *Journal of Chemical Theory and Computation*, 6(6), 1843.
- Niehaus, T. A., Hofbeck, T., & Yersin, H. (2015). Charge-transfer excited states in phosphorescent organo-transition metal compounds: a difficult case for time dependent density functional theory? *RSC Advanced*, 5, 63318.
- Peach, M. J. G., Benfield, P., Helgaker, T., & Tozer, D. J. (2008). Excitation energies in density functional theory: An evaluation and a diagnostic test. *Journal of Chemical Physics*, 128, 044118.
- Presti, D., Labat, F., Pedone, A., Frisch, M. J., Hratchian, H. P., Ciofini, I., Menziani, M. C., & Adamo, C. (2014). Computational Protocol for Modeling Thermochromic Molecular Crystals: Salicylidene Aniline As a Case Study. *Journal of Chemical Theory and Computation*, 10, 5577.
- Runge, E., & Gross, E. K. U. (1984). Density-Functional Theory for Time-Dependent Systems. *Physical Review Letters*, 52, 997.
- Sanchez-Garcia, E., Doerr, M., & Thiel, W. (2010). QM/MM study of the absorption spectra of DsRed.M1 chromophores. *Journal of Computational Chemistry*, 31(8), 1603.
- Santorio, F., Lami, A., Improta, R., Bloino, J., & Barone, V. (2008). Effective method for the computation of optical spectra of large molecules at finite temperature including the Duschinsky and Herzberg–Teller effect: The QxQx band of porphyrin as a case study. *Journal of Chemical Physics*, 128, 224311.

- Santos, A. R., Escudero, D., González, L., & Orellana, G. (2015). Unravelling the Quenching Mechanisms of a Luminescent RuII Probe for CuII. *Chemistry – An Asian Journal*, 10(3), 622.
- Scalmani, G., Frisch, M. J., Mennucci, B., Tomasi, J., Cammi, R., & Barone, V. (2006). Geometries and properties of excited states in the gas phase and in solution: Theory and application of a time-dependent density functional theory polarizable continuum model. *Journal of Chemical Physics*, 124, 094107.
- Send, R., Kühn, M., & Furche, F. (2011). Assessing Excited State Methods by Adiabatic Excitation Energies. *Journal of Chemical Theory and Computation*, 7(8), 2376.
- Senn, H. M., & Thiel, W. (2009). QM/MM Methods for Biomolecular Systems. *Angewandte Chemie International Edition*, 48(7), 1198.
- Söderhjelm, P., Husberg, C., Strambi, A., Olivucci, M., & Ryde, U. (2009). Protein Influence on Electronic Spectra Modeled by Multipoles and Polarizabilities. *Journal of Chemical Theory and Computation*, 5, 649.
- Srebro, M., & Autschbach, J. (2012). Does a Molecule-Specific Density Functional Give an Accurate Electron Density? The Challenging Case of the CuCl Electric Field Gradient. *Journal of Physical Chemistry Letters*, 3, 576.
- Tomasi, J., Mennucci, B., & Cammi, R. (2005). Quantum Mechanical Continuum Solvation Models. *Chemistry Reviews*, 105, 2999.
- Ullrich, C. (2012). *Time-dependent density-functional theory: Concepts and applications* (Oxford graduate texts). New York: Oxford University Press.
- Vlček, A., Jr., & Zálíš, S. (2007). Modeling of charge-transfer transitions and excited states in d6 transition metal complexes by DFT techniques. *Coordination Chemistry Reviews*, 251(3–4), 258. A Special Issue Highlighting the Many Aspects of the Electronic Spectroscopy of Inorganic Compounds.
- Wang, F., Ziegler, T., van Lenthe, E., van Gisbergen, S., & Baerends, E. J. (2005). The calculation of excitation energies based on the relativistic two-component zeroth-order regular approximation and time-dependent density-functional with full use of symmetry. *The Journal of Chemical Physics*, 122(20), 204103.
- Wiggins, P., Gareth Williams, J. A., & Tozer, D. J. (2009). Excited state surfaces in density functional theory: A new twist on an old problem. *Journal of Chemical Physics*, 131, 091101.
- Ziegler, T., Krykunov, M., Seidu, I., & Park, Y. (2015). Constricted Variational Density Functional Theory Approach to the Description of Excited States. In N. Ferré, M. Filatov, & M. Huix-Rotllant (Eds.), *Density-functional methods for excited states* (Topics in Current Chemistry, vol 368, pp. 61–95). Berlin/Heidelberg: Springer. doi:10.1007/128_2014_611.

Norio Yoshida and Katsura Nishiyama

Contents

Introduction: Role of the Solvent in Chemistry	964
Distribution Function Theory of Molecular Liquids	965
Solvent Effects on the Electronic Structure of Solvated Molecules	971
Solvation of Biomolecules	974
Summary and Future Perspective	976
Bibliography	977

Abstract

The majority of organic and inorganic chemical reactions relevant for synthesis or industrial applications take place in the solution phase. Similarly, almost all biological reactions proceed in an aqueous environment. Theoretical descriptions of chemical reactions in solution differ in the way they treat the surrounding solvent molecules. Their large number excludes an explicit treatment. Nevertheless, the elucidation of the solvated structure plays a key role in obtaining an accurate description of chemical reactions as well as the activity of biomolecules. In this chapter, we review recent theoretical developments for the description of solvated molecules during chemical reactions, employing statistical physics to treat the solvent. The reference interaction-site model (RISM) theory, and some of its extensions, coupled with recent quantum mechanical theories, are described, and applications to various systems in solution, from conventional chemical reactions to the activities of biomolecules in biological systems, are presented. The solvent

N. Yoshida (✉)

Department of Chemistry, Graduate School of Science, Kyushu University, Fukuoka, Japan
e-mail: noriwo@chem.kyushu-univ.jp

K. Nishiyama (✉)

Faculty of Education, Shimane University, Matsue, Japan
e-mail: katsura_nishiyama@edu.shimane-u.ac.jp

distribution around the solute and solute–solvent interactions is critical for the reactions and structures of such systems. The theories described here offer the possibility to obtain detailed information on the molecular origin of solvation and may contribute to discover and elucidate novel phenomena in chemistry, physics, and bioscience.

Introduction: Role of the Solvent in Chemistry

A liquid or solution may be considered as a system composed of a large number of molecular particles. The macroscopic properties of this system can be finally traced back to the microscopic properties of the particles. Statistical mechanics plays an essential role in the emergence of macroscopic properties from microscopic ones (Hirata 2003). Statistical mechanics can also be applied to describe chemical reactions in solution, where a solute or a reaction center is located in the surrounding solvent particles, the number of which is of the order of the Avogadro constant.

How does a solvent behave during a chemical reaction in solution? In the case of conventional organic or inorganic reactions in solution, solid or liquid reactants are dissolved in the solvent, or liquid reactants are mixed. The system may be heated or cooled, if necessary. After the reaction has finished, products are separated or extracted. In this series of reactions, the solvent – including the case of liquid–liquid reactions – functions as the heat bath supplying the activation energy to the reactants. This feature is unique to liquid reactions. In the gas-phase, in contrast, reactions are mostly initiated by collisional interactions between the reacting molecules, which are essentially independent of the surrounding molecules (Levine 2005).

In addition to conventional organic or inorganic reactions, the solvent also plays a significant role in biological systems. Almost all biomolecules require water in order to develop their biological functions. Many biomolecules are not able to maintain their structure outside of the aqueous environment. Removing a biomolecule from the water where the system is functioning and transferring it into another solvent may lead to an immediate collapse of the molecular structure of the biomolecule, thus rendering it incapable to exert its usual biological function. Here, we list proteins, deoxyribonucleic acid (DNA), and sugars as examples of such biomolecules. The hydrogen bonds between these biomolecules and the water molecules, the hydrophobic or hydrophilic interactions among biomolecules, as well as those in biomolecule–cofactors are absolutely critical to express their usual functions.

Given such many-body, complicated interactions between the components, how can we describe chemical reactions in solution and biomolecules? A method on which such a treatment may be based is provided by statistical mechanics, which describes the connection between the properties of microscopic particles and the behavior of macroscopic systems. Historically, statistical mechanics has been firstly applied to the solid-state properties of metals and magnetic materials. On the other hand, theories of statistical mechanics to handle the liquid state have been developed

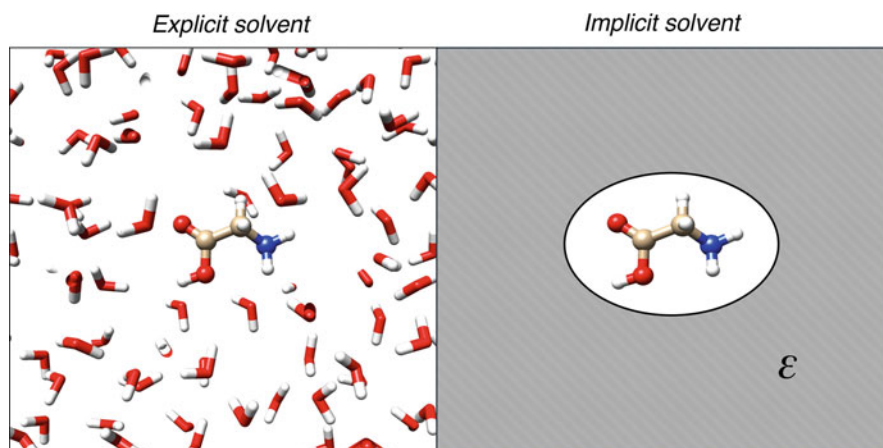


Fig. 1 Explicit (*left panel*) and implicit (*right panel*) solvent models

since the 1970s (Hansen and McDonald 2006). Another class of models treat the solvent as a dielectric continuum, where the molecular nature of the solvent is hidden (Fig. 1) (Tomasi et al. 2005). Such a continuum model is quite intuitive, and these models can be easily implemented as a computational algorithm. However, for many phenomena we need more elaborate theoretical frameworks which allow a more realistic description of interactions in these systems, such as hydrogen bond interactions.

In this chapter, we survey theories targeting liquids and molecular assemblies that have been developed based on statistical mechanics. The remainder of this chapter is organized as follows. Section “[Distribution Function Theory of Molecular Liquids](#)” covers the theory of distribution functions used to describe the structure of molecular liquids, providing also some background of the theory of statistical mechanics. In section “[Solvent Effects on the Electronic Structure of Solvated Molecules](#)” we summarize how statistical mechanics can be coupled with quantum mechanical treatments in order to describe the effect of the solvent on the electronic structures of solvated molecules. For biological systems, hydration is of critical importance, and some topics of biological functions employing recent theoretical formulations are presented in section “[Solvation of Biomolecules.](#)” Finally, section “[Summary and Future Perspective](#)” provides a summary and outlook to future developments.

Distribution Function Theory of Molecular Liquids

We begin with the theory of distribution functions for a pure solvent. Particles in the liquid state can move around inside the system and interact strongly with each other. This is a prominent feature of the liquid state, which distinguishes it from gas or solid phases. Therefore, in the liquid state, a rigorous description of the position

of particle i , as r_i , is not practical. We need to introduce statistical and probability theories to describe liquid-state structures.

The distribution probability of a molecule is given by a distribution function. The pair density distribution function is defined as:

$$\rho(\mathbf{r}, \mathbf{r}') = \langle v(\mathbf{r}) v(\mathbf{r}') \rangle. \quad (1)$$

where $\langle \bullet \rangle$ stands for an ensemble average and $v(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$ is the density field at the position \mathbf{r} inside a molecular distribution. The pair density distribution function defined above describes the probability to find a molecule at \mathbf{r}' if at the same time another molecule is present at \mathbf{r} . When the distance between \mathbf{r}' and \mathbf{r} is infinite, the probability of finding both molecules can be simply given as the product of the single molecular densities:

$$\lim_{|\mathbf{r}-\mathbf{r}'| \rightarrow \infty} \rho(\mathbf{r}, \mathbf{r}') \rightarrow \rho(\mathbf{r}) \rho(\mathbf{r}'). \quad (2)$$

In particular, in a homogeneous liquid without an external field acting on it, the molecular density is independent of \mathbf{r} , i.e., $\rho(\mathbf{r}) = \rho$. Using this property we can define a normalized distribution function by:

$$g(\mathbf{r}, \mathbf{r}') = \frac{\rho(\mathbf{r}, \mathbf{r}')}{\rho^2}, \quad (3)$$

which is called a pair distribution function. Furthermore, we define the total correlation function as:

$$h(\mathbf{r}, \mathbf{r}') = g(\mathbf{r}, \mathbf{r}') - 1. \quad (4)$$

The pair correlation function expresses the correlation of the density fluctuation between the positions \mathbf{r} and \mathbf{r}' . An equation governing these correlation functions is the Ornstein–Zernike (OZ) equation:

$$h(\mathbf{r}, \mathbf{r}') = c(\mathbf{r}, \mathbf{r}') + \rho \int c(\mathbf{r}, \mathbf{r}'') h(\mathbf{r}'', \mathbf{r}') d\mathbf{r}''. \quad (5)$$

Here, $c(\mathbf{r}, \mathbf{r}')$ is the direct correlation function, which is defined in the context of the OZ equation. Originally, Ornstein and Zernike proposed the equation rather intuitively. However, further studies have shown that the OZ equation can be derived rigorously from the theory of distribution functions (Hansen and McDonald 2006).

The OZ equation has two unknown functions, $h(\mathbf{r}, \mathbf{r}')$ and $c(\mathbf{r}, \mathbf{r}')$. To solve the OZ equation, we therefore need another relation that “closes” the set of equations. A representative closure is the hypernetted chain (HNC) approximation:

$$g(\mathbf{r}, \mathbf{r}') = \exp \left\{ -\frac{1}{k_B T} u(\mathbf{r}, \mathbf{r}') + h(\mathbf{r}, \mathbf{r}') - c(\mathbf{r}, \mathbf{r}') \right\}, \quad (6)$$

where k_B , T , and $u(\mathbf{r}, \mathbf{r}')$ are the Boltzmann constant, thermodynamic temperature, and the pair interaction potential, respectively. Two kinds of procedures are known to derive the closure: one procedure is based on Feynman diagrams; the other one employs the functional Taylor expansions.

The pair distribution function, which provides information on the structure of the liquid, can be obtained numerically. In this numerical process, we solve the combination of the OZ equation and closure by iteration, assuming additionally an appropriate pair potential between the particles.

The distribution function of a liquid consisting of polyatomic molecules is a function both of the position of the molecules, \mathbf{r} , as well as their orientation Ω . The pair density distribution function of a molecular liquid is defined as:

$$\rho(\mathbf{r}, \mathbf{r}', \Omega, \Omega') = \left\langle \sum_i \sum_{j \neq i} \delta(\mathbf{r} - \mathbf{r}_i) \delta(\mathbf{r}' - \mathbf{r}_j) \delta(\Omega - \Omega_i) \delta(\Omega' - \Omega_j) \right\rangle. \quad (7)$$

The pair distribution function is then derived as:

$$g(\mathbf{r}_{12}, \Omega_1, \Omega_2) = \left(\frac{\Omega}{\rho} \right)^2 \rho(\mathbf{r}_{12}, \Omega_1, \Omega_2), \quad (8)$$

where we have used the relations $\mathbf{r}_{12} = \mathbf{r}_2 - \mathbf{r}_1$ and $\Omega \equiv \int d\Omega$. The molecular OZ (MOZ) equation is then given by:

$$h(\mathbf{r}_{12}, \Omega_1, \Omega_2) = c(\mathbf{r}_{12}, \Omega_1, \Omega_2) + \frac{\rho}{\Omega} \int c(\mathbf{r}_{13}, \Omega_1, \Omega_3) h(\mathbf{r}_{32}, \Omega_3, \Omega_2) d\mathbf{r}_3 d\Omega_3. \quad (9)$$

A method to solve the MOZ equation is to use expansions in spherical harmonics. In particular, the spherical harmonic expansion of the pair correlation function is written as:

$$h(\mathbf{r}_{12}, \Omega_1, \Omega_2) = \sum_{\substack{mnl \\ \mu\nu}} h_{\mu\nu}^{mnl}(r_{12}) \sum_{\mu'\nu'\lambda'} \begin{pmatrix} m & n & l \\ \mu' & \nu' & \lambda' \end{pmatrix} R_{\mu\mu'}^m(\Omega_1) R_{\nu\nu'}^n(\Omega_2) R_{0\lambda'}^l(\hat{\mathbf{r}}_{12}), \quad (10)$$

where $R_{\mu\mu'}^m$ is Wigner's generalized harmonic and $\hat{\mathbf{r}}$ denotes the orientation of the vector \mathbf{r} (Blum 1972, 1973; Blum and Torruella 1972). With this expansion, the MOZ equation can be reduced to:

$$\begin{aligned} \tilde{h}_{\mu\nu}^{mnl}(k) &= \tilde{c}_{\mu\nu}^{mnl}(k) + \rho \sum \frac{(2l+1)}{(2n'+1)} (-)^{m+n+n'} \begin{Bmatrix} l'' & l' & l \\ m & n & n' \end{Bmatrix} \begin{pmatrix} l'' & l' & l \\ 0 & 0 & 0 \end{pmatrix} \\ &\times \sum_{\nu'} (-)^{\nu'} \tilde{h}_{\mu\nu'}^{mnl'}(k) \tilde{c}_{\nu'v}^{n'l''}(k) \end{aligned} \quad (11)$$

where the underlined subscript denotes a negative value. The parentheses $\{\dots\}$ and (\dots) indicate the 6j and 3j symbols, respectively. In this context, $\tilde{h}_{\mu\nu}^{mnl}(k)$ is given by the Hankel transform of $h_{\mu\nu}^{mnl}(r)$,

$$\tilde{h}_{\mu\nu}^{mnl}(k) = 4\pi i \int r^2 j_l(kr) h_{\mu\nu}^{mnl}(r) dr, \quad (12)$$

where j_l is an l th order spherical Bessel function. Coupling the MOZ equation with a corresponding closure, we can obtain the molecular pair correlation function. A commonly used closure for this purpose is a molecular HNC closure:

$$g(\mathbf{r}_{12}, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2) = \exp \left[-\frac{1}{k_B T} u(\mathbf{r}_{12}, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2) + h(\mathbf{r}_{12}, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2) - c(\mathbf{r}_{12}, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2) \right]. \quad (13)$$

If the spherical harmonic expansion is expanded up to infinite order, the MOZ equation is exact. In practice, however, we have to cut off the expansion at a finite order, which is usually low. As a consequence, the MOZ equation has some difficulties when it is applied to molecules with complicated geometries.

Another method to obtain the correlation function of molecular liquids is to use an interaction-site model for the pair correlation function (Hirata 2003). The site–site pair correlation function can be obtained by integration of the molecular pair correlation function over the orientation of molecules:

$$h_{\alpha\gamma}(r) = \frac{1}{\Omega^2} \iint h(\mathbf{r}_{12}, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2) \delta(\mathbf{r}_{12} + l_{2\gamma}(\boldsymbol{\Omega}_2) - l_{1\alpha}(\boldsymbol{\Omega}_1) - \mathbf{r}) d\boldsymbol{\Omega}_1 d\boldsymbol{\Omega}_2, \quad (14)$$

where $l_{1\alpha}$ denotes the vector connecting the molecular center and the interaction-site α on molecule 1. On the other hand, a superposition approximation is applied to the direct correlation function (Chandler and Andersen 1972):

$$c(\mathbf{r}_{12}, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2) = \sum_{\alpha\gamma} c_{\alpha\gamma}(r). \quad (15)$$

With these equations, the MOZ equation can be reduced to the site–site OZ equation:

$$h_{\alpha\gamma} = \sum_{\alpha'\gamma'} [\omega_{\alpha\alpha'} * c_{\alpha'\gamma'} * \omega_{\gamma'\gamma} + \omega_{\alpha\alpha'} * c_{\alpha'\gamma'} * \rho_{\gamma'} h_{\gamma'\gamma}], \quad (16)$$

where $*$ denotes the convolution integral and $\omega_{\alpha\alpha'}$ is the intramolecular correlation function:

$$\omega_{\alpha\alpha'}(r) = \delta_{\alpha\alpha'} \delta(r) + \frac{1}{4\pi l_{\alpha\alpha'}^2} (1 - \delta_{\alpha\alpha'}) \delta(r - l_{\alpha\alpha'}), \quad (17)$$

with $l_{\alpha\alpha'}$ the distance between the interaction sites α and α' . This expression is called the reference interaction-site model (RISM). A closure relation that can be used for RISM is also given by the site–site correlation function. We use the HNC closure:

$$g_{\alpha\gamma}(r) = \exp \left[-\frac{1}{k_{\text{B}}T} u_{\alpha\gamma}(r) + h_{\alpha\gamma}(r) - c_{\alpha\gamma}(r) \right]. \quad (18)$$

A more useful closure relation was proposed by Kovalenko and Hirata, which we call the Kovalenko–Hirata (KH) closure (Kovalenko and Hirata 1999a, b, 2000):

$$g_{\alpha\gamma}(r) = \begin{cases} \exp [d_{\alpha\gamma}(r)] & \text{for } d_{\alpha\gamma}(r) < 0 \\ 1 + d_{\alpha\gamma}(r) & \text{for } d_{\alpha\gamma}(r) \geq 0 \end{cases} \quad (19)$$

$$d_{\alpha\gamma}(r) = -\frac{1}{k_{\text{B}}T} u_{\alpha\gamma}(r) + h_{\alpha\gamma}(r) - c_{\alpha\gamma}(r).$$

Using the KH closure leads to a large improvement of the convergence behavior. The KH closure has been applied for example to structures of supercritical water, yielding very reasonable physical properties such as density fluctuations (Matsugami et al. 2014).

The RISM theory has been applied to various chemical and physical processes in solution. However, it is difficult for the RISM theory to handle large solute molecules with large structural anisotropies, such as proteins. To treat the solvation structure of large and anisotropic solute molecules, a three-dimensional (3D)-RISM theory has been proposed (Beglov and Roux 1997; Kovalenko and Hirata 1999b). In the 3D-RISM formalism, the molecular pair correlation function of the solute–solvent system is integrated over the orientation of solvent molecules, which yields the molecular-site or three-dimensional pair correlation function:

$$h_{\gamma}(\mathbf{r}) = \frac{1}{\Omega} \int h(\mathbf{r}_{12}, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2) \delta(\mathbf{r}_{12} + l_{2\gamma}(\boldsymbol{\Omega}_2) - \mathbf{r}) d2, \quad (20)$$

where the position of molecule 1 is located at the origin, and the molecule has a fixed orientation. The MOZ equation can be further reduced to:

$$h_{\gamma} = \sum_{\gamma'} [c_{\gamma'} * \omega_{\gamma'\gamma} + c_{\gamma'} * \rho_{\gamma'} h_{\gamma'\gamma}], \quad (21)$$

where $h_{\gamma'\gamma}$ on the right-hand side is the solvent site–site pair correlation function, which is evaluated using RISM theory. The closure relation is also given by the 3D expression, i.e., the HNC closure is:

$$g_{\gamma}(\mathbf{r}) = \exp \left[-\frac{1}{k_{\text{B}}T} u_{\gamma}(\mathbf{r}) + h_{\gamma}(\mathbf{r}) - c_{\gamma}(\mathbf{r}) \right], \quad (22)$$

and the KH closure is:

$$g_\gamma(\mathbf{r}) = \begin{cases} \exp[d_\gamma(\mathbf{r})] & \text{for } d_\gamma(\mathbf{r}) < 0 \\ 1 + d_\gamma(\mathbf{r}) & \text{for } d_\gamma(\mathbf{r}) \geq 0 \end{cases} \quad (23)$$

$$d_\gamma(\mathbf{r}) = -\frac{1}{k_B T} u_\gamma(\mathbf{r}) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r})$$

The 3D-RISM theory has been successfully applied to the solvation of biomolecules and nanomaterials (Yoshida et al. 2009).

Following the notable breakthrough of the 3D-RISM theory, an extension of the MOZ theory for anisotropic molecules in solute–solvent systems has also been proposed, which we call the extended MOZ (XMOZ) theory (Ishizuka and Yoshida 2013). In the XMOZ theory, the solute molecule is fixed at the origin of the coordinates. The spherical harmonic expansion of the pair correlation function is only applied to the orientation of the solvent molecules, whereas the coordinates of the solvent molecules remain as Cartesian coordinates. Hence, the molecular pair correlation function of the solute–solvent system is expressed as:

$$h(\mathbf{r}_{12}, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2) = h(\mathbf{r}_2, \boldsymbol{\Omega}_2) = \sum_{\substack{n \\ vv'}} h_{vv'}^n(\mathbf{r}_2) R_{vv'}^n(\boldsymbol{\Omega}_2). \quad (24)$$

The XMOZ equations in terms of the harmonic expansion coefficients are:

$$\tilde{h}_{vv'}^n(\mathbf{k}) = \tilde{c}_{vv'}^n(\mathbf{k}) + \sum_{\substack{m \\ \mu\mu'}} \tilde{c}_{\mu\mu'}^m(\mathbf{k}) (-)^{\mu-\mu'} \tilde{X}_{vv'\underline{\mu}\underline{\mu}'}^{nm}(\mathbf{k}) \quad (25)$$

$$\tilde{X}_{vv'\underline{\mu}\underline{\mu}'}^{nm}(\mathbf{k}) = \frac{\rho}{2m+1} \sum_{l\lambda'} \begin{pmatrix} m & n & l \\ \mu' & v' & \lambda' \end{pmatrix} \tilde{h}_{\mu v}^{mnl}(k) R_{\lambda'0}^l(\hat{\mathbf{k}}).$$

The closure relation for the XMOZ equations is basically the same as the one used for the MOZ theory. Thus, the HNC closure is given as:

$$g(\mathbf{r}_2, \boldsymbol{\Omega}_2) = \exp \left[-\frac{1}{k_B T} u(\mathbf{r}_2, \boldsymbol{\Omega}_2) + h(\mathbf{r}_2, \boldsymbol{\Omega}_2) - c(\mathbf{r}_2, \boldsymbol{\Omega}_2) \right]. \quad (26)$$

The molecular KH closure is:

$$g(\mathbf{r}_2, \boldsymbol{\Omega}_2) = \begin{cases} \exp[d(\mathbf{r}_2, \boldsymbol{\Omega}_2)] & \text{for } d(\mathbf{r}_2, \boldsymbol{\Omega}_2) < 0 \\ 1 + d(\mathbf{r}_2, \boldsymbol{\Omega}_2) & \text{for } d(\mathbf{r}_2, \boldsymbol{\Omega}_2) \geq 0 \end{cases} \quad (27)$$

$$d(\mathbf{r}_2, \boldsymbol{\Omega}_2) = -\frac{1}{k_B T} u(\mathbf{r}_2, \boldsymbol{\Omega}_2) + h(\mathbf{r}_2, \boldsymbol{\Omega}_2) - c(\mathbf{r}_2, \boldsymbol{\Omega}_2).$$

The applicability of the XMOZ theory is much more versatile compared with that of the MOZ theory. We should note, though, that the XMOZ theory has

a large computational cost at present. However, future contributions to reduce the computational cost can be envisaged, which will make this theory useful for calculations of biomolecules.

Solvent Effects on the Electronic Structure of Solvated Molecules

The majority of chemical reactions and almost all biochemical processes proceed in solution, the latter processes taking place in an aqueous phase. The solvent may significantly affect the progress of the reaction, as well as the properties of the solute. Thus, a knowledge of the electronic structure of solvated molecules is required, which in general is the central concern of theoretical and computational chemistry. Conventional procedures of quantum chemistry, which explicitly treat all of the molecules in the system in a quantum-chemical manner, are hardly applicable to liquids and solutions because the number of solvent molecules in the system is in the range of the Avogadro constant. For the same reason, the Schrödinger equation including the solute molecule and a large number of solvent molecules cannot be solved in practice. Thus, quantum–classical hybrid approaches have been developed. In those approaches, the solute molecule is treated quantum mechanically, while the solvent molecules are treated classically.

A popular method to treat solvated molecules is to approximate the solvent as a dielectric continuum. The quantum mechanics/molecular mechanics (QM/MM) method is also often employed, where the solvent is treated with molecular mechanics. Another approach to treat the solvent is an integral equation theory, which can be applied to molecular liquids. A pioneering work on a hybridized method of the integral equation and electronic state theories is the RISM–self-consistent field (RISM–SCF) theory introduced by Ten-no et al. (1993, 1994). Sato et al. (1996) reformulated the RISM–SCF method and derived a method to compute the analytical free energy gradient. Further extensions of this theory have been proposed, such as the 3D-RISM or MOZ theory, and have been applied to various chemical reactions in solution (Kovalenko and Hirata 1999b; Sato et al. 2000; Yoshida and Kato 2000; Yoshida 2007b).

In the hybrid method of the integral equation and electronic structure theories, first, the integral equation is solved under the electric field that is produced by the solute electronic structure. Then, an initial solvent distribution is determined and the solute electronic structure is calculated under the influence of this solvent distribution. This process is iterated, until the solute electronic structure and solvent distribution are finally determined in a self-consistent manner. In this method, the solvent distribution function and the solute electronic-density distribution connect integral equation theory and electronic structure theories. As mentioned above, RISM, 3D-RISM, and MOZ theories use different formulae for the solvent distribution functions and different methods to determine the solute electrostatic potential acting on the solvent. In a practical calculation, we need to develop a methodology to link these distributions and structures between the different theoretical frameworks.

Sato et al. (1996) derived the variational condition for the RISM–SCF framework and analytical first derivative of the Helmholtz energy of the system. In their work, they established a sound theoretical foundation for the RISM–SCF framework, which has been used to implement various properties, such as analytical free energy derivatives. 3D-RISM–SCF and MOZ theories also employ similar theoretical frameworks (Kovalenko and Hirata 1999b; Sato et al. 2000; Yoshida and Kato 2000).

We define the overall Helmholtz free energy of the system, A , as the sum of the energy of the solute, E_{solute} , and the excess chemical potential, $\Delta\mu$:

$$A = E_{\text{solute}} + \Delta\mu. \quad (28)$$

E_{solute} can be obtained via ab initio MO calculations as follows:

$$E_{\text{solute}} = \sum_{a>b} \frac{Z_a Z_b}{R_{ab}} + \sum_{ij} \gamma_{ij} h_{ij} + \frac{1}{2} \sum_{ijkl} \Gamma_{ijkl} \langle \phi_i | g_{kl} | \phi_j \rangle, \quad (29)$$

where h_{ij} and $\langle \phi_i | g_{kl} | \phi_j \rangle$ are one- and two-electron integrals, respectively. Z_n (with $n = a, b$) and R_{ab} denote the nucleus charge of the atom n and the distance between these atoms, respectively. Further, γ and ϕ_j are the density matrix and an atomic basis function, respectively. Singer and Chandler (1985) introduced $\Delta\mu$ as follows, assuming the HNC approximation as the closure:

$$\begin{aligned} \Delta\mu &= -\frac{\rho}{k_B T} \sum_{\alpha,s} \int d\mathbf{r} \left(c_{\alpha s}(r) - \frac{1}{2} h_{\alpha s}^2(r) + \frac{1}{2} h_{\alpha s}(r) c_{\alpha s}(r) \right) \\ &= -\frac{\rho}{k_B T} \sum_{\alpha,s} \int d\mathbf{r} \left\{ e^{-\frac{1}{k_B T} u_{\alpha s}(r) + t_{\alpha s}(r)} - 1 - t_{\alpha s}(r) - h_{\alpha s}(r) t_{\alpha s}(r) + \frac{1}{2} h_{\alpha s}^2(r) \right\} \\ &\quad - \frac{1}{(2\pi)^3 k_B T} \int d\mathbf{k} \left\{ -\sum_{\alpha,s} \hat{c}_{\alpha s}(k) \rho \hat{h}_{\alpha s}(k) + \frac{1}{2} \sum_{\alpha,\gamma,s,s'} \hat{c}_{\alpha s}(k) \hat{c}_{\gamma s'}(k) \hat{\omega}_{\alpha\gamma}(k) \hat{\chi}_{s s'}(k) \right\}. \end{aligned} \quad (30)$$

Here, Greek subscripts represent the solute site, while s and s' denote solvent sites. $u_{\alpha s}$ is the interaction potential between the solute and solvent sites, α and s , respectively. The potential $u_{\alpha s}$ may, for example, be given as a sum of the Coulomb interaction between partial charges and the Lennard–Jones potential between two sites:

$$u_{\alpha s}(r) = \frac{q_\alpha q_s}{r} + 4\varepsilon_{\alpha s} \left\{ \left(\frac{\sigma_{\alpha s}}{r} \right)^{12} - \left(\frac{\sigma_{\alpha s}}{r} \right)^6 \right\}, \quad (31)$$

where ε and σ are the Lennard–Jones parameters and q_x with $x = a$ and s is the partial charge on site x . The solute partial charge q_α is determined by:

$$q_\alpha = \sum_{ij} \gamma_{ij} \langle \phi_i | b_\alpha | \phi_j \rangle. \quad (32)$$

We note that q_α depends on the solute charge distribution. In this expression, b_α is an operator to estimate the partial charge, which is called the electrostatic potential (ESP) fitting. A least-square fitting process is carried out to reproduce the ESP created by the solute electrons on the solvent molecules. Thus, q_α changes with a changing solvent environment.

So far, we have treated mainly RISM–SCF theory; extensions of this theory to incorporate 3D-RISM and MOZ formulations, i.e., 3D-RISM–SCF and MOZ–SCF, have also been proposed. Recently, such frameworks have been extended to treat the electronic structure and the solvation effects of biomolecules, such as proteins and DNA. To estimate the electronic structure of biomolecules, QM/MM theory can be employed (Gao 1996).

In a general QM/MM method, the reactive moiety of proteins, for example, is treated by QM, and the other sites are treated by MM. Solvents are handled by MM or are approximated by the continuum model. Yoshida et al. (2011) proposed the QM/MM/RISM theory, where a QM/MM method is combined with the 3D-RISM–SCF theory. In this procedure, the macromolecule is treated with QM/MM and the solvent is treated using 3D-RISM, respectively. The electronic structure, which is calculated by QM, and the solvent distribution are determined self-consistently, as is done in the 3D-RISM–SCF theory.

A fragment molecular orbital (FMO)/3D-RISM method has been proposed, where the 3D-RISM–SCF theory and the FMO method are combined (Yoshida 2014). The FMO method allows us to evaluate the electronic structure of whole macromolecules by fragmentation of the molecule (Kitaura et al. 1999). Such an attempt to apply the FMO/3D-RISM method to biomolecules is expected to be developed further, along with significant upgrades of related theoretical methodologies.

To conclude this section, we show an example of recent theoretical explanation for experimental spectroscopic results. Experimentally, absorption and fluorescence spectra of an organic dye molecule can be measured as a function of the solvent polarity. The difference of the absorption and fluorescence spectra, or Stokes shift, is correlated with the solute dipole moment, which has been known as the Mataga–Lippert equation for about 60 years (Lippert 1955; Mataga et al. 1955). With the aid of RISM–SCF theory, the correlation is rationalized in terms of the solute electronic structure and the precise distribution of solvent (Nishiyama et al. 2012, 2013). Specifically, it was found that the solvent located in close vicinity of the solute (0.4 nm) significantly contributes to the Stokes shift, whereas solvent in the second solvation shell (2 nm) plays an important role in solvent rearrangement.

Moreover, using the RISM–SCF theory it was possible to reproduce the Stokes shift of particular solvents, such as tetrahydrofuran, where conventional theory using a continuum model was unable to explain the experimental data. To date, the Mataga–Lippert equation has been employed by numerous experimentalists to estimate the solute dipole moment in the photoexcited state, because of its simplicity and intuitive understanding. Therefore, the RISM–SCF theory fulfills the requirements from experimentalists in an adequate format. As another example, the

3D-RISM theory yielded clear descriptions of the distinct ion-binding modes and reinforced the proposed mechanisms based on the experimental observations.

Solvation of Biomolecules

Biomolecules are, nearly without exception, surrounded by water. This fact is essential for biomolecules to retain their structure and to perform their biological functions. Therefore, information on the solvation structure or, more specifically, the hydration structure of biomolecules is critical to understand their structures and functions. Recent developments in the statistical mechanics of the liquid phase represented by the 3D-RISM theory have been dedicated to the elucidation of solvation structures of biomolecules (Hirata 2003; Yoshida et al. 2009). Such attempts can also be applied to the analysis of biological functions. In this section, we introduce very recent research on molecular recognition using the latest developments based on statistical mechanics.

When applying the 3D-RISM theory to treat solvated biomolecules, the biomolecules are regarded as “solute molecules” surrounded by “solvents,” i.e., water molecules and ions. As has been mentioned earlier, the 3D-RISM theory yields what chemical species are distributed around what specific atoms or moieties of the particular biomolecule and to what extent. As input for such a calculation, we have to provide a solute–solvent pair correlation potential, the solvent density, and the temperature.

Figure 2 shows ion distributions of the human lysozyme mutant, determined by the 3D-RISM theory (Yoshida et al. 2006, 2007a). This theoretical result corresponds to the experimental study by Kuroki and Yutani (1998). The comparison establishes that 3D-RISM theory is able to predict the selective ionic bonds for proteins, where ions are located in a highly anisotropic environment. It is

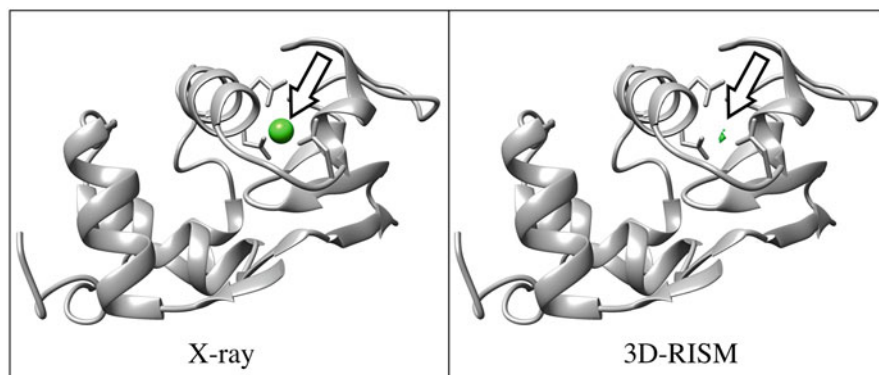


Fig. 2 Position of Ca^{2+} binding site of the Q86D/A92D mutant of human lysozyme determined using X-ray crystallography (Kuroki and Yutani 1998) and 3D-RISM theory (Yoshida et al. 2006) in the *left* and *right* panels, respectively

noted here that these results could not be found by a simple electrostatic field approach. Because the ion recognition by protein or, more commonly, the molecular recognition by protein is a thermodynamic process which is derived by a free energy change due to the recognition. Not only direct electrostatic interactions between protein and ion but those between protein–water and ion–water and also the entropy change and the reorganization of the solvent distribution should be considered. For example, the A92D mutant of the human lysozyme shows greater affinity to Na^+ than to Ca^{2+} , in spite of its smaller charge. This may be because Ca^{2+} suffers a greater dehydration penalty upon binding compared with the case of Na^+ .

Another application example of the 3D-RISM theory is ion permeation of the KcsA channel. The KcsA channel is a channel protein derived from bacteria, which functions as a control of K^+ concentration inside the cell. The KcsA channel has an excellent selectivity for K^+ to permeate through the channel, even though concentrations of Na^+ on both sides of the cell are very large (Hill 2001).

Experiments in 2009 indicated that these ions, K^+ and Na^+ , have different distributions or bonding modes (Thompson et al. 2009). Figure 3 shows the distributions of ions inside the KcsA channel obtained using the 3D-RISM theory, which properly reproduces the experimental results (Phongphanphanee et al. 2014a, b). Furthermore, the theory determines the water distribution in the channel, which provides the Li^+ and K^+ distributions in more detail. In the stable structure, Li^+ takes a six-coordinated structure: four carbonyl oxygens are coordinated in a square geometry and two waters are in the above and below positions. In contrast, K^+ is not in immediate contact with water molecules; this ion is positioned in the center of a cage provided by eight carbonyl oxygens. The predicted bonding modes for these ions correspond well with experimental and simulation results.

The 3D-RISM theory has also been used to describe the molecular recognition processes of various biomolecules, such as the escaping process of carbon monoxide from myoglobin (Kiyota et al. 2009), proton exclusion mechanism of aquaporin (Phongphanphanee et al. 2008), and proton permeation mechanisms at the M2

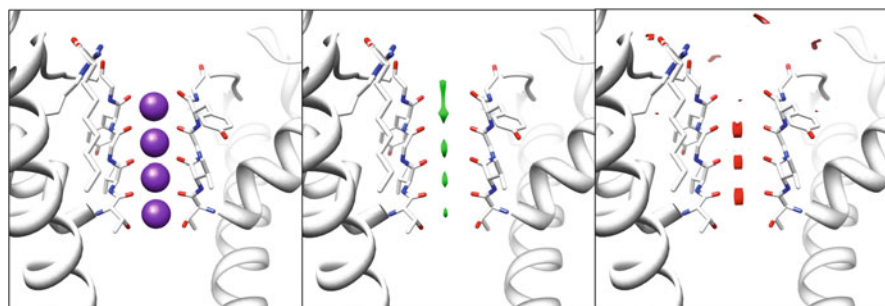


Fig. 3 K^+ position in the KcsA channel taken from the experimental result (Zhou et al. 2001) is presented as the *purple spheres* in the *left panel*. The isosurface plot of the ion distributions of K^+ and Li^+ evaluated by 3D-RISM theory are colored by *green* and *red* in the *center* and *right panels*, respectively

channel of influenza A (Phongphanphanee et al. 2010). In addition to small molecules such as water and ions, the theory can also be applied to realistic drug recognition (Imai et al. 2009; Kiyota et al. 2011).

In the example mentioned above, the 3D-RISM calculations were performed based on the static X-ray structure of protein. For a large biomolecule, the effects of the dynamics and fluctuations of the molecular structure may be far from negligible and need to be considered. In order to include the effects of fluctuations of the protein structure, a combination of 3D-RISM with molecular dynamics (MD) simulation has been proposed, which may reveal more detailed information on solvated biomolecules for a larger variety of systems (Luchko et al. 2010; Miyata and Hirata 2008; Miyata et al. 2010, 2011). In the proposed methods, the multiple time step technique has been employed to reduce the computational cost. More recently, Omelyan and Kovalenko (2015) developed an efficient hybrid algorithm employing a generalized solvation force extrapolation approach to accelerate the multiple time step MD. With this method, the folding of a miniprotein, starting from a fully denatured state, could be simulated successfully.

Another way to take into account the fluctuation of the structure of a solvated protein has been proposed by Kim and Hirata (2013). In this method, the fluctuations of the protein, coupled with that of the solvent, is treated using methodologies based on the generalized Langevin equation and 3D-RISM theory with the aid of linear response theory.

Relying on these methods, it may be soon possible to treat biological events accompanied with large structural changes of protein molecules computationally.

Summary and Future Perspective

In this chapter, we have outlined the development of a theoretical treatment to describe molecular liquids and assemblies based on statistical mechanics. The theory of molecular liquids has been developed to cover the wide area of physics, chemistry, and biology.

Examples of the application of this theoretical description include chemical reactions in solution (Sato 2013). The treatment of chemical reactions in solution has been developed successfully in the last decades by combining electronic structure theory and the molecular theory of liquids. The recent progress in theoretical methods enables a detailed explanation of novel phenomena that have been found by experiment, by including a molecular view of individual molecules participating in the reaction.

Another important aspect of the recent molecular theory should include the description of biological processes (Yoshida et al. 2009). To our knowledge, the 3D-RISM theory is one of the most developed and practical frameworks among the theoretical procedures available so far that can be applied to biological processes taking place in solution.

In the future, further close collaborations between experimentalists and theoreticians will be required to discover and analyze novel phenomena. We believe that

such collaborations among scientists in the field of solution chemistry and physics and bioscience will be extremely fruitful. One may also envisage that current attempts of theoretical development may be used to solve social problems we are facing currently, such as energy problems, environment, and food problems.

Bibliography

- Beglov, D., & Roux, B. (1997). An integral equation to describe the solvation of polar molecules in liquid water. *Journal of Physical Chemistry B*, *101*(39), 7821–7826.
- Blum, L. (1972). Invariant expansion. 2. Ornstein–Zernike equation for nonspherical molecules and an extended solution to the mean spherical model. *Journal of Chemical Physics*, *57*(5), 1862–1869.
- Blum, L. (1973). Invariant expansion. 3. General solution of mean spherical model for neutral spheres with electrostatic interactions. *Journal of Chemical Physics*, *58*(8), 3295–3303.
- Blum, L., & Torruella, J. (1972). Invariant expansion for two-body correlations: Thermodynamic functions, scattering, and the Ornstein–Zernike equation. *Journal of Chemical Physics*, *56*(1), 303–310.
- Chandler, D., & Andersen, H. C. (1972). Optimized cluster expansions for classical fluids. 2. Theory of molecular liquids. *Journal of Chemical Physics*, *57*(5), 1930–1937.
- Gao, J. (1996). Methods and applications of combined quantum mechanical and molecular mechanical potentials. In K. B. Lipkowitz & D. B. Boyd (Eds.), *Reviews in computational chemistry* (Vol. 7). Hoboken: Wiley.
- Hansen, J. P., & McDonald, I. R. (2006). *Theory of simple liquids*. Amsterdam: Academic.
- Hill, B. (2001). *Ion channels of excitable membranes*. Sunderland: Sinauer Associates.
- Hirata, F. (2003). *Molecular theory of solvation*. Dordrecht: Kluwer.
- Imai, T., Oda, K., Kovalenko, A., Hirata, F., & Kidera, A. (2009). Ligand mapping on protein surfaces by the 3D-RISM theory: Toward computational fragment-based drug design. *Journal of the American Chemical Society*, *131*(34), 12430–12440.
- Ishizuka, R., & Yoshida, N. (2013). Extended molecular Ornstein–Zernike integral equation for fully anisotropic solute molecules: Formulation in a rectangular coordinate system. *Journal of Chemical Physics*, *139*(8), 084119.
- Kim, B., & Hirata, F. (2013). Structural fluctuation of protein in water around its native state: A new statistical mechanics formulation. *Journal of Chemical Physics*, *138*, 054108 (1–11).
- Kitaura, K., Ikeo, E., Asada, T., Nakano, T., & Uebayasi, M. (1999). Fragment molecular orbital method: An approximate computational method for large molecules. *Chemical Physics Letters*, *313*, 701–706.
- Kiyota, Y., Hiraoka, R., Yoshida, N., Maruyama, Y., Imai, T., & Hirata, F. (2009). Theoretical study of CO escaping pathway in myoglobin with the 3D-RISM theory. *Journal of the American Chemical Society*, *131*(11), 3852–3853.
- Kiyota, Y., Yoshida, N., & Hirata, F. (2011). A new approach for investigating the molecular recognition of protein: Toward structure-based drug design based on the 3D-RISM theory. *Journal of Chemical Theory and Computation*, *7*(11), 3803–3815.
- Kovalenko, A., & Hirata, F. (1999a). Potential of mean force between two molecular ions in a polar molecular solvent: A study by the three-dimensional reference interaction site model. *Journal of Physical Chemistry B*, *103*(37), 7942–7957.
- Kovalenko, A., & Hirata, F. (1999b). Self-consistent description of a metal–water interface by the Kohn–Sham density functional theory and the three-dimensional reference interaction site model. *Journal of Chemical Physics*, *110*(20), 10095–10112.
- Kovalenko, A., & Hirata, F. (2000). Potentials of mean force of simple ions in ambient aqueous solution. I. Three-dimensional reference interaction site model approach. *Journal of Chemical Physics*, *112*(23), 10391–10402.

- Kuroki, R., & Yutani, K. (1998). Structural and thermodynamic responses of mutations at a Ca^{2+} binding site engineered into human lysozyme. *Journal of Biological Chemistry*, 273(51), 34310–34315.
- Levine, R. D. (2005). *Molecular reaction dynamics*. Cambridge: Cambridge University Press.
- Lippert, E. (1955). Dipolmoment und Elektronenstruktur von angeregten Molekülen. *Zeitschrift für Naturforschung*, 10a, 541–545.
- Luchko, T., Gusarov, S., Roe, D. R., Simmerling, C., Case, D. A., Tuszynski, J., & Kovalenko, A. (2010). Three-dimensional molecular theory of solvation coupled with molecular dynamics in amber. *Journal of Chemical Theory and Computation*, 6(3), 607–624.
- Mataga, N., Kaifu, Y., & Koizumi, M. (1955). The solvent effect on fluorescence spectrum, change of solute–solvent interaction during the lifetime of excited solute molecule. *Bulletin of the Chemical Society of Japan*, 28, 690–691.
- Matsugami, M., Yoshida, N., & Hirata, F. (2014). Theoretical characterization of the “ridge” in the supercritical region in the fluid phase diagram of water. *Journal of Chemical Physics*, 140(10), 104511.
- Miyata, T., & Hirata, F. (2008). Combination of molecular dynamics method and 3D-RISM theory for conformational sampling of large flexible molecules in solution. *Journal of Computational Chemistry*, 29(6), 871–882.
- Miyata, T., Ikuta, Y., & Hirata, F. (2010). Free energy calculation using molecular dynamics simulation combined with the three dimensional reference interaction site model theory. I. Free energy perturbation and thermodynamic integration along a coupling parameter. *Journal of Chemical Physics*, 133(4), 044114.
- Miyata, T., Ikuta, Y., & Hirata, F. (2011). Free energy calculation using molecular dynamics simulation combined with the three-dimensional reference interaction site model theory. II. Thermodynamic integration along a spatial reaction coordinate. *Journal of Chemical Physics*, 134(4), 044127.
- Nishiyama, K., Watanabe, Y., Yoshida, N., & Hirata, F. (2012). Solvent effects on electronic structures of coumarin 153: Parallel studies by means of spectroscopy and RISM–SCF calculations. *Journal of the Physical Society of Japan*, 81, SA016.
- Nishiyama, K., Watanabe, Y., Yoshida, N., & Hirata, F. (2013). Solvent dependence of Stokes shift for organic solute–solvent systems: A comparative study by spectroscopy and reference interaction-site model–self-consistent-field theory. *Journal of Chemical Physics*, 139(9), 094503.
- Omelyan, I., & Kovalenko, A. (2015). MTS-MD of biomolecules steered with 3D-RISM-KH mean solvation forces accelerated with generalized solvation force extrapolation. *Journal of Chemical Theory and Computation*, 11, 1875–1895.
- Phongphanphanee, S., Yoshida, N., & Hirata, F. (2008). On the proton exclusion of aquaporins: A statistical mechanics study. *Journal of the American Chemical Society*, 130(5), 1540–1541.
- Phongphanphanee, S., Rungrotmongkol, T., Yoshida, N., Hannongbua, S., & Hirata, F. (2010). Proton transport through the influenza A M2 channel: Three-dimensional reference interaction site model study. *Journal of the American Chemical Society*, 132, 9782–9788.
- Phongphanphanee, S., Yoshida, N., Oiki, S., & Hirata, F. (2014a). Distinct configurations of cations and water in the selectivity filter of the KcsA potassium channel probed by 3D-RISM theory. *Journal of Molecular Liquids*, 200, 52–58.
- Phongphanphanee, S., Yoshida, N., Oiki, S., & Hirata, F. (2014b). Probing “ambivalent” snug-fit sites in the KcsA potassium channel using three-dimensional reference interaction site model (3D-RISM) theory. *Pure and Applied Chemistry*, 86(2), 97–104.
- Sato, H. (2013). A modern solvation theory: Quantum chemistry and statistical chemistry. *Physical Chemistry Chemical Physics*, 15, 7450–7465.
- Sato, H., Hirata, F., & Kato, S. (1996). Analytical energy gradient for the reference interaction site model multiconfigurational self-consistent-field method: Application to 1,2-difluoroethylene in aqueous solution. *Journal of Chemical Physics*, 105(4), 1546–1551.

- Sato, H., Kovalenko, A., & Hirata, F. (2000). Self-consistent field, ab initio molecular orbital and three-dimensional reference interaction site model study for solvation effect on carbon monoxide in aqueous solution. *Journal of Chemical Physics*, *112*(21), 9463–9468.
- Singer, S. J., & Chandler, D. (1985). Free energy functions in the extended RISM approximation. *Molecular Physics*, *55*(3), 621–625.
- Ten-no, S., Hirata, F., & Kato, S. (1993). A hybrid approach for the solvent effect on the electronic structure of a solute based on the RISM and Hartree–Fock equations. *Chemical Physics Letters*, *214*(3–4), 391–396.
- Ten-no, S., Hirata, F., & Kato, S. (1994). Reference interaction site model self-consistent-field study for solvation effect on carbonyl compounds in aqueous solution. *Journal of Chemical Physics*, *100*(10), 7443–7453.
- Thompson, A. N., Kim, I., Panosian, T. D., Iverson, T. M., Allen, T. W., & Nimigean, C. M. (2009). Mechanism of potassium-channel selectivity revealed by Na⁺ and Li⁺ binding sites within the KcsA pore. *Nature Structural and Molecular Biology*, *16*(12), 1317–1324.
- Tomasi, J., Mennucci, B., & Cammi, R. (2005). Quantum mechanical continuum solvation models. *Chemical Review*, *105*(8), 2999–3093.
- Yoshida, N. (2007a). Analytical free energy gradient for the molecular Ornstein–Zernike self-consistent-field method. *Condensed Matter Physics*, *10*(3), 363–372.
- Yoshida, N., Phongphanphanee, S., & Hirata, F. (2007b). Selective ion binding by protein probed with the statistical mechanical integral equation theory. *Journal of Physical Chemistry B*, *111*(17), 4588–4595.
- Yoshida, N. (2014). Efficient implementation of the three-dimensional reference interaction site model method in the fragment molecular orbital method. *Journal of Chemical Physics*, *140*(21), 214118.
- Yoshida, N., & Kato, S. (2000). Molecular Ornstein–Zernike approach to the solvent effects on solute electronic structures in solution. *Journal of Chemical Physics*, *113*(12), 4974–4984.
- Yoshida, N., Phongphanphanee, S., Maruyama, Y., Imai, T., & Hirata, F. (2006). Selective ion-binding by protein probed with the 3D-RISM theory. *Journal of the American Chemical Society*, *128*(37), 12042–12043.
- Yoshida, N., Imai, T., Phongphanphanee, S., Kovalenko, A., & Hirata, F. (2009). Molecular recognition in biomolecules studied by statistical-mechanical integral-equation theory of liquids. *Journal of Physical Chemistry B*, *113*(4), 873–886.
- Yoshida, N., Kiyota, Y., & Hirata, F. (2011). The electronic-structure theory of a large-molecular system in solution: Application to the intercalation of proflavine with solvated DNA. *Journal of Molecular Liquids*, *159*(1), 83–92.
- Zhou, Y. F., Morais-Cabral, J. H., Kaufman, A., & MacKinnon, R. (2001). Chemistry of ion coordination and hydration revealed by a K⁺ channel-Fab complex at 2.0 Å resolution. *Nature*, *414*, 43–48.

Part III

Solid States and Nanomaterials

Manthos Papadopoulos and Heribert Reis

Photoactive Semiconducting Oxides for Energy and Environment: Experimental and Theoretical Insights

23

Malgorzata Makowska-Janusik and Abdel-Hadi Kassiba

Contents

Introduction	985
Experimental Insights on Materials and Related Features	987
Titanium Dioxide (TiO ₂)	987
BiVO ₄ Nanostructures for Photocatalysis	993
Theoretical Background	998
Modeling of Crystalline Structures: Preamble	998
DFT Methodology and Related Functionals	1001
Cluster Approach to Model Nanostructures	1005
Structural Features of TiO ₂ , NiTiO ₃ , and BiVO ₄ Crystals	1006
Computational Parameters	1006
Physical Properties of NiTiO ₃ and BiVO ₄ Crystals	1009
Structural Aspects	1009
Electronic Properties of BiVO ₄	1009
Electronic Properties of NiTiO ₃	1012
Structural and Electronic Properties of TiO ₂ , NiTiO ₃ , and BiVO ₄ Nanosized Systems ..	1013
BiVO ₄ Nanostructures	1021
Conclusions	1023
Bibliography	1024

M. Makowska-Janusik (✉)
Institute of Physics, Faculty of Mathematics and Natural Science, Jan Dlugosz University in
Czestochowa, Czestochowa, Poland
e-mail: m.makowska@ajd.czest.pl

A.-H. Kassiba
Institute of Molecules and Materials of Le Mans – UMR-CNRS 6283, Université du Maine,
Le Mans, France
e-mail: kassiba@univ-lemans.fr

Abstract

This chapter reports experimental investigations and theoretical approaches devoted to analyze the electronic, optical, and vibrational properties occurring in semiconducting photoactive materials in their bulk form or as nanosized objects. An original cluster approach was developed to describe the physical features of nanoparticles and the impact of their surface on photoinduced charge transfer phenomena.

Photovoltaic and photocatalysis are two major applications for both clean environment and sustainable sources of energy as well. In this context, semiconducting nanocrystals from defined oxide families have attracted increasing interest during the last decade for their promising potential in renewable energy applications because of their versatile and coupled optical and electronic properties. More specifically, the semiconductor oxides as titanium oxide and bismuth vanadate-based materials can be tailored in nanosized and mesoporous structures. The electronic and optical features may be modulated in large extents when dye molecules are used as sensitizing vectors enhancing the efficiency of solar cells, energy storage, and photocatalyst. These applications depend on the reactive surface area, morphology and nanostructuring, doping and sensitizing agents, as well as controlled vacancy rates acting on the charge transfer peculiarities.

In parallel to experimental investigations dedicated to selected forms of photoactive semiconducting oxides, original theoretical approaches were developed to analyze the key features of the considered functional systems. Using an integrated approach associating theoretical models and numerical simulations, the influence of the size and morphology of nanoparticles on their electronic and optical properties was pointed out. The cluster approach methodology was developed to simulate the electronic properties of semiconducting nanocrystalline materials as isolated objects or functionalized by organic dye molecules. The construction of the system proceeds by the crystal structure frozen in the cluster core while the surface is modified according to the environmental interactions. Theoretically, it was proved that nanostructures exhibit patchwork properties coming from the bulk material including core crystal structure and from surface effects caused by environment. On the other hand, the role of doping of the considered structures by metallic elements was investigated on the photoactivity mechanism involved in nanoparticles. The nature of vacancies located close to the dopants plays a crucial role on the electronic and optical features in the photoactive materials. Thus, the theoretical approaches and the carried out numerical simulations contribute to draw quantitative insights of the physical properties of functional semiconducting oxides in agreement with relevant experimental analyses.

Introduction

Photoactive semiconducting oxides are subject of intensive investigations during the last two decades with a major aim to realize active media for photovoltaic and photocatalytic processes. The applications concern important societal challenges, such as green energies based on alternative and nonfossil resources and environmental problems related to water and atmosphere purification and preservation. Particularly, hydrogen production from water dissociation and depollution of environment from contaminants are major concerns of the twenty-first-century society. Increased industrial activity due to the emergence of new economic powers along with important demographic pressure contributes to high energy demand. Average estimates indicate energy resources exploited nowadays consummate around 80 % from nonrenewable energy origin even if the part of other energy sources (solar, wind) increase continuously in industrial countries during the last 20 years. Similarly, the improvement of living conditions in emerging countries through technological developments requires an increasing energy demand and water consumption. It is obvious that the factors listed above (water, energy, industry, demography, development) generate an impact on the environment. The increasing air and water pollution is a crucial health and societal problem which requires control and regulation and then the societal issues must find upstream solutions to prevent crises. Beyond the precautionary principles and good practices from the societies, scientific approaches at large scale should be conducted to support changes in the demand for energy and environmental preservation required for suitable evolution of the human life standards.

In this context, the scientific activity related to photocatalytic processes has been exploited intensively for more than two decades in the active research area dealing with energy and environment. In particular, harvesting efficiently the solar spectrum in the visible light-driven photocatalysis is subject of important developments to create photoactive materials with an implementation of experimental protocols devoted to increase the efficiency of the photocatalysts. These mechanisms may be applied to the production of clean energy sources such as hydrogen production obtained by water splitting. They also apply to the degradation of organic pollutants or bacterial organisms that can affect water reserves or some critical atmospheres as required in medical environments. Other large-scale applications of photocatalytic reaction principles were tested on road surfaces containing photocatalysts to destroy polluting gases such as oxynitrides or hydrocarbons. These examples demonstrate the wealth of applications that rely on photocatalysis from the point of view of the active material designs as well as for the optimization of photocatalytic reactions. Among the wide classes of active materials, two families of semiconducting oxides are worthy of interest and were subject of important developments. They are

related, on the one hand, to titanium dioxide (TiO_2) materials as pure or associated with organic or inorganic components following the searched applications from photovoltaics or photocatalysis. On the other hand, bismuth vanadate (BiVO_4) and related structures are emerging smart materials for photocatalysis.

The experimental developments on such classes of photoactive materials are supported by theoretical and numerical models able to predict and analyze quantitatively the experimental properties of bulk and nanosized materials from charge transfer peculiarities to optical absorption and structural organization. Original approaches were developed on clusters with different size giving rise to behaviors related to molecular scales, nanosized clusters, or to macroscopic media. For photovoltaic or photocatalytic processes activated by light irradiation in semiconducting materials, the phenomena of charge transfer and the optical absorption features are of primary importance, and the simulation methods can contribute to shed light on the strength of the phenomena and orient the ways of their optimization. For such approaches, quantum chemistry codes and software packages based on quantum mechanics are active and hot topics for the simulations in material sciences. The density functional theory (DFT) method is an efficient tool to predict the physical properties of materials. The built of realistic geometries of the materials, the optimization of the structures and atomic position, as well as the evaluation of the related properties by DFT approach are currently applied to wide classes of functional materials including semiconducting photoactive media.

Several models were explored on selective classes of photoactive materials such as TiO_2 , BiVO_4 , and NiTiO_3 dedicated to photocatalysis and photovoltaic applications. For titanium dioxide structures, original theoretical studies have been realized to model anatase structure of nanocrystals with sizes ranging from 0.7 nm up to 2.0 nm organized as $(\text{TiO}_2)_n$ clusters. Also, the systems like BiVO_4 or close ones as NiTiO_3 were modeled by applying the cluster approach developed in our previous work (Makowska-Janusik and Kassiba 2012). For the nanostructures, the influence of surface and environment at their electronic properties was studied. Similar theoretical and numerical modeling studies were applied to predict and describe quantitatively the key physical properties of BiVO_4 and NiTiO_3 . Numerical approaches based on cluster model and quantum-chemical calculations were implemented to explain the electronic properties of nanoparticles and to elucidate the influence of surface effects on photocatalytic-driven mechanism such as charge transfer. The developed approaches are emanating from former works developed on bulk BiVO_4 structures leading to their electronic properties (Walsh et al. 2009; Yin et al. 2011; Cooper et al 2014; Ma and Wang 2014).

The implemented theoretical approach applies to model the features of mesoporous structure offering high specific surfaces which enhance photocatalytic and photovoltaic properties. The experimental investigations ensuring the synthesis of such structures and the evaluation of related vibrational, electronic, and optical features are confronted to the predicted properties obtained by numerical simulations.

Experimental Insights on Materials and Related Features

Titanium Dioxide (TiO₂)

Photoactivity of Titanium Dioxide

Among the efficient materials for their photocatalytic applications, titanium dioxide (TiO₂) is undeniably the most known and investigated. The green energy such as hydrogen production and the new generation of solar cells as well as the photocatalysis dedicated to the preservation of the environment have been massively exploiting active materials based on TiO₂. Particularly, photovoltaic applications were investigated using TiO₂ in its pure or doped forms by metallic or nonmetallic ions or alternatively associated with other components including metallic clusters or organic groups. Various morphologies were considered such as micrometric or nanometric powders (Ishigaki et al. 2003), mesoporous structures or photoactive gels (Pattier et al. 2010), and thin films (Nam et al. 2004) or nanocomposites. The morphology and structural organization of titanium dioxide plays key roles on its electronic and optical features and was concerned by exhaustive investigations by a large scientific community. Thus, nanostructured titanium oxide as thin films, nanoparticles (Reyes-Coronado et al. 2008), or mesoporous organizations (Kassiba et al. 2012) were realized with the relevant parameters leading to the requested structural organization (rutile, anatase, amorphous), morphologies (nanoparticles, nanodots, nanotubes), as well as self-assembled media with controlled porosity. The functionalization of titanium dioxide surface by active vectors was realized in order to improve the photovoltaic performances of devices based on such class of composites (Melhem et al. 2013). However, such elaboration processes are relevant to graft other functional groups on high specific surface materials. This is the case of hybrid networks composed by inorganic nanocrystals – polymer blend for new generation of hybrid solar cells (Lin et al. 2009). Other structures based on titanium oxides were also considered to improve the photo-conversion efficiency by high specific surfaces between the active components of solar cells (Hwang et al. 2009). Thus, sol–gel batches based on titanium oxide gels (Pattier et al. 2010) were realized, and their photoactivity under UV was investigated in the aim to realize new generation of hybrid solar cells. Nevertheless, despite the major limitations of TiO₂ due to a wide band gap (3.2 eV), large-scale applications have been conducted and extensively used as appropriate structures for photovoltaic and photocatalysis. For applications, the photoactivity enhancement can be achieved by mechanisms as doping by metallic or nonmetallic ions able to modulate the electronic structure. The combination of TiO₂ in composites with semiconductor structures such as ZnO (Chandiran et al. 2014) and CdS (Biswas et al. 2008) has been made, and the photoinduced effects were compared to those involved in pure materials. However, even with an increase in the efficiency by such experimental modifications, the UV radiation required for the photoactivity is costly,

and it is necessary to develop alternative media with visible driven photoactivity or for high-efficient solar cells. Also, the well-known chemical procedure based on dyes sensitized solar cells (DSSC) has created the most investigated systems based on TiO_2 and offers the best efficiency compared to heterostructure-based devices. Other hybrid nanostructured architectures were realized by using titanium dioxide and graphene to ensure challenging energy sources in new-generation of Li-ion batteries (Wang et al. 2014). Recently other approaches have created oxygen-deficient titanium dioxide associated with conjugated polymer. The realized functional media shows a noticeable improvement of efficiency of the based solar cell and the optoelectronic device such as OLEDs.

In addition to these structural or composition of the functional materials based on titanium dioxide, the doping process is worthy of interest to modulate the electronic and optical responses of the materials. Also, several strategies were used for doping with metallic (Fe, Ag, Mo, Eu, etc.) or nonmetallic elements (N, C, etc.). Exhaustive investigations of the doping effects were carried out to analyze the electronic, optical, and magnetic features of the substitutional ions and to identify their local environments and the quality of their distribution in the host titanium dioxide structures. Such studies have been made on mesoporous architectures where several chemical reactions were also involved and catalyzed by oxygen being adsorbed in voids or incorporated in the bulk as anti-sites. This contributes to create new electronic states allowed in the band gap of the doped material and were discussed in several works as illustrated in the forthcoming part. However, instead of the doping effect, electronic and optical features of photoactive gels composed by titanium oxide can be also modified by UV irradiation leading to photoinduced electrons localized as Ti^{3+} centers (Pattier et al. 2010). The mechanism of the Ti^{3+} formation is well described by a series of reactions starting by the creation of a photoinduced electron–hole pairs. This process makes titanium oxide as efficient as dye sensitized one and can be also as crucial as doping agents may do on the electronic behavior. Such photoactive media ($\text{Ti}_i^{4+} \xrightarrow{h\nu} \text{Ti}_i^{3+}$) was tested for new generation of solar cells (Pattier et al. 2010). The experimental investigations need quantitative analysis by using theoretical and modeling approaches in order to improve the photovoltaic or photocatalytic activities based on doped titanium dioxide-based materials (Makowska-Janusik et al. 2014). Several theoretical works have focused on the study of the electronic and optical features of titanium oxide possessing different stoichiometry and applying different computational approaches. The performed numerical simulations deal mainly with the electronic properties of the infinite crystal (Stausholm-Møller et al. 2010). Other performed approaches consider titanium oxide clusters with sizes up to 1–2 nm possessing crystalline core structure emanating from anatase or rutile (Makowska-Janusik et al. 2014). In all cases, the surface structure, morphology, and organization are inherent problem in numerical simulations on nano-objects with regard to their key role on the predicted properties.

Synthesis of Mesoporous TiO₂ Material and Selected Investigations

The synthesis of mesoporous TiO₂ is dedicated to create high specific surface. Titanium oxide-based mesoporous structures were synthesized by using the sol-gel method (Kassiba et al. 2012; Mei et al. 2010) with precursors as titanium tetraisopropoxide and thiourea for nitrogen doping. Surfactants as pluronic P123 are usually used to realize porous (structured) titanium dioxide with or without nitrogen doping. The materials can be obtained in the form of mesoporous powders or thin films. The as-formed systems were submitted to annealing treatments at temperatures up to 500 °C in order to improve the crystalline structure of the titanium dioxide with a stable polytype such as anatase.

The experimental investigations were carried out to identify the crystalline structures involved at defined synthesis conditions and after treatments by using X-ray diffraction and particularly wide-angle X-ray scattering (WAXS) technique well adapted for mesoporous structures. The major information concern also the porosity features and in the case of doped structures the doping efficiency through the location of substitutional ions in the crystalline sites of the host structure. In this aim, electron paramagnetic spectroscopy (EPR) is a relevant technique to characterize the active electronic centers in mesoporous structures. Beyond to identify the nature of the active centers, EPR informs on the local environments of doping agents and their electronic, optical, and magnetic properties. The microstructure of the samples plays a key role in the stability of the paramagnetic centers and their incorporation in the network (gel, mesoporous) (Pattier et al. 2010; Kassiba et al. 2012). Thus, comparative investigations performed on mesoporous materials inform on the efficiency of the doping process and on the concentration and nature of the involved oxygen-based radicals and vacancies. The large specific surfaces of these nanostructured systems enhance the concentration of oxygen- or nitroxide-based radicals and contribute to lower the doping efficiency.

X-ray diffraction measurements were conducted on TiO₂ powders with porous morphologies (structured) or without pores (nonstructured) with or without doping. Figure 1 (top) shows WAXS patterns recorded on nonstructured samples with and without nitrogen doping (N-TiO₂-400-4.5 and TiO₂-400-4.5); the samples were annealed at 400 °C during 4.5 h under air (Mei et al. 2010). The analysis of the diffraction line features by using the MAUD software program (Ferrari and Lutterotti 1994) evaluates the involved crystalline polytypes and their respective fractions. The structural investigations reveal the role of the doping and the morphology (structured or nonstructured) on the stabilized crystalline polytypes. Thus, the appropriate annealing conditions have been defined at 500 °C during 12 h to realize pure anatase structure with suitable substitutional doping in the crystal sites. In this case, the same WAXS patterns show similar crystalline features whatever the samples such as structured or not and doped or free from any doping agents.

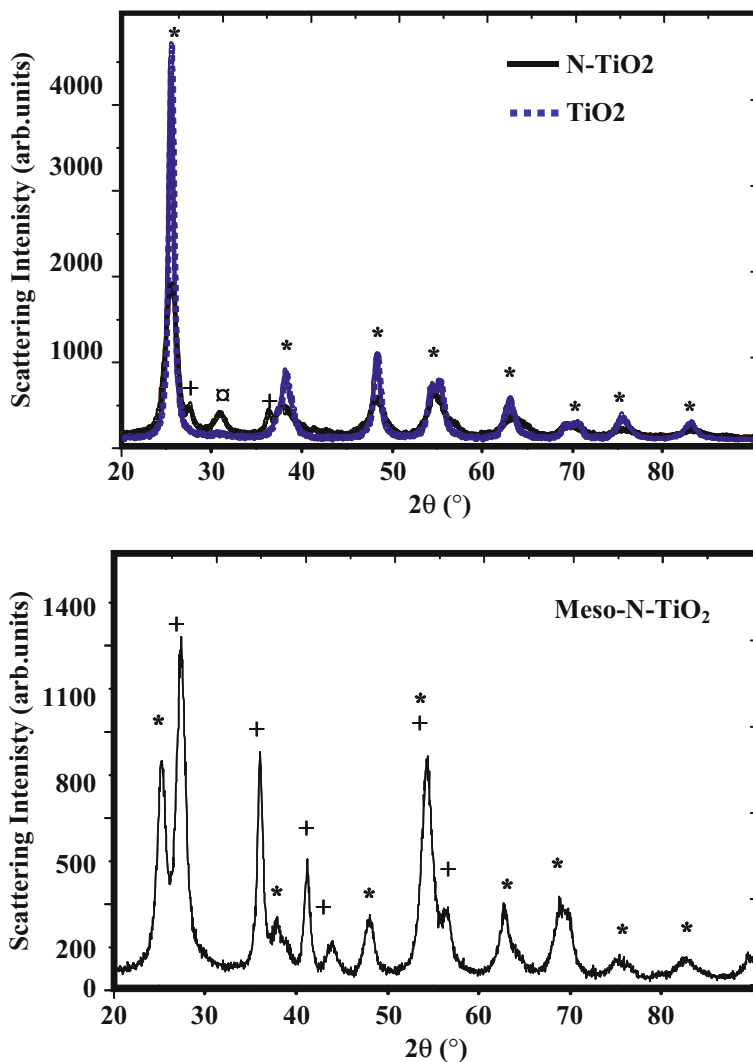


Fig. 1 WAXS patterns showing the involved crystalline structures in nonstructured TiO₂ samples (*top*) with nitrogen doping or without (*:anatase, +: rutile, ⊠: brookite) and in mesoporous sample (*bottom*) doped with nitrogen and annealed for 4.5 h at 400 °C

Sol-gel method and spin-coating technique were used to realize mesoporous thin films (Pattier et al. 2010; Pattier 2010). The structural analysis can be made by X-ray reflectivity which informs on the ordering of pore lattices as a function of the thermal treatment of the films. Thus, characterizations by X-ray reflectivity were performed on mesoporous samples as a function of their annealing temperature and the time duration of this process. The existence of ordered porous structure

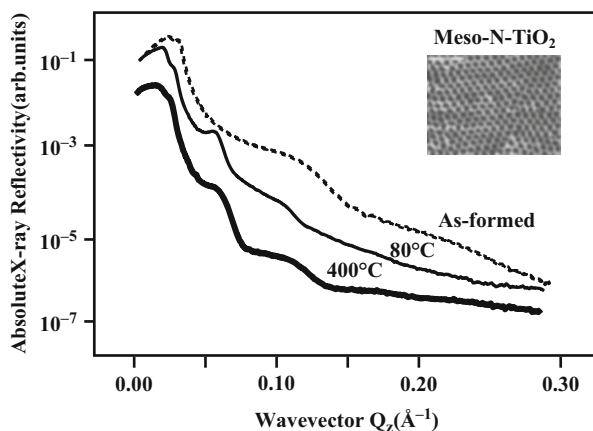
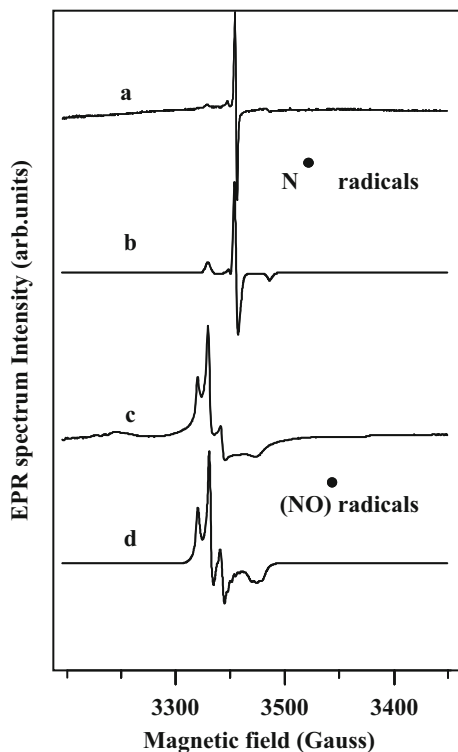


Fig. 2 X-ray reflectivity measurements performed on thin films of mesoporous TiO_2 with nitrogen doping. The organization of the pores is uniform at defined annealing temperature and time duration. The period of the pore on the layer changes from 10.9 nm for the as-formed sample to 5.4 nm in mesoporous film annealed at 400°C during 30 min (Pattier et al. 2010)

is demonstrated as well as the evolution of the pore sizes with the annealing conditions (Fig. 2). Evaluations of the pore sizes were obtained by adjusting the X-ray reflectivity curves. The pore size was estimated to 10.9 nm in the as-formed sample and change to 5.4 nm when the annealing is performed at 400°C during 4.5 h. This behavior is well understood by the contraction of the porous phase after extraction of the surfactant by annealing (Pattier et al. 2010; Kassiba et al. 2012).

When doping is performed, the location of the substitutional ions in defined crystal sites or alternatively the existence of clusters or aggregates made from doping species can be identified by EPR. Such experiments were carried out to analyze the nitrogen doping from thiourea precursors as well as all active electronic centers (oxygen radicals, titanium ions, nitro-oxide radicals, etc.) which can contribute to the conductivity or act as traps for charge carriers. For the considered mesoporous structures, the existence of high specific surface favors the location of doping elements at the surface. Thus, Fig. 3 shows the EPR spectra recorded on the mesoporous samples and nonporous samples, annealed under air or argon at 400°C during 30 min. These experiments show that the atmosphere has a drastic effect on the nature of involved paramagnetic centers. Instead of the observation of nitrogen radicals, a variety of paramagnetic species, e.g., O_2^- , O_3^- , NO , Ti^{3+} , was demonstrated. However, the nature of paramagnetic radicals and the role of the annealing temperatures and its duration were clarified on the considered N- TiO_2 -doped samples. The nitrogen doping is effective but without to be incorporated in oxygen sites. Resolved signals implying nitrogen was observed for mesoporous samples annealed at 500°C during 12 h with the formation of NO radicals located on the surface of the porous structure. For nonporous samples, the paramagnetic species such as oxygen radicals were suppressed for the same

Fig. 3 EPR spectra of N-TiO₂ samples annealed at 500 °C. Nitrogen ions are substituted in the crystalline sites of TiO₂ only for samples without porous structures (a-experiment, b-simulation). For the porous one, only surface (NO)* radicals dominate the paramagnetic species behind the observed EPR signal (c-experiment, d-simulation)



treatment (500 °C, 12 h) with the formation of nitrogen radicals (N^{*}). The developed EPR signal which appears after annealing is characterized by narrow lines (Fig. 3) which indicate well-crystallized sites around the nitrogen ions. Thus, only in the case of nonstructured (nonporous) TiO₂, nitrogen ions are indeed incorporated in the bulk structures as substitutional species. This contrast with the observations on structured (mesoporous) samples where nitrogen remains associated with oxygen as NO radicals located on the pore surfaces. As a matter of fact, doping with nitrogen is inefficient in mesoporous structures due to the tendency of doping ions to migrate at the outermost pore surfaces.

The applications of titanium dioxide-based structures in photovoltaic devices or for visible light-driven photocatalytic reactions require appropriate electronic structure and band gap leading to maximum absorption of the solar spectrum. Intrinsically, TiO₂ possesses wide band gap (3.1–3.2 eV) and only UV radiation can contribute to create photoinduced charge transfers from valance to conduction band. The electronic structure of TiO₂ crystalline structure is characterized by O-2p orbitals which form the valence band, while Ti³⁺ (3d¹) states form the conduction band. In nitrogen-doped samples, it was shown from former work (Melhem et al. 2013) that reduction of the band gap occurs and a spectral extension of the

absorption of solar radiation. Quantitative analysis is outlined in the theoretical approach developed below and supported by numerical simulations.

BiVO₄ Nanostructures for Photocatalysis

The semiconducting oxide BiVO₄ Bismuth vanadate (BiVO₄) is currently a subject of intensive investigations devoted to create several morphologies such as nanoparticles, thin films, or composites associating metallic clusters. As thin films, BiVO₄ represents an ideal photo-anode with a high quantum efficiency in the decomposition of water and hydrogen production (Kim and Choi 2014). The application as photocatalyst is also a very active research area, for example, in water depollution from organic contaminants (Wang et al. 2012). The realization of nanostructures with large specific surfaces by using suitable deposition conditions of *rf*-sputtered films or by the growth of mesoporous architectures or nanoparticles are well-improved approaches. The synthesis methods and annealing modify in a large extent the structural and morphological characteristics of the BiVO₄ (Venkatesan et al. 2012). The physical properties can be modulated by several procedures to enhance the efficiency of the reaction mechanisms in visible light-driven photocatalytic processes or for photovoltaic conversion based on the photoinduced charge transfers (Zhang et al. 2007; Yu and Kudo 2006; Ng et al. 2010). Indeed, the efficiency of the phenomena depends on the charge densities involved in the semiconducting structure, on the optical band gap, as well as on the lifetimes of photogenerated excitons. However, the intrinsic properties of BiVO₄ give rise to limited performances as, for example, in the photodegradation of organic molecules under visible light. A challenging strategy consists of doping the considered structures by suitable elements such as metallic ions incorporated in the structure or clusters contributing to realize composites with improved photocatalytic properties. Suitable doping was realized by using Cu, Mo, or other ions (Merupo et al. 2015a, b) as substitutional ions or using silver nanoparticles associated with BiVO₄ to enhance visible light photocatalytic degradation of organic dyes in aqueous solutions. Other complex BiVO₄-based materials were also realized to enhance the photoactivity by increasing the rates of photogenerated carriers as well as their lifetimes required for photocatalysis. Thus, nanostructured-conducting media such as graphene layers have been associated with BiVO₄ to limit the rate of recombination of photogenerated electron-hole pairs in water dissociation (Ng et al. 2010) or for degradation of organic dyes (Gawande and Thakare 2012). The association of TiO₂ and BiVO₄ with particular bridging agents was also investigated leading to remarkable photoactivity (Xie et al. 2015). In similar approaches, nanocomposites were also synthesized based on other oxides (Bajaj et al. 2013; Madhusudan et al. 2011).

The relevant methodology ensuring the best composition and features of the efficient photocatalysts based on BiVO₄ associates exhaustive experimental contribution for the synthesis and characterizations and theoretical approaches with numerical simulations to validate and analyze quantitatively the phenomena. The

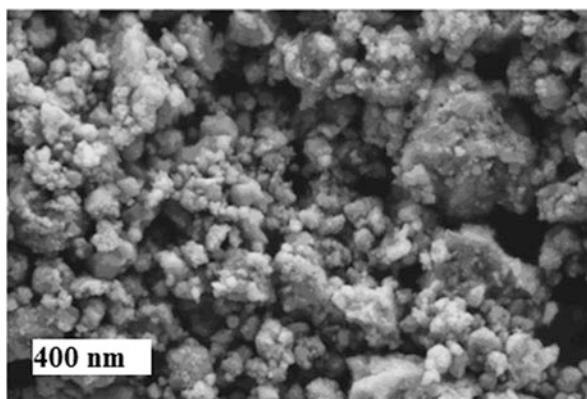
developed numerical approaches were partially conducted in our former works devoted to functional semiconducting systems (Makowska-Janusik and Kassiba 2012). The methodology proceeds by modeling at macroscale the structural, vibrational, and electronic properties of bulk crystal. Prediction and quantitative description are achieved for the key physical properties of BiVO_4 . In the second step, numerical approaches based on cluster model and quantum-chemical calculations were implemented to explain the electronic properties including quantum confinement effects and related optical changes of nanoparticles and elucidate the influence of surface effects on photocatalytic-driven mechanism such as charge transfers.

Structural Properties of BiVO_4

BiVO_4 crystallizes in three main polytypes such as tetragonal zircon, monoclinic scheelite, and tetragonal scheelite structures. For applications in photocatalysis, it was shown that the monoclinic structure offers the best photocatalytic efficiency (Merupo et al. 2015b; Venkatesan 2014). On the other hand, the specific surface of nanomaterials is an important parameter which acts on the reaction area at the interfaces between the photocatalysts and its reactive environment.

Suitable methods for the synthesis of BiVO_4 were developed including ball milling or hydrothermal for micro- and nanosized particles as well as *rf*-sputtering for thin films with nanostructured surface profiles. For ball milling technique, mechanochemical processes were carried out with initial pure reactants Bi_2O_3 and V_2O_5 . Suitable operating conditions were chosen (time of reactions, speed of rotating reaction chambers, nature of the used grinding balls) to obtain homogeneous powders with defined stoichiometry and stable crystalline structure (Merupo et al. 2015b). An example of powder obtained by mechanochemical process during short milling time (6 h) is shown in Fig. 4 by field effect scanning electron microscopy (FE-SEM) (Venkatesan 2014). The microstructure exhibits individual aggregates made from small nanoparticles with sizes in the order of 20 nm.

Fig. 4 FE-SEM image of BiVO_4 powders obtained by ball milling



Structural investigations were carried out to identify the involved crystalline phase as a monoclinic scheelite in agreement with lattice parameters $a = 5.1950 \text{ \AA}$, $b = 11.701 \text{ \AA}$, and $c = 5.092 \text{ \AA}$ (JCPDS no. 14-0688) (Fig. 5). The vibrational properties give complementary information due to their sensitivity to local distortions of chemical bonding involved in defined environments of the materials. The investigations by Raman spectrometry determine the vibrational modes allowed by the symmetry of the crystal structure. The analysis of the corresponding Raman bands including their full width at half maximum (FWHM) or the shift of their wave number position are related to the quality of the crystalline order of the atomic environments submitted to stretching or bending. Thus, Raman active modes (Fig. 6) of monoclinic *clinobisvanite* bismuth vanadate are involved at wave number positions at $211, 327, 369, 710,$ and 831 cm^{-1} (Venkatesan et al. 2012). These active

Fig. 5 XRD pattern of ball-milled BiVO_4 , its refinement by Fullprof software is in agreement with the monoclinic scheelite structure according to the diffraction data reference (JCPDS no. 14-0688)

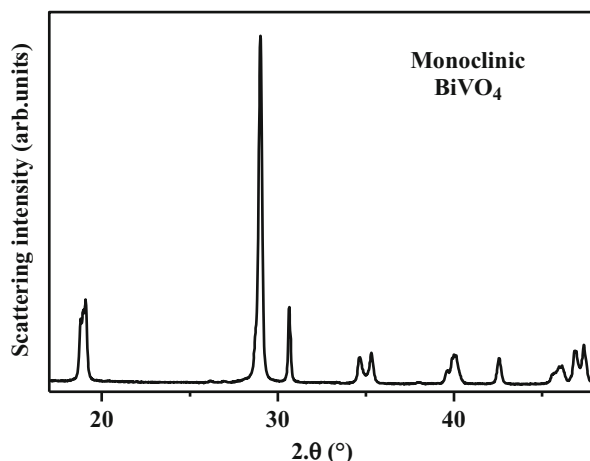
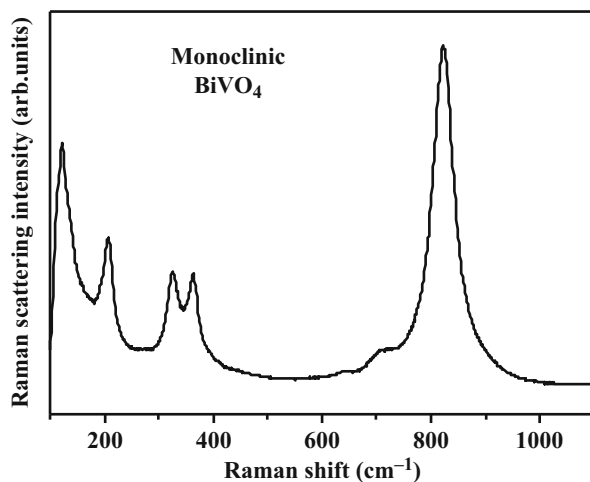


Fig. 6 Raman spectrum of BiVO_4 obtained by mechanochemical method

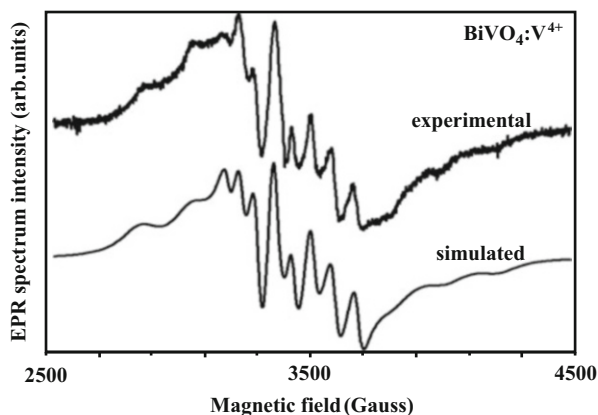


modes are associated with stretching and bending modes of VO_4 groups involved in the unit cell of the monoclinic BiVO_4 structure. The intense band at 831 and the weak shoulder at 710 cm^{-1} are assigned to the symmetric A_g and asymmetric stretching modes of VO_4 groups. The modes at 369 and 327 cm^{-1} correspond to the symmetric and antisymmetric bending modes for the vanadium oxide tetrahedrons. Such modes are sensitive to bond angles and length distortions within the VO_4 groups. The low wave number mode at 211 cm^{-1} is assigned to an external mode associated with Bi-VO_4 .

Electronic Active Centers of BiVO_4

The existence of high specific surfaces in the ball-milled synthesized powders and crystalline defects such as oxygen vacancies favor the appearance of active electronic centers. F -centers from electrons trapped in the vacancy sites or valence change of vanadium ions as V^{4+} ($3d^1$) play a crucial role on the photoinduced charge transfer process. Both centers are characterized by an effective electron spin $S = 1/2$ leading to an EPR signal which can be used to evaluate the nature of the center, its local environment, thermal stability, and behavior under light irradiation. For the ball-milled sample, the characteristic experimental EPR spectrum is reported in Fig. 7. The simulation of the experimental spectrum is made by superimposition of an EPR signal from isolated V^{4+} ions in orthorhombic local environments, while a second signal is related to vanadium clustering. These results are inferred from the magnetic interactions between close V^{4+} ions which alter the EPR line widths (FWHM) and average the anisotropy of the environments. Vanadium pairs or clustering of V^{4+} ions accounts for the EPR signal with the features of a broad background absorption line. The structured details observed on the EPR signal such as several sharp lines are emanating from isolated ions with anisotropic local environments (Venkatesan et al. 2013). In one hand, magnetic interactions are characterized by g -tensor reflecting an orthorhombic local symmetry around V^{4+} . This is understood by the occurrence of similar electronic structure and local environment of V^{4+} in the sample BiVO_4 irrespective to the sample microstructures.

Fig. 7 EPR spectra (experiment and simulation) related to isolated V^{4+} ions due to charge compensation from oxygen vacancies and from clustering of vanadium ions on the surfaces of BiVO_4 particles (Venkatesan et al. 2013)



The location of V^{4+} within crystal sites of the particle cores accounts for the experimental observations. On the other hand, the clustering of V^{4+} is probably involved at the outermost particle surfaces.

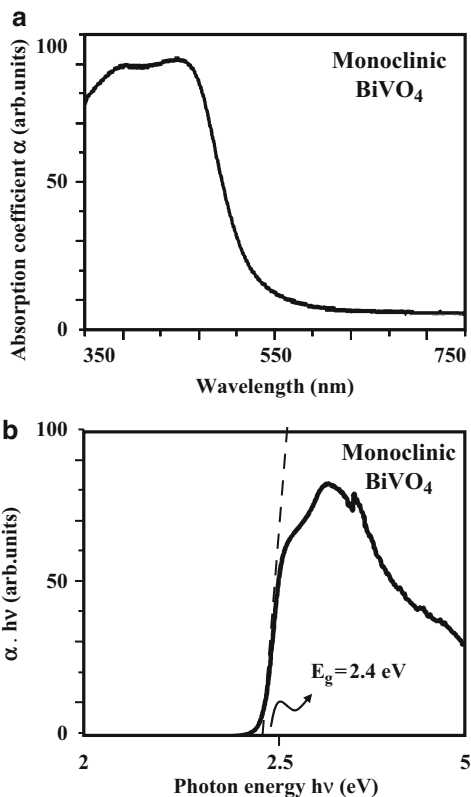
The exhaustive analysis of the magnetic interactions has been carried out in the former report (Venkatesan et al. 2013) and points out the high probability of occupation of excited orbitals of vanadium ions. The photoinduced charge transfer under light irradiation can be then favored and contribute to the efficiency of photocatalytic reactions which exploit $BiVO_4$ with such active electronic centers. Additional investigations have been carried out in order to characterize the evolution of the electrical conductivity in $BiVO_4$ samples with different microstructures as function of V^{4+} concentration which is found in the range 10^{16} - 10^{17} ion/gram. As the molecular bonding $V^{5+}-O^{2-}$ accounts for the charge carrier motions within the crystalline structure, the presence of high or low concentration of V^{4+} did not alter the effective conductivity. This fact excludes the possibility for such electronic active centers to act as trap for charge carriers. The photoinduced processes are expected to be enhanced in samples with higher concentration of V^{4+} ions.

Optical Behavior of $BiVO_4$

The optical absorption spectrum informs on the electronic structure and can be used to evaluate the band gap of materials. Previous theoretical investigations were developed on $BiVO_4$ which possesses the features of direct band gap with the extreme of the valence (VB) and conduction band (CB) away from the center of the Brillouin zone (Walsh et al. 2009). Indeed, electronic band structure calculations demonstrate the occurrence of direct transitions between VB and CB at the Brillouin zone boarder, i.e., at the Z point associated with (0.000, -0.500, 0.500) point. The electronic states which compose the valence band (VB) and the conduction band (CB) were identified. Thus, the contributing orbitals to VB consist in interacting of Bi (6 s) and O (2p) states. For the CB, the main contribution occurs from V (3d) orbitals which can interact with Bi (6p) and O (2p). The optical band gap is then defined by transitions between (VB) and (CB) with finite probability at the Z point.

The experimental investigations are traduced in Fig. 8 which reports the UV-vis absorption spectrum of $BiVO_4$ versus the photon energy. The photon absorption range indicates the possibility of investigated structure to act as visible light-driven photocatalyst. In order to evaluate precisely the electronic band gap, the so-called Tauc plot is usually exploited through the equation $(\alpha \cdot h\nu)^2 = C \cdot (h\nu - E_g)$ where α , C , and E_g represent, respectively, the absorption coefficient, a proportionality constant, and the direct band gap energy. The asymptotic extrapolation of the plot with the photon energy ($h\nu$) axis gives the direct band E_g with a good approximation. By this representation we evaluate the band gap of the considered ball-milled powder with the monoclinic sheelite structure of $BiVO_4$ to $E_g = 2.40(5)$ eV. It is worth to notice that the optical behavior depends on the microstructure of the material, on its crystalline order, and on the eventual electronic active centers as discussed above notably for V^{4+} . Annealing treatment of the powders at temperatures up to 500 °C is relevant. As an example, as-formed milled powder and

Fig. 8 (Top) Experimental UV–vis optical absorption spectrum of ball-milled powders of BiVO_4 . (Bottom) Tauc plot allows determining precisely the direct band gap of the sample



that obtained after annealing show different absorption edges of the BiVO_4 (Fig. 9) with a net tendency to lower the band gap for annealed samples. These experimental results will be discussed based on the theoretical calculations developed for the mentioned materials.

Theoretical Background

Modeling of Crystalline Structures: Preamble

Theoretical models and numerical simulation programs were extensively developed in the last decades with the aim to describe the physical properties of semiconducting materials and to offer qualitative and quantitative supports for experiments. These approaches are used to predict physical properties (electronic, optic, magnetic, etc.) in a wide class of functional materials. In this context, electronic band structures have been evaluated, and several models were developed to ensure the calculations within reasonable time duration of the required computational

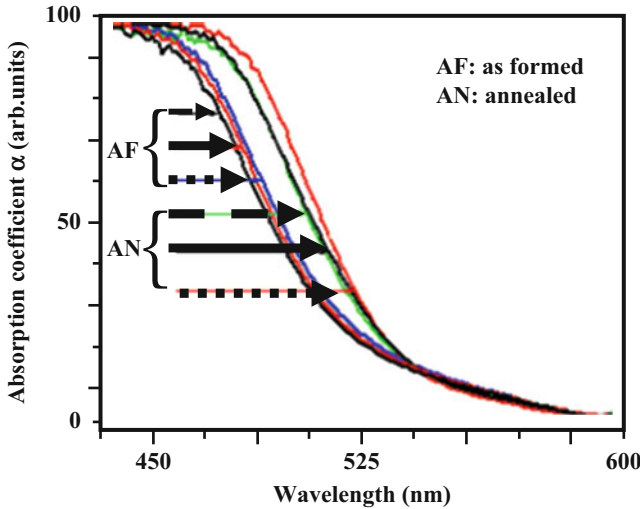


Fig. 9 Optical absorbance spectra of ball-milled BiVO₄ with different milling times modifying the microstructure (Venkatesan 2014) and after post-annealing treatment (A). Annealing contributes to widen the absorption band to higher wavelengths (lower band gap)

operations. In all situations, the starting point deals with the Hamiltonian describing mutual interactions between electrons and nuclei and summarized as follows:

$$\begin{aligned}
 H = & \sum_i \frac{p_i^2}{2m_i} + \sum_j \frac{p_j^2}{2m_j} + \frac{1}{2} \sum_{i,i'} \frac{e^2}{4\pi\epsilon_0 |r_i - r_{i'}|} - \\
 & - \sum_{i,j} \frac{Z_j e^2}{4\pi\epsilon_0 |r_i - R_j|} + \frac{1}{2} \sum_{j,j'} \frac{Z_j Z_{j'} e^2}{4\pi\epsilon_0 |R_j - R_{j'}|}
 \end{aligned}
 \tag{1}$$

where r_i and $r_{i'}$ indicate the position of the i -th and i' -th electron, R_j and $R_{j'}$ are the positions of the nucleus, Z_j denotes the atomic number of the nucleus, and p_i , p_j are, respectively, the momentum operators of the electrons and nuclei. However, as written, the eigenstates and eigenvalues of such Hamiltonian cannot be found simply and same approximations must be introduced to solve the problem. One of the approaches consists of separation of the electrons into the core and valence parts with the possibility to frozen the first one.

The frozen-core approximation is thus fundamental in the pseudopotential approach. In this frame, the introduced pseudopotential reproduces the true potential outside the core region (valence electrons) while smoothing of such potential for the core electrons inside some defined radius and contributing then to eliminate the oscillations of the electron wave function inside the core region. According to norm-conserving condition, the pseudopotential $V_i(r)$ can be composed by a long-range local part (V^{LR}) and a short-range semi-local one (V^{SR}) as follows:

$$V^{NC-PP} = V^{LR} + V^{SR},
 \tag{2}$$

where $V^{LR} = -(Z_v e^2)/r$ is a radial function, and the semi-local part is the deviation from the all-electron potential inside the core region. In the same context, the ultrasoft pseudopotential approximation (Vanderbilt 1990) uses more than one projector for each momentum and further smooth the electron wave function.

Within the Born–Oppenheimer approximation and for crystalline structure, the mean-field approximation should be implemented to solve the Schrodinger equation. In such approach, each electron feels the same average potential $V(r)$ and its Schrodinger equation takes the form:

$$\left(\frac{p^2}{2m} + V(r) \right) \Phi_N(r) = E_N \Phi_N(r), \quad (3)$$

where $\frac{p^2}{2m} + V(r)$ denotes one-electron Hamiltonian (H_{eI}) identical for all valence electrons. $V(r)$ possesses the symmetry of the lattice and contains both electron–electron and core–electron interactions. Additionally, the potential $V(r) = V(r + R_n)$ has periodicity of the lattice, where R is the lattice vector. Bloch theorem states that the eigenstates of a periodic Hamiltonian can be written as a product of periodic function with a plane wave of momentum k restricted to the first Brillouin zone (BZ).

The quantum-chemical calculations of the crystal structure are based on a supercell method. The wave function in periodic potential named as Bloch function is defined as (Payne et al. 1992):

$$\Phi_k(r) = \exp(i\mathbf{k} \cdot \mathbf{r}) u_k(r), \quad (4)$$

where $u_k(r) = u_k(r + R)$ is a periodic function with the translational period equal to R as it was mentioned for the potential $V(r)$. As defined, the plane wave character of the wave function is seen with an amplitude modulated by the periodic function $u_k(r)$. Therefore, the wave function satisfies:

$$\Phi_k(r + R) = \exp(i\mathbf{k} \cdot R) \Phi_k(r) \quad (5)$$

Since the $u_k(r)$ has the periodicity of the lattice, it can be expanded in Fourier series:

$$u_k(r) = \sum_j A_j(j) \exp(iG_j \cdot r), \quad (6)$$

where A_j is the minimum translation vector and G_j is the reciprocal lattice vector.

When translation parameter T_R acts at the Bloch function:

$$T_R \Phi_k(r) = \Phi_k(r + R) = \exp(i\mathbf{k}R) \Phi_k(r) \quad (7)$$

This is an eigenfunction of the T_R with its eigenvalue equal to $\exp(i\mathbf{k}R)$. According to the commutator rules, the eigenfunction of the one-electron Hamiltonian (Eq. 3) may be defined as the sum of Bloch functions:

$$\Phi(x) = \sum_k A_k \Phi_k(x) = \sum_k A_k \exp(i\mathbf{k}x) u_k(x) \quad (8)$$

Thus the one-electron wave functions can be indexed by k , which are the wave vectors of the plane waves forming the “backbone” of the Bloch function. A plot of the electron energies in (Eq. 3) versus k is known as the electronic band structure of the crystal.

The electronic states for the crystal structure are calculated over the numbers of k -points. The symmetry considerations suggest that only k -points within the irreducible segment of the Brillouin zone should be taken into account (Srivastava and Weaire 1987; Payne et al. 1992). Using the mentioned method, it can be obtained an accurate approximation of the electronic potential and the total energy of an insulator by calculating electronic states at restricted number of k -points. The calculations for metallic systems require a more dense set of k -points to determine the Fermi level accurately. One of the most popular schemes for generating k -points proposed by Monkhorst and Pack (1976, 1977) produces a uniform grid along the three axes in reciprocal space. They defined a uniform set of special k -points as:

$$k = u_p b_1 + u_r b_2 + u_s b_3, \quad (9)$$

where b_1 , b_2 , and b_3 are the reciprocal lattice vectors, and u_p , u_r , and u_s are numbers from the sequence:

$$\begin{aligned} u_r &= (2r - q - 1) / 2q \\ r &= 1, 2, 3, \dots, q \end{aligned} \quad (10)$$

The total number of k -points is q^3 , but the number of the irreducible k -points can be drastically decreased by symmetry.

DFT Methodology and Related Functionals

Another approach widely developed to compute the electronic structure of materials consists of the density functional theory (DFT) used in wide variety of approaches (Bermudez 2010; Na-Phattalung et al. 2006; Le et al. 2005; Chretien and Metiu 2011; Martinsovich et al. 2010; Persson et al. 2003; Albaret et al. 2000; Walsh et al. 1999; Qu and Kroes 2006, 2007; Gałyńska and Persson 2013; Sahoo et al. 2011; Barnard and Zapol 2004a, b; Blagojevic et al. 2009; Hamad et al. 2005; Albuquerque et al. 2012; Zhang et al. 2005 ; Labat et al. 2007). Initially, DFT was developed by Hohenberg and Kohn (Hohenberg and Kohn 1964) and then by Levy (1979). They define the charge density ρ as the key parameter in energy functions instead of the electron wave functions. Thus, the total energy is composed of the kinetic energy $T[\rho]$ of a system of noninteracting particles superimposed to the electrostatic part due to Coulomb interactions $U[\rho]$ and to the many-body interaction $E_{xc}[\rho]$ called as the exchange and correlation energy. To achieve computational tasks within reasonable time delays, the exchange-correlation energy requires some approximation. In this aim, DFT methodology was implemented

within two different approximations, namely, local (LDA) and gradient-corrected (GGA) one.

The LDA approximation is based on the exchange-correlation energy of uniform electron gas (Hedin and Lundqvist 1971; Ceperley and Alder 1980). The local density approximation assumes that the charge density varies slowly on an atomic scale. The total exchange-correlation energy can be obtained by integrating over the uniform electron gas:

$$E_{XC}^{LDA}[\rho] \cong \int \rho(r) \varepsilon_{XC}[\rho(r)] dr, \quad (11)$$

where $\varepsilon_{xc}[\rho]$ represents the exchange-correlation energy per particle in the considered uniform electron gas. Slater proposed that the exchange part of the mentioned potential may be defined as equal to $\varepsilon_{XC}=\rho^{1/3}$ (Slater 1951) where in fact the correlation is not included. Generally, the exchange-correlation energy is described by the approximation implemented by Vosko, Wilk, and Nusair (VWN) (Vosko et al. 1980) or by Perdew and Zunger (PW) (Perdew and Zunger 1981).

To investigate the character of the non-homogeneity of the electron gas, the LDA approximation is not sufficient. This phenomena is taken into account by the density-gradient expansion referred to the nonlocal spin-density approximation (NLS) based on the GGA functional and depends on the derivative $d\rho/dr$ as well as on ρ . This provides a considerable increase in the accuracy of predicted energies and structures, but with an additional computational time delay. The mentioned methodology was implemented by Perdew and Wang in 1991 (Perdew et al. 1992) with the energy defined as:

$$E_{XC}^{GGA}[\rho] \cong \int \rho(r) \varepsilon_{XC}[\rho(r), \nabla\rho(r)] dr \quad (12)$$

One of the GGA functionals is the Perdew–Burke–Ernzerhof (PBE) one (Perdew et al. 1996), which is used in several problems related to material sciences. However the BLYP functional (Becke 1988) is exploited to evaluate accurately the energy parameters of organic groups leading to more relevant results than the PBE model. The situation is quite different for the correlation energy related to metallic systems where the BLYP potential gives unsatisfactory results.

So far, DFT methodology contributes to precise insights on the electronic properties of different class of materials (Hohenberg and Kohn 1964). This seems related to the development of new and more accurate approaches through the used functionals as well as the versatility of numerical methods with regard to different classes of computed systems. As a major drawback of all functionals, the underestimation of the calculated electronic band gap (Hybertsen and Louie 1989) for semiconductors and insulators is frequently encountered due to the large uncertainty in the calculations of the electron self-interaction energies. The pure LDA (Kohn and Sham 1965) or the GGA approximation (Langreth and Perdew 1980) or even the hybrid functional (Stephens et al. 1994; Perdew et al. 1996;

Heyd et al. 2003) may suffer from lack of precision for the so called “strongly correlated” systems.

Hybrid functionals were introduced by Becke (1993) replacing a fraction of GGA exchange with HF exchange. Such an approach represents a judicious compromise between Hartree–Fock exchange and DFT exchange and correlation part. The hybrid version of PBE and BLYP was proposed as PBE0 (Perdew et al. 1996), HSE (Krukau et al. 2006), and B3LYP (Becke 1993). B3LYP is currently the most popular functional used for organic groups. The formulation is based on Becke 88 exchange functional and LYP correlation functional (Becke, three parameter, Lee–Yang–Parr):

$$E_{XC}^{B3LYP} = E_{XC}^{LDA} + a_0 (E_X^{HF} - E_X^{LDA}) + a_x (E_X^{B88} - E_X^{LDA}) + a_c (E_C^{LYP} - E_C^{VWN}), \quad (13)$$

where $a_0 = 0.20$, $a_x = 0.72$, and $a_c = 0.81$.

The PBE0 functional can be traduced as follows:

$$E_{XC}^{PBE0} = 0.25E_X^{HF} + 0.75E_X^{PBE} + E_C^{PBE} \quad (14)$$

and HSE uses an error function screened Coulomb potential to calculate the exchange part of the energy in order to improve the computational evaluation, especially for metallic systems:

$$E_{XC}^{HSE} = 0.25E_X^{SR}(\mu) + 0.75E_X^{PBE,SR}(\mu) + E_X^{PBE,LR}(\mu) + E_C^{PBE}. \quad (15)$$

The separation of the electron–electron interaction into a short- and long-ranged part, labeled SR and LR respectively, is realized only in the exchange interactions.

It is generally known that classical DFT potentials do not reproduce the correct far-nucleus asymptotic behavior (Almbladh and Pedroza 1984; Levy et al. 1984) and underestimate the excitation energies notably for charge transfer (Dreuw et al. 2003) as also noticed for calculation based on time-dependent DFT (TDDFT) functional (Appel et al. 2003; Dreuw et al. 2003). It is presumed that classical functionals overestimate local contributions and underestimate nonlocal ones. The long-range electron–electron exchange interaction is neglected in pure functionals and is supposed that the appropriate inclusion of the mentioned interaction into DFT calculations improves the obtained results (Tawada et al. 2004). Therefore, the potential solution leads to the partitioning of the total exchange energy into short-range and long-range contributions:

$$E_X = E_X^{sr} + E_X^{lr} \quad (16)$$

Such methodology is called long-range correction (LC) scheme (Iikura et al. 2001). It uses the standard error function to modify and improve the exchange functional for evaluating the long-range electron–electron interaction by the HF exchange

integral. In LC approach, the repulsion electron operator is also divided into short-range and long-range part and can be defined as:

$$\frac{1}{r_{ij}} = \frac{1 - [\alpha + \beta \operatorname{erf}(\mu r_{ij})]}{r_{ij}} + \frac{\alpha + \beta \operatorname{erf}(\mu r_{ij})}{r_{ij}}, \quad (17)$$

where $r_{ij} = |r_i - r_j|$ is the distance between two electrons located in position r_i and r_j . The parameter μ represents the weighting factor which controls the separation between short-range exchange functional and the long-range part of HF exchange integral. The optimal μ factor corresponds to the atomic bond length and hence is expected to be dissimilar for different atoms. When μ is equal to zero, the LC DFT calculation corresponds to pure DFT methodology, and conversely $\mu = \infty$ corresponds to the standard HF calculation. The α and β parameters define the exact exchange percentage between short- and long-range exchange functionals. Since $\alpha + \beta = 1$, the exchange potential in LC functionals has the exact asymptotic behavior. Also, the parameter α determines the contribution of the HF exchange at the short-range region. At a distance of about $2/\mu$ the short-range interactions become negligible. The LC-BLYP method uses $\alpha = 0$ and $\beta = 1$.

The CAM-B3LYP functional (Yanai et al. 2004) (Coulomb-attenuating method applied to B3LYP) also eliminates the long-range self-interaction error. In this case, the parameters are assigned to $\alpha = 0.19$ and $\beta = 0.46$, and hence the exact asymptote of the exchange potential is lost. To sum up, the parameter α allows incorporating the HF exchange contribution over the whole range of distance. Similarly, the parameter β defines the DFT counterpart over the whole range by a factor of $1 - (\alpha + \beta)$.

Another major problem of standard LDAs and GGAs is that they could not describe the system when the electrons tend to be localized and strongly interacting, such as for transition metal oxides or rare earth elements and compounds. One of the successful models to describe correlated electrons in solids is the Hubbard model (Anisimov et al. 1991) based on extended LDA approach to what is termed LDA+U. It decreases the electron self-interaction error by selectively adding an energy correction to localized electron states as for d or f orbitals, where the self-interaction is particularly large. The additional interactions is usually considered only for high localized atomic-like orbitals on the same site, as the ‘‘U’’ interactions in Hubbard models:

$$E^U = \frac{1}{2}U \sum_{i \neq j} n_i n_j. \quad (18)$$

In the multiband Hubbard model, the effective LDA+U energy functional is written as:

$$E^{LDA+U} = E^{LDA} + E^U + E^{dc}, \quad (19)$$

where E^{LDA} denotes standard LDA energy functional and the $E^{dc} = UN(N-1)/2$ is called the double-counting term, to remove the energy

contribution of these orbitals which are twice introduced by the Hubbard term ($N = \sum_i n_i$).

In the LDA + U approach, the Kohn–Sham equation is supplemented by the nonlocal potential. From the other side, the DFT/Hubbard method fails to compute the correct energy difference in systems with localized/correlated and delocalized/uncorrelated electronic states (Jain et al. 2011). Also, the LDA/GGA + U methodology was successfully applied to compute the electronic properties of different ternary oxides such as CuAlO₂ (Laskowski et al. 2009), CuAl₂O₄ (Liu and Liu 2011), Pr₂Ti₂O₇ or Ce₂Ti₂O₇ (Sayede et al. 2013), and MnFe₂O₄ (Huang and Cheng 2013). Based on the outlined theoretical framework and as discussed below, the influence of the Hubbard parameter U on the electronic properties of NiTiO₃ was investigated on the crystalline structure in bulk material or in nanosized clusters taking into account the strong correlation of the d-orbital electrons.

Cluster Approach to Model Nanostructures

The study of nanosized systems is of great interest with regard to the critical role of the specific active surface in physical phenomena. Particularly, interfaces between the active material and the surrounding media contribute critically to the efficiency of heterogeneous catalysis. In this context, theoretical modeling and numerical simulations to predict electronic, optical, structural, vibrational, and related properties of nanoparticles are challenging tasks. Among the numerical approaches for prediction of the physical properties of nanoscale system, the cluster approach is worthy of interest with respect to combined aspects from molecular physics methodology as well as from that applied to macroscopic systems. Such approach requires DFT methodology supported by the semiempirical PM6 and PM7 parameterized approach due to the cluster size.

Some of quantum-chemical methodologies use bulk crystal approach to predict the electronic properties of nanosized materials by building various surfaces and by reducing the periodicity of the crystal lattice (Martsinovich et al. 2010). Other methods proceed by full optimization of atomic positions of nano-clusters without defined crystal structure (Qu and Kroes 2006; Galyńska and Persson 2013). In the present work, the cluster approach is developed within a methodology based on crystalline structure inside the core of the investigated nanoparticles surrounded by a reconstructed surface according to surface energy requirement and interactions with the environment. The nanostructures are built from their native bulk crystals with specified atomic positions as involved in defined space groups. Modeling of bulk systems is of primary importance to settle the methodology for the structural, vibrational, and electronic properties of nanostructures.

In order to simulate realistic structures able to traduce experimental data, the clusters should be constructed with the appropriate size (Makowska-Janusik and Kassiba 2012). However, experimental observations are usually made on nanoparti-

cles with diameters in the range of a few nanometers, but numerical modeling based on quantum-chemical methods are highly time consuming and cannot be applied on such large systems. In this context, semiempirical approaches are more relevant, and the parameterized PM6 or PM7 methods are preferred because their bases include parameterized data for frequently studied atoms. The following section reports selected semiconducting nanostructures which were modeled based on the above cited approach.

Structural Features of TiO_2 , NiTiO_3 , and BiVO_4 Crystals

The crystal structures of investigated semiconducting oxides were computed for two different configurations of the material. One of them was considered like bulk monocrystal and the second one was created in nanosized crystal future. In this case the TiO_2 , NiTiO_3 , NiTiO_3 and BiVO_4 crystals were chosen. To build all structures, the Materials Studio program package was used. The structural parameters of all investigated crystals are presented in Table 1. These data are in agreement with those from experimental and other theoretical works (Burdett et al. 1987; Sleight et al. 1979; Zhao et al. 2011). The example of semiconducting NiTiO_3 crystal unit cell and its nanostructure are shown in Fig. 10.

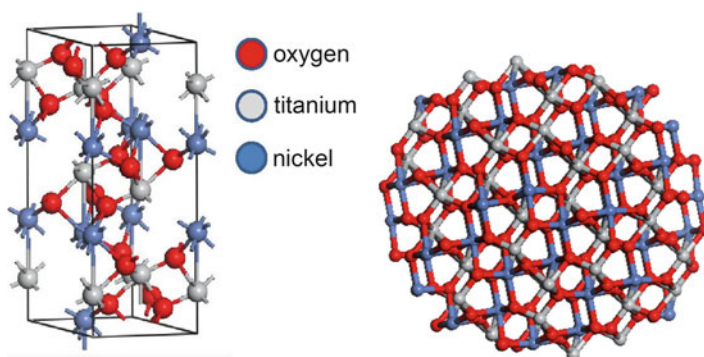
Computational Parameters

The electronic properties of all investigated single crystal were calculated using DFT method. The quantum-chemical calculations were performed using Cambridge Serial Total Energy Package (CASTEP), the module of Materials Studio program. The CASTEP is based on total energy plane wave pseudopotential method. In the first step, the geometries of the investigated structures were optimized with respect to the total energy minimization within the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm (Pfrommer et al. 1997). The symmetries of all structures were frozen during the geometry optimization procedure, but in some cases the size of the unit cell was allowed to be modified. The convergence parameters were chosen for the energy as low as 2×10^{-5} eV/atom, the force on the atom was less than 0.01 eV/Å, the stress on the atom limited below 0.02 GPa, and the atomic displacement less than 5×10^{-4} Å.

The electronic exchange-correlation energy was treated within the frame of the generalized gradient approximation (GGA) using Perdew–Burke–Ernzerhof (PBE) (Perdew et al. 1996) potential. The calculations were performed with the cutoff energy of the plane wave basis set equal to 350 eV for BiVO_4 and 500 eV for TiO_2 and NiTiO_3 crystals. To accelerate the computational runs, the functional such as ultrasoft pseudopotential was used for the calculations related to the following

Table 1 Structural parameters of investigated crystals: TiO₂ in anatase phase, NiTiO₃ in ilmenite phase, and the BiVO₄ in monoclinic scheelite structure

TiO ₂ anatase phase (Burdett et al. 1987)				
Space group	I4 ₁ /amd, local symmetry			
Cell parameters	a = b = 3.782 Å, c = 9.502 Å, γ = 90°			
Atom position	Atom name	x/a	y/b	z/c
	Ti ₁	0	0	0
	Ti ₂	0	0.5	0.25
	O ₁	0	0	0.2066
	O ₂	0	0	-0.2066
	O ₃	0	0.5	0.4566
	O ₄	0	0.5	0.0434
NiTiO ₃ – ilmenite phase				
Space group	(No. 148)			
Cell parameters	a = b = 5.029 Å c = 13.795 Å α = β = 90° γ = 120°			
Atom position	Atom name	x/a	y/b	z/c
	Ni	0	0	0.3499
	Ti	0	0	0.1441
	O	0.3263	0.0214	0.2430
BiVO ₄ – monoclinic scheelite phase (Sleight et al. 1979; Zhao et al. 2011)				
Space group	C2/c space group and its point group is			
Cell parameters	a = 7.247 Å b = 11.697 Å c = 5.090 Å γ = 134.226°			
Atom position	Atom name	x/a	y/b	z/c
	Bi	0	0.134	0.750
	V	0	0.370	0.250
	O ₁	0.261	0.051	0.380
	O ₂	0.354	0.208	0.861

**Fig. 10** Unit cell of NiTiO₃ crystal (*left*) and the unpassivated nanocrystal made by atomic group (NiTiO₃)₈₂ (*right*) (Ruiz Preciado et al. 2015)

electronic configuration: Ti $3s^2 3p^6 3d^2 4s^2$, Ni $3d^8 4s^2$, Bi $6s^2 6p^3$, V $3s^2 3p^6 3d^3 4s^2$, and O $2s^2 2p^4$. The numerical sampling integration over the Brillouin zone was carried out using the Monkhorst–Pack method with a $8 \times 8 \times 8$ and $5 \times 5 \times 7$ special k -point mesh for $(\text{TiO}_2, \text{NiTiO}_3)$ and BiVO_4 , respectively. The total energy was assumed to converge when the self-consistent field (SCF) tolerance is equal to 10^{-5} eV/atom.

The electronic properties were computed using the same parameters as used during the geometry optimization procedure. The calculations were performed in spin restricted as well as spin unrestricted procedures applying GGA/PBE potential. For the TiO_2 and NiTiO_3 crystals, the Kohn–Sham equation was also solved using the GGA/PBE functional with the Hubbard parameters.

As for bulk systems, Materials Studio software was also used to build the nanostructures. The clusters possess spherical shape and were realized by limiting defined volume within the bulk crystal. According to stoichiometry, the clusters were built as n -multiple of the (TiO_2) , (NiTiO_3) , or (BiVO_4) units with sizes ranging from 0.7 nm up to 2.6 nm. For NiTiO_3 , the clusters consist of $(\text{NiTiO}_3)_2$ up to $(\text{NiTiO}_3)_{183}$ units. The atomic positions within the cluster were frozen to the ones related to bulk crystal. All clusters were modeled without any external saturation. Only the $(\text{TiO}_2)_n$ clusters were saturated by using hydrogen atoms or by hydroxyl ($-\text{OH}$) groups.

The electronic properties of clusters were computed by using the parameterized PM6 and PM7 method (Stewart 2007). This procedure allows pertinent evaluation of physical features of clusters with large sizes and with involvement of heavy atoms. The numerical simulations were then performed using MOPAC (Stewart 2012) and GAMESS (Schmidt et al. 1993) packages. All the calculations were carried out under restricted Hartree–Fock (RHF) regime with the SCF convergence criterion sets at 10^{-5} Hartree and no more than 500 iterations. The eigenvalues were calculated for the frozen atomic positions in the clusters with the molecular point group C1.

On the other hand, the DFT calculations were performed by using DMol3 software (Delley 1990, 2000) with implemented PBE and B3LYP functionals applying DND basis set. No symmetry rules were imposed during the electronic structure analysis and the SCF density convergence criterion was fixed to 10^{-5} eV. In order to reduce the required computational time, all core electrons were represented by effective core pseudopotentials (ECP) (Dolg et al. 1987; Bergner et al. 1993). The electronic properties of the $(\text{TiO}_2)_n$ clusters were also calculated applying DFT methodology implemented in GAMESS program package with the 6-31G basis set and the effective core potential named Stevens–Basch–Krauss–Jasien–Cundari (SBKJC) (Stevens et al. 1984; Cundari and Stevens 1993) required to limit the number of contributing electron wave functions. The calculations were performed with different exchange–correlation (XC) potentials in generalized gradient approximation. The BLYP and B3LYP potentials have been chosen, and the results were compared to different long-correlated hybrid and CAM-B3LYP functional.

Physical Properties of NiTiO₃ and BiVO₄ Crystals

Structural Aspects

Quantum-chemical calculations of all investigated crystals proceed first by optimizing their geometries using the GGA/PBE functional. The evaluated atomic distances were relaxed, but their changes remain limited below 5 % compared to experimental results reported elsewhere (Rajalingam et al. 2012). The small discrepancies could be attributed to the fact that the present calculations were made at $T = 0$ K, whereas the experimental measurements are generally recorded at finite temperature. However, the small deviations between theoretical and experimental values indicate that GGA/PBE is satisfactory as a computational method to describe the structures of all investigated crystals.

Following the geometry optimization, the investigated ilmenite NiTiO₃ structure exhibits layered organization (see Fig. 10) with the Ti and Ni atoms forming the layers separated by oxygen atoms. Figure 11 shows that for the optimized BiVO₄ crystal structure, the Bi–O and V–O bonds are more symmetric compared to the starting geometry. All Bi–O and V–O bonds have tendency to align at the same values. The geometry of monoclinic scheelite structure tends to the tetragonal geometry during the total energy minimization procedure. The possible reason of such isotropic bond lengths for Bi–O may be due to strong correlation between the Bi (6 s) and O (2p) orbitals (Ma and Wang 2014).

Electronic Properties of BiVO₄

The electronic parameters of the geometry optimized structure of BiVO₄ were calculated using DFT/PBE potential. The coordinates of the special points of the BZ

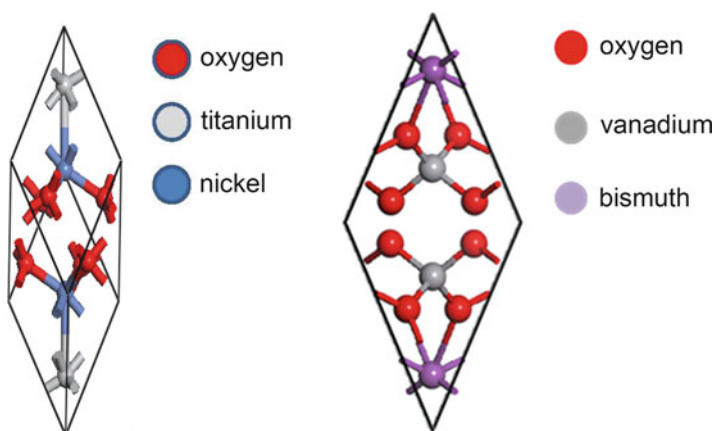


Fig. 11 Primitive unit cell of NiTiO₃ single crystal (*left panel*) and of the BiVO₄ single crystal (*right panel*)

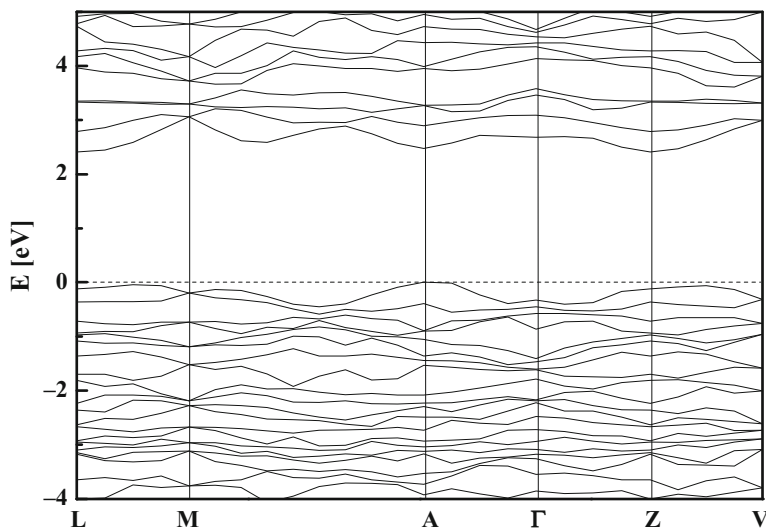


Fig. 12 Band structures calculated for frozen BiVO_4 monoclinic structure

are (in units of the reciprocal lattice vectors) L $(-1/2, 0, 1/2)$; M $(-1/2, 1/2, 1/2)$, A $(-1/2, 0, 0)$, Γ $(0, 0, 0)$, M $(0, -1/2, 1/2)$, and V $(0, 0, 1/2)$. The features of the obtained energy dispersion curves are similar to former results of Zhao et al. (Zhao et al. 2011), but the value of the band gap is significantly different. The minimum band gap obtained for such structure is equal to 2.25 eV compared to 2.17 eV reported in Zhao et al. study. The energy band calculations were also performed for BiVO_4 structure without any geometry optimization. The energy gap calculated for the frozen structure shows a departure of 0.16 eV higher than the band gap obtained by the same methodology for the relaxed geometry. Figure 12 reports the energy band structure calculated for the frozen BiVO_4 monoclinic structure where the band gap value is equal to 2.41 eV. The obtained energy band structure has the same shape for both relaxed and frozen structures. The valence band maximum (VBM) is located at the A point of BZ, while the minimum of conduction band (CBM) is realized at the Z point. The evaluated indirect band gap agrees with experimental values predicted by Rajalingam et al. (2012) and contrasts with former evaluation performed by Walsh et al. (2009) where they characterize the BiVO_4 crystal as direct semiconducting material.

The main contributions to the valence band states for the frozen structure are from O (2p) orbitals slightly hybridized with Bi (6 s) and V (3d) states (see Fig. 13). Metal oxide semiconductors such as ZnO or TiO_2 usually have wide band gaps because the valence band has mainly O 2p character and is located at a positive potential (Kudo et al. 1999). The hybridization between the Bi (6 s) state and the O (2p) state at the top of the VB is indeed responsible for the relatively smaller band gap of BiVO_4 monoclinic scheelite structure (Walsh et al. 2009). In contrary to the frozen structure, the partial DOS of optimized monoclinic BiVO_4 underlines that the VBM

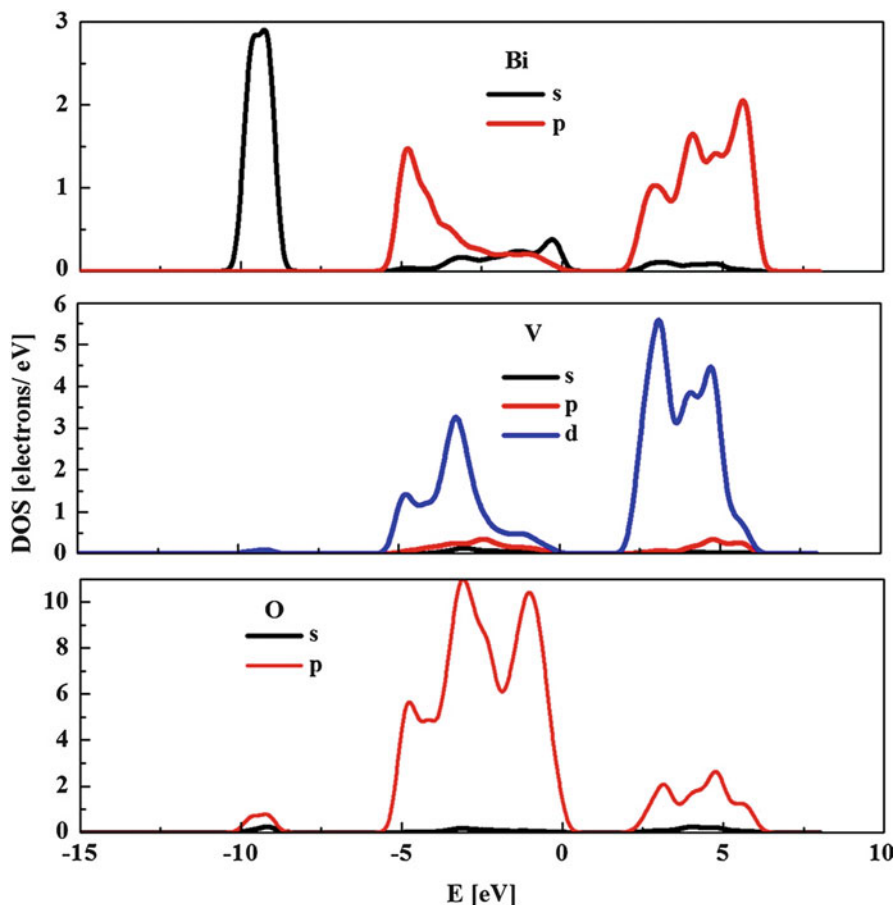


Fig. 13 Partial density of states of monoclinic BiVO_4 frozen structure

possesses contribution mainly from V (3p) states hybridized with O (2p) electrons. The mentioned oxygen–vanadium interaction decreases the calculated energy gap. In both frozen and optimized structure, the bismuth and oxygen contributions to the VBM are similar.

The conduction band of the frozen BiVO_4 crystal is dominated by V (3d) states with small contribution from Bi (6p) and O (2p) states. The same data were also obtained for relaxed crystal. These results are in agreement with the data measured by XPS (Payne et al. 2011). The oxygen and bismuth contribute to antibonding orbitals forming the CB states. The energy band structure obtained for frozen BiVO_4 single crystal gives better results compared to the experimental data. This indicates that the no-relaxed structure gives sound analysis for the electronic features. The distortions of Bi–O and V–O bond lengths enhance the impact of the Bi (6 s) states and contribute to more dispersion in the valence band.

Electronic Properties of NiTiO₃

The same calculations performed for the frozen and relaxed NiTiO₃ single crystal give unrealistic energy band gap values. Therefore, the electronic properties of the mentioned crystal were calculated at the spin-polarized approach for a primitive unit cell applying Hubbard parameters. These calculations were performed in k space within the BZ directions where the coordinates of the special points are (in units of the reciprocal lattice vectors) F (1/2, 1/2, 0), Γ (0, 0, 0), K (1/4, 1/4, 1/4), and Z (1/2, 1/2, 1/2). The DFT/PBE method with or without Hubbard approximation was taken into account. Such functional when used in the spin-polarized regime underestimates the electronic energy band gap. The obtained value equal to 0.77 eV is far from experimental data which is in the range 2.12–2.18 eV (Singh et al. 1995; Wang et al. 2013). The reason of such large discrepancy is probably due to the presence of titanium and nickel atoms which possess strongly correlated electrons. An improvement of the predicted results requires the use of Hubbard approximation. Figure 14 reports the electron density calculated by DFT/PBE potential and shows the covalent character of the O–Ti, O–Ni, and Ti–Ni bonds. This behavior makes us confident in the relevance of the statement about the studied material characterized by strongly correlated electrons. The Hubbard-modified Hamiltonian is required to obtain band energy gaps in agreement with experimental data.

The different Hubbard parameter (U) values were chosen for the valence Ti (3d) and Ni (3d) electrons and the energy gap values were calculated. The chosen U values summarized in Table 2. show that the close agreement between estimated band gap and experimental results are obtained for the Hubbard parameters equal

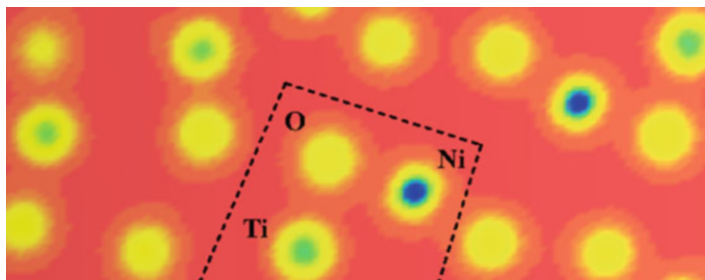


Fig. 14 Electron density plotted in a selected plane with the constitutive elements Ti, O, and Ni atoms (Ruiz Preciado et al. 2015)

Table 2 Hubbard U parameter values chosen for the Ti d and Ni d orbitals with the calculated energy gap value by using PBE + U methodology

Ti d	Ni d	E_g [eV]
0	0	0.77
2.5	2.5	1.75
3.5	4.5	2.33
4.5	3.5	1.94
4.5	4.5	2.46

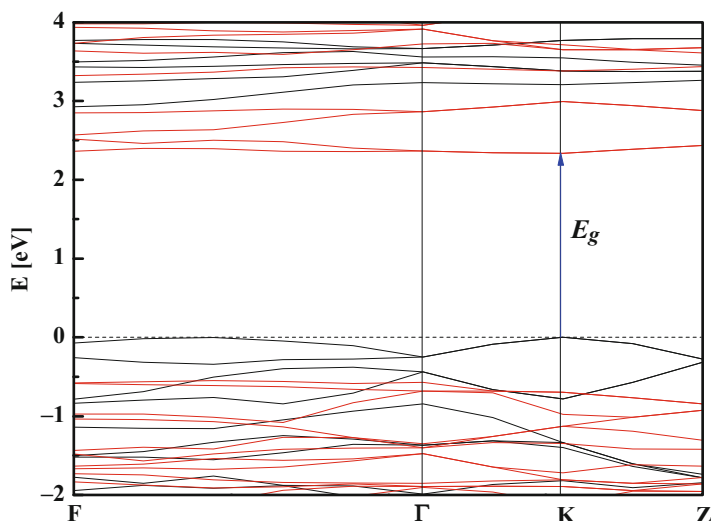


Fig. 15 Electron band structure calculated by PBE/LDA+U methodology with the energy levels created by alpha electron – *black* and energy levels created by beta electron – *red* (Ruiz Preciado et al. 2015)

to 3.5 eV and 4.5 eV for Ti (3d) and Ni (3d) electrons, respectively. These results underline that the PBE + U method provides a satisfactory qualitative electronic structure calculations with the Hubbard parameters U_{Ti} and U_{Ni} in a reasonable range. The electronic band structure computed with $U_{Ti} = 3.5$ eV and $U_{Ni} = 4.5$ eV is presented in Fig. 15. and indicates a direct semiconducting nature of $NiTiO_3$ crystal with an energy gap at 2.33 eV. The top of valence band is build by alpha electrons and the bottom of conduction band is composed by beta ones. The alpha electrons are involved in hybridized orbitals based on the Ni (3d) and O (2p) (see Fig. 16). The bottom of conduction band is also composed by the hybridization of the Ti (3d) and Ni (3d) states in agreement with former report by Salvador et al. (1982).

In both investigated materials, $NiTiO_3$ and $BiVO_4$, the valence band is created by the oxygen O (2p) electron possessing anion character, while the conduction band contains mostly V (3d) and Bi (6p) of the monoclinic $BiVO_4$ and Ti (3d) and Ni (3d) cations in the ilmenite $NiTiO_3$ crystal. Similarly, as discussed below, the valence band of TiO_2 is dominated by the O (2p) electrons, while the conduction band is populated mainly by the unoccupied Ti (3d) electron states.

Structural and Electronic Properties of TiO_2 , $NiTiO_3$, and $BiVO_4$ Nanosized Systems

TiO_2 Nanostructures

The built of nanocrystalline structures with defined geometries was described in Chapter, “[Cluster approach to model the nanocrystals and nanostructures](#)”. The

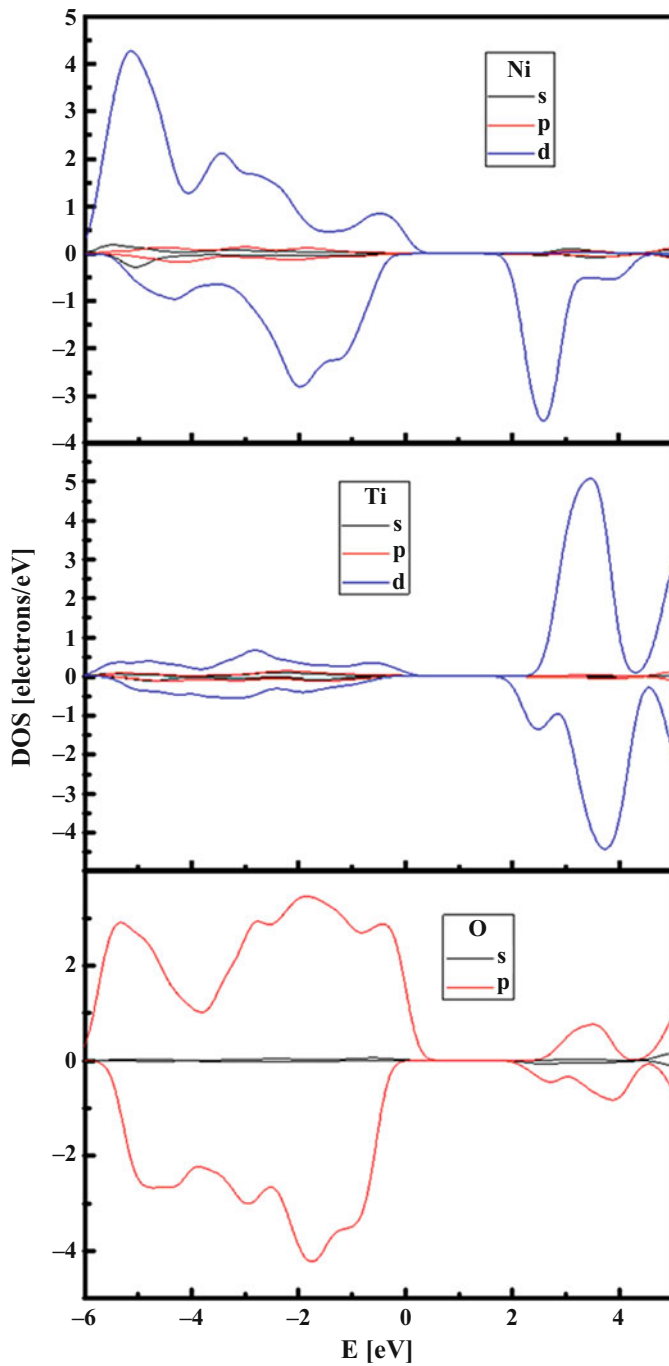


Fig. 16 Electron density of state (*DOS*) calculated for NiTiO₃ crystal using PBE+U method (Ruiz Preciado et al. 2015)

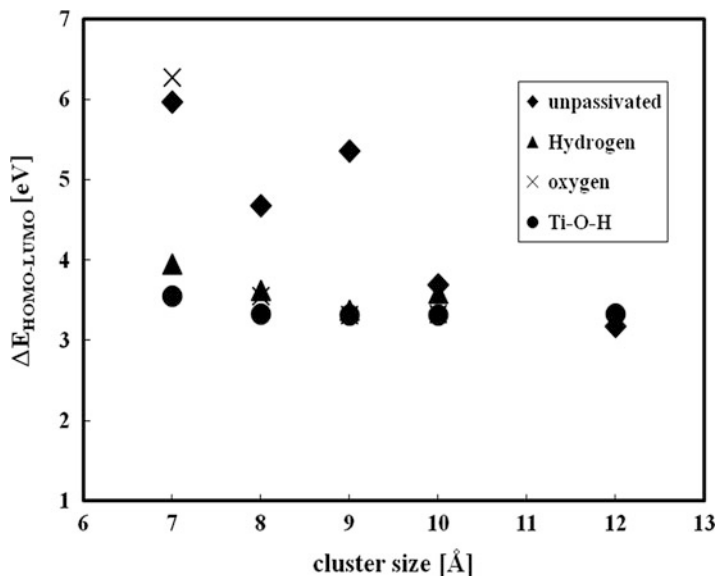


Fig. 17 HOMO–LUMO energy gap splitting versus size obtained by using parametrized PM6 semiempirical method for $(\text{TiO}_2)_n$ applying different passivation elements (Makowska-Janusik et al. 2014)

computational procedure considers $(\text{TiO}_2)_n$ clusters with size ranging from 0.6 up to 2.0 nm possessing indivisible number of TiO_2 units as $n = 2 \div 140$. All the clusters were spherical in agreement with the work of Kshirsagar et al. (Kshirsagar and Kumbhojkar 2008) establishing that small crystallites exhibit a spherical shape. The HOMO–LUMO energy gap splitting ($\Delta E_{\text{HOMO-LUMO}}$) was calculated for $(\text{TiO}_2)_n$ clusters using the semiempirical PM6 parameterized method. The electronic parameters were evaluated versus the cluster size and the different surface passivation nature such as hydrogen, oxygen, or $-\text{TiOH}$ groups. The obtained results point out the occurrence of quantum size confinement effect for all clusters. The size effect and the nature of surface passivation show correlated impact on the $\Delta E_{\text{HOMO-LUMO}}$ values for all clusters (see Fig. 17). Indeed, the most pronounced size effects were obtained with different features for unpassivated and oxygen-passivated clusters. Passivation by oxygen contributes to drastic shift of $\Delta E_{\text{HOMO-LUMO}}$ with a particular strength for the clusters with sizes less than 0.8 nm. These results are different than those from the work of Sahoo et al. (2011) probably due to our approach in the present work taking into account the cluster surface reconstruction which was implemented in accordance to the energy minimization procedure and environmental impact. According to our former analysis (Makowska-Janusik et al. 2014), in the surface reconstruction process, atoms at $(\text{TiO}_2)_n$ outermost cluster shell were moved according to the conjugated gradient algorithm with RMS gradient convergence equal to $0.01 \text{ kcal } \text{Å}^{-1} \text{ mol}^{-1}$.

Table 3 HOMO–LUMO energy gap splitting calculated by using different DFT functionals for unpassivated $(\text{TiO}_2)_n$ anatase clusters versus its size variation

Number of $(\text{TiO}_2)_N$ units n	Cluster size [nm]	BLYP	B3LYP	LC-BLYP $\mu = 0.33$	CAM-B3LYP	LC-BLYP $\mu = 0.80$
		$\Delta E_{\text{HOMO-LUMO}}$ [eV]				
2	0.6	0.97	1.04	4.63	2.92	4.76
9	0.8	0.77	0.83	3.62	2.28	3.97
18	1.0	0.47	0.50	2.94	1.85	3.03
40	1.2	0.42	0.45	2.91	1.46	3.03

These calculations indicate the important role of saturating atoms and surface reconstruction for small clusters, but for the clusters with diameter larger than 1.0 nm, the saturation effects can be omitted. This is convenient from the point of view of computer simulations because choosing realistic physical passivation is always problematic, even if it can bring satisfactory results (Kassiba et al. 2002). Therefore, for the detailed studies, only the unpassivated $(\text{TiO}_2)_n$ clusters were taken into account.

An evaluation of the $\Delta E_{\text{HOMO-LUMO}}$ of $(\text{TiO}_2)_n$ clusters was also calculated for structures with various sizes and different DFT functionals. The calculations were performed by using pure BLYP and hybrid B3LYP functional, as well as the different LC and CAM-B3LYP potentials. The obtained data are reported in Table 3. The HOMO–LUMO energy gap values obtained by BLYP and B3LYP functionals are quite small, but their evolution as a function of cluster size reproduces the quantum confinement effect. It is known that the band gap in semiconductors is underestimated by more than 30–50 % within the LDA approximation (Burke 2012). The calculations show also that the evaluated energy gap value can be substantially smaller than the real gap for semiconducting solids (Grüning et al. 2006). To overcome these discrepancies, a possibility consists of treating the exchange term of KS theory as orbital dependent by using hybrid functional (Seidl et al. 1996). However, in bulk crystal approach, the B3LYP potential does not match with the prediction of electronic features of the $(\text{TiO}_2)_n$ clusters.

For representative clusters, the use of LC-BLYP, LC-BOP, LC-BECK, or LC-BVWN functionals with a range separation parameter $\mu = 0.33$, gives similar values of the HOMO–LUMO energy gap splitting. In Table 3 only data obtained for LC-BLYP potential are presented. The LC and CAM functionals give better predictions of the electronic features than the pure BLYP or B3LYP hybrid functionals. However, LC-BLYP and CAM-B3LYP underestimate the HOMO–LUMO energy gap splitting compared to the band gap energy of experimental bulk anatase which is about 3.02 eV (Amtout and Leonelli 1995). The more relevant results are obtained by using the LC-BLYP functional, leading to a HOMO–LUMO gap of 2.94 eV and 2.91 eV for clusters with diameters of 1.0 nm [$(\text{TiO}_2)_{18}$] and 1.2 nm [$(\text{TiO}_2)_{40}$], respectively. Thus, the LC functional smoothly replaces the DFT exchange by the HF exchange at long inter-electron distances. Furthermore, it is possible to obtain larger HOMO–LUMO energy gap values by using different

Table 4 HOMO–LUMO energy gap splitting calculated for unpassivated $(\text{TiO}_2)_{18}$ anatase cluster with diameter equal to 1.0 nm using LC-BLYP functional versus range separation parameter μ

DFT functional	$\mu = 0.10$	$\mu = 0.33$	$\mu = 0.50$	$\mu = 0.80$
	$\Delta E_{\text{HOMO-LUMO}}$ [eV]			
LC-BLYP	2.88	2.94	2.99	3.03

range separation parameter μ . Several works on the LC scheme suggested that the separation parameter should be larger than the standard one equal to 0.33 in order to reproduce atomization energies or barrier height energies, for example (Gerber and Ángyán 2005). In this case, the HOMO–LUMO energy gap splitting was calculated for $(\text{TiO}_2)_{18}$ clusters of 1 nm diameter using different range separation parameters, within the LC-BLYP potential (see Table 4).

The $\Delta E_{\text{HOMO-LUMO}}$ value increases with the separation parameter μ . A relevant evaluation of the HOMO–LUMO energy gap is obtained for a μ parameter equal to 0.8 (see Table 3 last column). Higher values of μ from 0.9 to 1.5 were also examined, but they did not influence significantly the band gap. These results are consistent with former works of Iikura et al. (Iikura et al. 2001). It is important to notice that an increase of μ leads to a higher contribution of long-range exchange interactions from the HF exchange potential limiting also the self-exchange and self-correlation effect for the interacting electrons at the long inter-charge distance. The performed calculations are consistent with data obtained for strongly correlated systems with the electron density distributed close to the nuclei, as it is observed for ionic bonds. In the considered TiO_2 system, Ti atoms possess octahedral coordination, and the strong ionic character of the bonds results in localized charges around the O^{2-} anions. The HOMO orbital is dominated by the O (2p) electrons, while the LUMO is populated mainly by the unoccupied Ti (3d) electron states. The same results were obtained for the valence and conduction bands of TiO_2 crystal.

It was found that LC functionals in DFT method quantitatively reproduces energies of frontier orbitals including HOMO and LUMO. It is shown that for the strongly correlated systems, the LC functional with appropriate long-range separation factor may accurately reproduce the orbital energies in agreement with the prediction of Singh and Tsuneda (Singh and Tsuneda 2013). Within the developed approach, the performed calculations indicate that the choice of the separation coefficient μ modulates critically and in large extent the electronic properties of the considered system.

NiTiO₃ Nanostructures

The implementation of LC functional to calculate the electronic parameters of NiTiO_3 clusters did not evaluate coherently the HOMO–LUMO energy gap splitting. Reliable results were obtained by applying the semiempirical methodology. In Fig. 18, the evaluation of the energy difference between HOMO and LUMO calculated for the NiTiO_3 clusters is presented and the size quantum confinement effect is clearly seen. Thus, the results obtained by using PM6 methodology

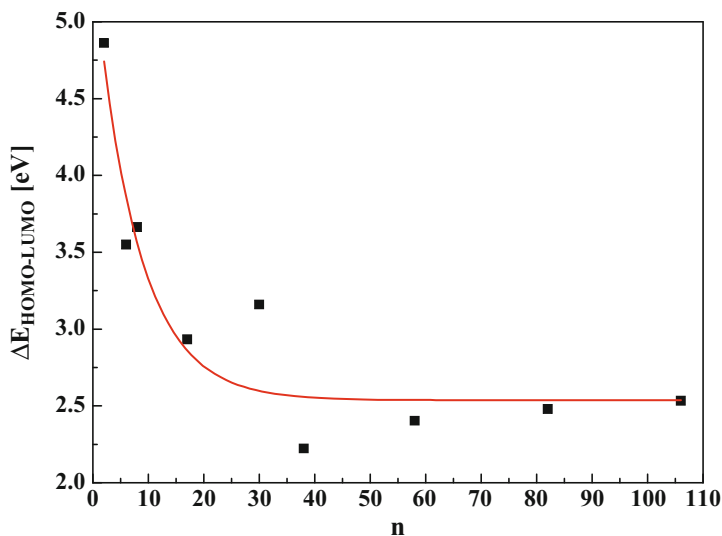
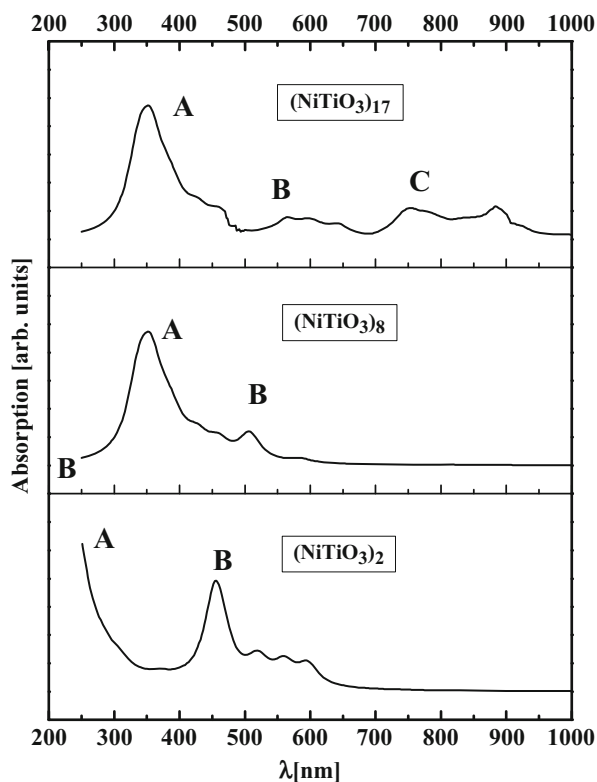


Fig. 18 Evaluation of the $\Delta E_{\text{HOMO-LUMO}}$ energy splitting versus number n of NiTiO_3 units calculated by PM6 methodology for the different $(\text{NiTiO}_3)_n$ clusters (Ruiz Preciado et al. 2015)

reproduce quantitatively the energy gap of nanosized NiTiO_3 . In the work of Vijayalakshmi and Rajendran (Vijayalakshmi and Rajendran 2012), it was specified that the size reduction of the NiTiO_3 nanoparticle gives blue shift in the absorption spectra compared to the bulk material.

To evaluate the electronic properties of NiTiO_3 nanostructures versus their size, UV–vis absorption spectra were calculated by PM6 methodology. Three different clusters were used, namely, $(\text{NiTiO}_3)_2$, $(\text{NiTiO}_3)_8$, and $(\text{NiTiO}_3)_{17}$. Large-sized clusters were not evaluated because of the computer memory problem. The calculated spectra are presented in Fig. 19. These results for the cluster $(\text{NiTiO}_3)_{17}$ agree with the experimental data obtained on NiTiO_3 powder (Fig. 20 top panel). The increase of the cluster size shifts the spectral A peak position to red range (see Fig. 19). The peak B which is well pronounced for the smallest cluster is also shifted to red side but in less marked rate for larger particles. The additional broad peak (C) at the position of 750–900 nm appears for the cluster $(\text{NiTiO}_3)_{17}$. The two peaks B and C are not observed in the spectrum calculated for the bulk NiTiO_3 single crystal (Fig. 20 down panel). Probably this behavior results from surface influence on the electronic properties of nanocrystal material. In our previous work (Ruiz Preciado et al. 2015), it was shown that the HOMO and LUMO orbitals observed for the investigated NiTiO_3 clusters are located at the separated atoms. The valence band minimum and conduction band maximum are diffused through the different atoms as it was shown in part concerning crystal electronic properties evaluation. The mentioned separation of HOMO and LUMO orbitals is not seen for the bulk materials where band structure exhibits different behavior. Probably, this explains

Fig. 19 The UV–vis absorption spectra calculated by the PM6 methodology for different clusters



why the B and C peaks are not visible in UV–vis absorption spectra of the single crystal.

The comparison between the theoretical UV–vis absorption spectrum calculated for the cluster $(\text{NiTiO}_3)_{17}$ and the experimental one (see Fig. 20, top panel) assigns quantitatively the origin of the observed features. The broad absorption edge situated at 410 nm is attributed to $\text{O}^{2-} \rightarrow \text{Ti}^{4+}$ charge transfer interaction. The lower energetic shoulder is associated with the crystal field splitting of NiTiO_3 , associated with the $\text{Ni}^{2+} \rightarrow \text{Ti}^{4+}$ charge transfer (Lopes et al. 2009). However, in contrast to experimental features, additional peaks are obtained on the theoretical UV–vis absorption spectra. These discrepancies are explained by the absence of temperature influence which causes structural relaxation of the nanoparticles. Indeed, taking into account such temperature influence, the electron–phonon interaction and the Franck–Condon rule lead to the calculated absorption peaks with more broad shapes.

Taking into account appropriate passivation of clusters and their reconstructed surfaces, the theoretical modeling of the nanosized structures reproduces their main physical features. This is illustrated on the vibrational properties through the calculated Raman spectra for selected clusters. Therefore, ilmenite crystal nanostructure

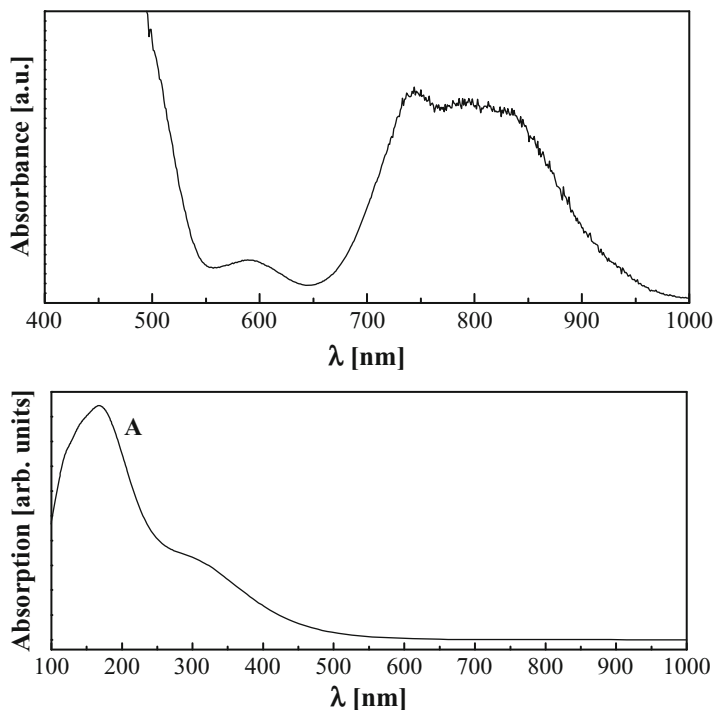


Fig. 20 The UV-vis absorption spectra measured for NiTiO₃ powder samples (*top panel*) and the spectrum calculated by PBE/LDA + U functional (*top right panel*) and the experimental data (*bottom right panel*)

with frozen atomic positions and nanoparticle with optimized geometry according to the total energy minimization was chosen. The studied structures possess common starting geometry indexed as (NiTiO₃)₁₇ cluster. The calculated and experimental Raman spectra are presented in Fig. 21 for the cluster (NiTiO₃)₁₇ with ilmenite structure. One may note that only one peak at position C is obtained. The origin of such vibrational band around 700 cm⁻¹ is attributed to the Ti–O–Ti stretching mode which can be slightly shifted in real crystal structure up to 720 cm⁻¹ (Su and Balmer 2000; Llabrés i Xamena et al 2003; Baraton et al. 1994). The peaks A, B, and D observed for the optimized nanostructure indicate the relaxed distance between Ni, Ti, and O bonds. Figure 21 reports also the convolution of the spectrum obtained for ilmenite nanocrystal and optimized structure with ratio 9/1. The strength of the developed theoretical and numerical modeling is clearly illustrated from the close correspondence between predicted Raman signatures and the experimental ones of nanopowder systems.

The obtained results such as vibrational or optical absorption support the fact that the nanocrystalline oxides give rise to two defined contributions from the core of particles and the external and outermost surface with amorphous or reconstructed

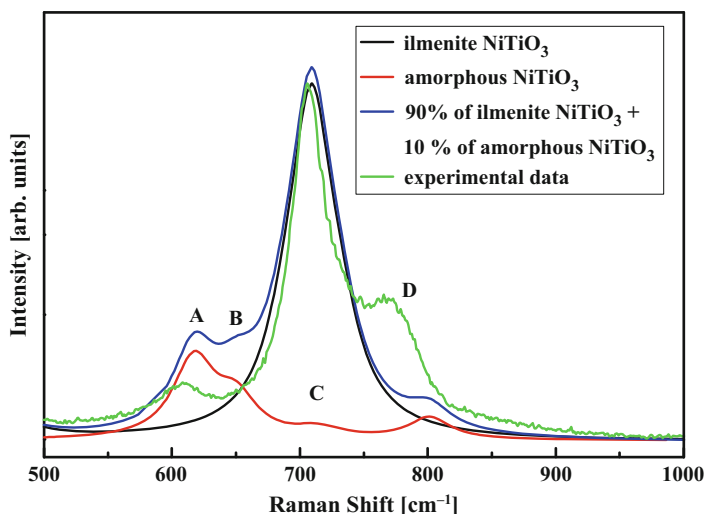


Fig. 21 Raman spectra of $(\text{NiTiO}_3)_{17}$ cluster calculated by the parametrized PM6 method and compared to experimental spectrum (Ruiz Preciado et al. 2015)

atomic positions. They traduce the properties of bulk nature and those traducing the surface broken periodicity.

BiVO_4 Nanostructures

Beyond NiTiO_3 , similar behavior was also noticed from the size quantum confinement effect involved in BiVO_4 nanostructures. Using the semiempirical PM6 parameterized method, the electronic structures of isolated $(\text{BiVO}_4)_n$ clusters were calculated. The number n of structural units was changed from $n = 1$ up to $n = 56$ leading to cluster diameter up to 2.0 nm. The calculated $\Delta E_{\text{HOMO-LUMO}}$ energy gap splitting is presented in Fig. 22. The predicted values of the HOMO–LUMO difference are divided into two parts as marked by rectangles and triangles. The first class of mentioned data can be fitted by an exponential curve. Qualitatively, this behavior is in agreement with quantum confinement size effect. The second class of data (triangles) represents the clusters with frozen structure as for bulk system. These data fluctuate with increasing the cluster size but have tendency to be aligned by one asymptotic value close to 2.4 eV. Such behavior is in agreement with the experimental value of the energy gap defined for bulk BiVO_4 . The area marked by dashed line contains cluster size range where the frozen bulk-like and surface-relaxed structures defining the property of nanoparticle compete one with the other. This area corresponds to cluster sizes ranging from 1.1 up to 1.8 nm where the surface and volume atoms contribute differently to the electronic band structure of the cluster. The origin of the calculated $\Delta E_{\text{HOMO-LUMO}}$ is the effect of emulation

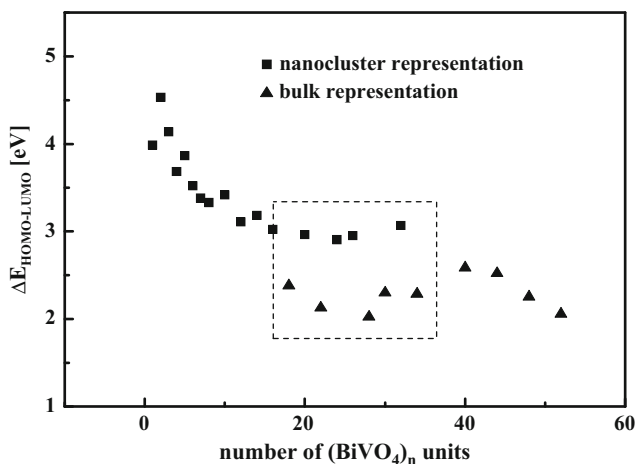


Fig. 22 The $\Delta E_{\text{HOMO-LUMO}}$ energy gap splitting versus size of BiVO_4 cluster represented by the number of $(\text{BiVO}_4)_n$ units. *Dashed line* indicates the intermediate area between bulk frozen atoms and nanoparticle behavior with contributions from core and surface atoms

between two competing mechanisms. One is believed to come from the crystalline cores of the nanometer-sized BiVO_4 and the other one is attributed to the surface layer. Similar mechanism was used to explain the photoexcitation effect occurring in the crystalline Si cores (Wilcoxon et al. 1999).

The electronic properties of the BiVO_4 nano-clusters computed by PM6 semiempirical method confirm that the HOMO orbital is built by the oxygen states, while LUMO orbital include vanadium states. It confirms that the limitation of the cluster size did not change drastically its electronic configuration and properties. The obtained results show that the electronic structure of nanoparticle is the consequence of bulk crystal material. The HOMO and LUMO orbitals agree with the VBM and CBM of frozen BiVO_4 single crystal, and performed calculations are in agreement with the works of Huang et al. (Huang et al. 2014) and Jo et al. (Jo et al. 2012).

The theoretical calculations are consistent with the experimental features reported in the first part of the presented chapter. This fact makes us confident in the developed numerical approach to analyze the electronic and optical effects of this class of semiconducting oxides. The quantum confinement effect is well demonstrated theoretically on the optical spectrum of nanosized cluster. However, our experimental synthesis did not allow so small nanocrystal sizes, and only the features 20–30 nm particles were indeed demonstrated experimentally. Dealing with the photoinduced charge transfer required for visible-driven photocatalysis, the experimental results point out the visible light absorption in a wide spectral range. The nanoparticle sizes 20–50 nm contribute to enhance the specific surface which is critical in heterogeneous photocatalysis. Indeed, as the contacts at the interfaces increase between photocatalysts and surrounding media, an improvement occurs on the efficiency of photocatalysis reactions. The benefit from nanosized clusters

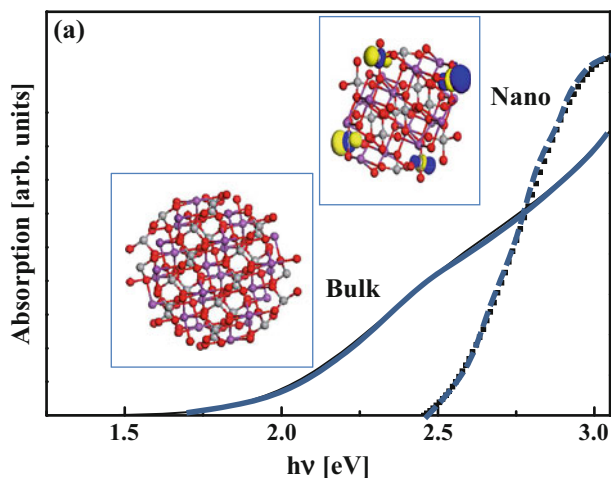


Fig. 23 Calculated optical absorption spectra of BiVO_4 as nanocrystalline object and infinite bulk system. The blue shift due to quantum confinement effect is traduced on the nanocrystal absorption curve

is thought to be related to drastic enhancement of specific surfaces. However, as it was shown from the calculations (Fig. 23), the spectral range of absorbed sun light is shifted to higher energies limiting then the absorption quantum efficiency. Thus, a balance must be achieved between the size reduction (specific surface) and the induced widening of the band gap.

Conclusions

Exhaustive experimental investigations and theoretical approaches supported by numerical simulations were dedicated to selected classes of semiconducting oxides such as TiO_2 , BiVO_4 , and NiTiO_3 characterized by promising visible-driven photoactive responses. The ambition consists of tailoring original smart materials from chemical composition, structural and morphological features for pphotovoltaic and photocatalysis applications. Original synthesis strategies were developed to obtain these materials as thin films, mesoporous structures, as well as nanoparticles. To create visible-driven photoactive media with higher efficiency, suitable doping was realized to modulate the electronic and optical properties. Structural and vibrational investigations were carried out and inform on the stabilized crystalline polytypes and their evolution with the synthesis and treatment conditions. The physical properties including electronics and optics, as well as the role of electronic active centers realized by doping or intrinsically induced by oxygen vacancies or form the stoichiometry departure, were also investigated, and their interpretation was made quantitatively with the support of exhaustive computational work. Quantum theory approaches based on DFT and semiempirical models were applied to predict the

physical peculiarities of photoactive semiconducting oxides. Original approaches were tested and the appropriate functionals (potentials) were defined coherently with the chemical nature of the materials. Computer simulations were performed for single crystals and implemented successfully for nanosized structures. For all chosen nanoparticles as TiO_2 , NiTiO_3 , and BiVO_4 , the appearance of the size quantum confinement effect was demonstrated on the electronic and optical features. The HOMO orbitals of all investigated nanocrystals are composed of the O (2p) states and the LUMO depends, respectively, on the V (3d) or Ti (3d) atoms. The hybridization of the O (2p) orbitals in the area of occupied orbitals with the other orbitals as Bi (6 s) decreases the energy gap for BiVO_4 . For structural aspects, surfaces of nanostructured objects are reconstructed due to the environmental influence inducing changes on the interatomic distances. Orbital hybridization contributing to extremes of valence and conduction bands is modified by such surface effects with particular strength for low-sized clusters as traduced from the carried out numerical simulations.

The close correlation between the experimental investigations on such semiconducting oxide systems and the theoretical approaches supported by numerical simulation paves the way to understand quantitatively the physical peculiarities of such photoactive systems. Charge transfer effects, role of doping, and structural arrangements as function for the material morphology are understood by such combined experimental investigations and theoretical models. Our ambition is to contribute by such methodology to design the relevant structures with improved photoactive properties able to be implemented in new technologies required for societal challenges like clean energy and safe environment.

Bibliography

- Albaret, T., Finocchi, F., & Noguera, C. (2000). Density functional study of stoichiometric and O-rich titanium oxygen clusters. *Journal Chemical Physics*, *113*, 2238–2249.
- Albuquerque, A. R., Garzim, M. L., dos Santos, I. M. G., Longo, V., Longo, E., & Sambrano, J. R. (2012). DFT study with inclusion of the Grimme potential on anatase TiO_2 : Structure, electronic, and vibrational analyses. *Journal of Physical Chemistry A*, *116*, 11731–11735.
- Almbladh, C. O., & Pedroza, A. C. (1984). Density-functional exchange-correlation potentials and orbital eigenvalues for light atoms. *Physical Review A*, *29*, 2322.
- Amtout, A., & Leonelli, R. (1995). Optical properties of Rutile near its fundamental band gap. *Physical Review B*, *51*, 6842–6851.
- Anisimov, V. I., Zaanen, J., & Andersen, O. K. (1991). Band theory and Mott insulators: Hubbard U instead of Stoner I. *Physical Review B*, *44*, 943–954.
- Appel, F., Gross, E. K. U., & Burke, K. (2003). Excitations in time-dependent density-functional theory. *Physical Review Letters*, *90*, 043005.
- Bajaj, R., Sharma, M., & Bahadur, D. (2013). Visible light-driven novel nanocomposite ($\text{BiVO}_4/\text{CuCr}_2\text{O}_4$) for efficient degradation of organic dye. *Dalton Transactions*, *42*, 6736–6744.
- Baraton, M. I., Busca, G., Prieto, M. C., Ricchiardi, G., & Sanchez Escribano, V. (1994). On the vibrational spectra and structure of FeCrO_3 and of the Ilmenite-Type compounds CoTiO_3 and NiTiO_3 . *Journal of Solid State Chemistry*, *112*, 9–14.

- Barnard, A. S., & Zapol, P. (2004a). A model for the phase stability of arbitrary nanoparticles as a function of size and shape. *Journal of Chemical Physics*, *121*, 4276–4283.
- Barnard, A. S., & Zapol, P. (2004b). Effects of particle morphology and surface hydrogenation on the phase stability of TiO₂. *Physical Review B*, *70*, 235403–13.
- Becke, A. D. (1988). Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, *38*, 3098–3100.
- Becke, A. D. (1993). Density-functional thermochemistry. iii. the role of exact exchange. *Journal of Chemical Physics*, *98*, 5648–5652.
- Bergner, A., Dolg, M., Kuechle, W., Stoll, H., & Preuss, H. (1993). Ab initio energy-adjusted pseudopotentials for elements of groups 13–17. *Molecular Physics*, *80*, 1431–1441.
- Bermudez, V. M. (2010). Ab initio study of the interaction of Dimethyl Methylphosphonate with Rutile (110) and Anatase (101) TiO₂ surfaces. *Journal Physical Chemistry C*, *114*, 3063–3074.
- Biswas, S., Hossain, M. F., & Takahashi, T. (2008). Fabrication of Grätzel solar cell with TiO₂/CdS bilayered photoelectrode. *Thin Solid Films*, *517*, 1284–1288.
- Blagojevic, V., Chen, Y. R., Steigerwald, M., Brus, L., & Friesner, R. (2009). A. Quantum chemical investigation of cluster models for TiO₂ nanoparticles with water-derived ligand passivation: Studies of excess electron states and implications for charge transport in the Gratzel cell. *Journal of Physical Chemistry C*, *113*, 19806–19811.
- Burdett, J. K., Hughbanks, T., Miller, G. J., Richardson, J. W., & Smith, J. V. (1987). Structural-electronic relationships in inorganic solids: powder neutron diffraction studies of the rutile and anatase polymorphs of titanium dioxide at 15 and 295 K. *Journal of the American Chemical Society*, *109*, 3639–3646.
- Burke, K. (2012). Perspective on density functional theory. *Journal of Chemical Physics*, *136*, 150901.
- Ceperley, D. M., & Alder, B. J. (1980). Ground state of the electron gas by a stochastic method. *Physical Review Letters*, *45*, 566–569.
- Chandiran, A. K., Abdi-Jalebi, M., Nazeeruddin, M. K., & Grätzel, M. (2014). Analysis of electron transfer properties of ZnO and TiO₂ photoanodes for Dye-Sensitized solar cells. *ACS Nano*, *8*, 2261–2268.
- Chretien, S., & Metiu, H. (2011). Electronic structure of partially reduced Rutile TiO₂(110) surface: Where are the unpaired electrons located? *Journal Physical Chemistry C*, *115*, 4696–4705.
- Cooper, J. K., Gul, S., Toma, F. M., Chen, L., Glans, P.-A., Guo, J., Ager, J. W., Yano, J., & Sharp, I. D. (2014). Electronic structure of monoclinic BiVO₄. *Chemistry of Materials*, *26*, 5365–5373.
- Cundari, T. R., & Stevens, W. J. (1993). Effective core potential methods for the lanthanides. *Journal of Chemical Physics*, *98*, 5555–5565.
- Delley, B. (1990). An all-electron numerical method for solving the local density functional for polyatomic molecules. *Journal of Chemical Physics*, *92*, 508–517.
- Delley, B. (2000). From molecules to solids with the DMol 3 approach. *Journal of Chemical Physics*, *113*, 7756–7764.
- Dolg, M., Wedig, U., Stoll, H., & Preuss, H. (1987). Energy-adjusted ab initio pseudopotentials for the first row transition elements. *Journal of Chemical Physics*, *86*, 866–872.
- Dreuw, A., Weisman, J. L., & Head-Gordon, M. (2003). Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange. *Journal of Chemical Physics*, *119*, 2943.
- Ferrari, M., & Lutterotti, L. (1994). Method for the simultaneous determination of anisotropic residual stresses and texture by X-ray diffraction. *Journal of Applied Physics*, *76*, 7246–7255.
- Gałyńska, M., & Persson, P. (2013). Emerging polymorphism in nanostructured TiO₂: Quantum chemical comparison of anatase, rutile, and brookite clusters. *International Journal of Quantum Chemistry*, *113*, 2611–2620.
- Gawande, S. B., & Thakare, S. R. (2012). Graphene wrapped BiVO₄ photocatalyst and its enhanced performance under visible light irradiation. *International Nano Letters*, *2*, 1–7.
- Gerber, I. C., & Ángyán, J. G. (2005). Hybrid functional with separated range. *Chemical Physics Letters*, *415*, 100–105.

- Grüning, M., Marini, A., & Rubio, A. (2006). Density functionals from many-body perturbation theory: The band gap for semiconductors and insulators. *Journal Chemical Physics*, *124*, 154108.
- Hamad, S., Catlow, C. R. A., Woodley, S. M., Lago, S., & Mejias, J. A. (2005). Structure and stability of small TiO₂ nanoparticles. *Journal of Physical Chemistry B*, *109*, 15741–15748.
- Hedin, L., & Lundqvist, B. I. (1971). Explicit local exchange correlation potentials. *Journal of Physics C*, *4*, 2064–2083.
- Heyd, J., Scuseria, G. E., & Ernzerhof, M. (2003). Hybrid functionals based on a screened Coulomb potential. *Journal of Chemical Physics*, *118*, 8207–8215.
- Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. *Physical Review B*, *136*, 864–871.
- Huang, J. R., & Cheng, C. (2013). Cation and magnetic orders in MnFe₂O₄ from density functional calculations. *Journal of Applied Physics*, *113*, 033912.
- Huang, Z. F., Pan, L., Zou, J. J., Zhang, X., & Wang, L. (2014). Nanostructured bismuth vanadate-based materials for solar-energy-driven water oxidation: a review on recent progress. *Nanoscale*, *6*, 14044–14063.
- Hwang, Y. J., Boukai, A., & Yang, P. (2009). High density n-Si/n-TiO₂ core/shell nanowire arrays with enhanced photoactivity. *Nano Letters*, *9*, 410–415.
- Hybertsen, M. S., & Louie, S. G. (1989). Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies. *Physical Review B*, *34*, 5390–5413.
- Iikura, H., Tsuneda, T., Yanai, T., & Hirao, K. (2001). A long-range correction scheme for generalized-gradient-approximation exchange functionals. *Journal of Chemical Physics*, *115*, 3540–3544.
- Ishigaki, T., Li, Y.-L., & Kataoka, E. (2003). Phase formation and microstructure of titanium oxides and composites produced by thermal plasma oxidation of titanium carbide. *Journal of the American Ceramic Society*, *86*, 1456–1463.
- Jain, A., Hautier, G., Ong, S. P., Moore, C. J., Fischer, C. C., Persson, K. A., & Ceder, G. (2011). Formation enthalpies by mixing GGA and GGA+U calculations. *Physical Review B*, *84*, 045115.
- Jo, W. J., Jang, J. W., Kong, K. J., Kang, H. J., Kim, J. Y., Jun, H., Parmar, K. P. S., & Lee, J. S. (2012). Phosphate doping into monoclinic BiVO₄ for enhanced photoelectrochemical water oxidation activity. *Angewandte Chemie International Edition*, *51*, 3147–3151.
- Kassiba, A., Makowska-Janusik, M., Boucle, J., Bardeau, J. F., Bulou, A., & Herlin-Boime, N. (2002). Photoluminescence features on the Raman Spectra of quasistoichiometric SiC nanoparticles: experimental and numerical simulations. *Physical Review B*, *66*, 155317.
- Kassiba, A., Pattier, B., Henderson, M., Makowska-Janusik, M., Mei, P., & Gibaud, A. (2012). Titanium oxide based mesoporous powders and gels: Doping effects and photogenerated charge transfer. *Materials Science and Engineering B*, *177*, 1446–1451.
- Kim, T. W., & Choi, K. S. (2014). Nanoporous BiVO₄ photoanodes with dual-layer oxygen evolution catalysts for solar water splitting. *Science*, *343*, 990–994.
- Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Physical Review A*, *140*, 1133–1138.
- Krukau, A. V., Vydrov, O. A., Izmaylov, A. F., & Scuseria, G. E. (2006). Influence of the exchange screening parameter on the performance of screened hybrid functionals. *Journal of Chemical Physics*, *125*, 224106.
- Kshirsagar, A., & Kumbhojkar, N. (2008). Empirical pseudo-potential studies on electronic structure of semiconducting quantum dots. *Bulletin of Materials Science*, *31*, 297–308.
- Kudo, A., Omori, K., & Kato, H. (1999). A novel aqueous process for preparation of crystal form-controlled and highly crystalline BiVO₄ powder from layered vanadates at room temperature and its photocatalytic and photophysical properties. *Journal of the American Chemical Society*, *121*, 11459–11467.
- Labat, F., Baranek, P., Domain, C., Minot, C., & Adamo, C. (2007). Density functional theory analysis of the structural and electronic properties of TiO₂ Rutile and anatase polytypes:

- Performances of different exchange-correlation functionals. *Journal of Chemical Physics*, *126*, 154703.
- Langreth, D. C., & Perdew, J. P. (1980). Theory of nonuniform electronic systems. I. Analysis of the gradient approximation and a generalization that works. *Physical Review B*, *21*, 5469–5493.
- Laskowski, R., Christensen, N. E., Blaha, P., & Palanivel, B. (2009). Strong excitonic effects in CuAlO_2 delafossite transparent conductive oxides. *Physical Review B*, *79*, 165209.
- Le, J. Y., Park, J., & Cho, J. H. (2005). Electronic properties of N- and C-doped TiO_2 . *Applied Physics Letters*, *87*, 011904.
- Levy, M. (1979). Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v -representability problem. *Proceedings of the National Academy of Sciences of the USA*, *76*, 6062–6065.
- Levy, M., Perdew, J. P., & Sahni, V. (1984). Exact differential equation for the density and ionization energy of a many-particle system. *Physical Review A*, *30*, 2745.
- Lin, Y. Y., Chu, T. H., Li, S. S., Chuang, C. H., Chang, C. H., Su, W. F., Chang, C. P., Chu, M. W., & Chen, C. W. (2009). Interfacial nanostructuring on the performance of polymer/ TiO_2 nanorod bulk heterojunction solar cells. *Journal of the American Chemical Society*, *131*, 3644–3649.
- Liu, Q. J., & Liu, Z. T. (2011). First-principles generalized gradient approximation + U study of cubic CuAl_2O_4 . *Applied Physics Letters*, *99*, 091902.
- Lopes, K. P., Cavalcante, L. S., Simoes, A. Z., Varela, J. A., Longo, E., & Leite, E. R. (2009). NiTiO_3 powders obtained by polymeric precursor method: Synthesis and characterization. *Journal of Alloys and Compounds*, *468*, 327.
- Ma, J., & Wang, L.-W. (2014). The role of the isolated 6s states in BiVO_4 on the electronic and atomic structures. *Applied Physics Letters*, *105*, 172102.
- Madhusudan, P., Ran, J., Zhang, J., Yu, J., & Liu, G. (2011). Novel urea assisted hydrothermal synthesis of hierarchical $\text{BiVO}_4/\text{Bi}_2\text{O}_3\text{CO}_3$ nanocomposites with enhanced visible-light photocatalytic activity. *Applied Catalysis B: Environmental*, *110*, 286–295.
- Makowska-Janusik, M., & Kassiba, A. (2012). Functional nanostructures and nanocomposites – Numerical modeling approach and experiment. In J. Leszczynski (Ed.), *Handbook of computational chemistry*. New York: Springer. ISBN 978-94-007-0711-5.
- Makowska-Janusik, M., Gladii, O., Kassiba, A., Herlin-Boime, N., & Bouclé, J. (2014). Cluster approach to model titanium dioxide as isolated or organic dye sensitized nanoobjects. *Journal of Physical Chemistry C*, *118*, 6009–6018.
- Martsinovich, N., Jones, D. R., & Troisi, A. (2010). Electronic structure of TiO_2 surfaces and effect of molecular adsorbates using different DFT implementations. *Journal Physical Chemistry C*, *114*, 22659–22670.
- Mei, P., Henderson, M., Kassiba, A., & Gibaud, A. (2010). EPR study of nitrogen-doped mesoporous TiO_2 powders. Original. *Journal of Physics and Chemistry of Solids*, *71*, 1–6.
- Melhem, H., Simon, P., Wang, J., Di Bin, C., Ratier, B., Leconte, Y., Herlin-Boime, N., Makowska-Janusik, M., Kassiba, A., & Bouclé, J. (2013). Direct photocurrent generation from nitrogen doped TiO_2 electrodes in solid-state dye-sensitized solar cells: Towards optically-active metal oxides for photovoltaic applications. *Solar Energy Materials and Solar Cells*, *177*, 624–631.
- Merupo, V. I., Velumani, S., Oza, G., Makowska-Janusik, M., & Kassiba, A. (2015a). Structural, electronic and optical features of molybdenum-doped bismuth vanadium oxide. *Materials Science in Semiconductor Processing*, *31*, 618–623.
- Merupo, V. I., Velumani, S., Ordon, K., Errien, N., Szade, J., & Kassiba, A. (2015b). Structural and optical characterizations of ball milled copper doped bismuth vanadium oxide (BiVO_4). *CrystEngComm*, *17*, 3366–3375. doi:[10.1039/C5CE00173K](https://doi.org/10.1039/C5CE00173K).
- Monkhorst, H. J., & Pack, J. D. (1976). Special points for Brillouin-zone integrations. *Physical Review B*, *13*, 5188–5192.
- Monkhorst, H. J., & Pack, J. D. (1977). Special points for Brillouin-zone integrations – a reply. *Physical Review B*, *16*, 1748–1749.

- Nam, H.-. J., Amemiya, T., Murabayashi, M., & Itoh, K. (2004). Photocatalytic activity of Sol-gel TiO₂ thin films on various kinds of glass substrates: The effects of Na⁺ and primary particle size. *Journal of Physical Chemistry B*, *108*, 8254–8259.
- Na-Phattalung, S., Smith, M. F., Kim, K., Du, M.-H., Wei, S.-H., Zhang, S. B., & Limpijumnong, S. (2006). First-principles study of native defects in anatase TiO₂. *Physical Review B*, *73*, 125205.
- Ng, Y. H., Iwase, A., Kudo, A., & Amal, R. (2010). Reducing graphene oxide on a visible-light BiVO₄ photocatalyst for an enhanced photoelectrochemical water splitting. *Journal of Physical Chemistry Letters*, *1*, 2607–2612.
- Pattier, B. (2010). Ph.D. dissertation. University of Maine, Le Mans France. cyberdoc.univ-lemans.fr/theses/2010/2010LEMA1019.pdf
- Pattier, B., Henderson, M., Pöppl, A., Kassiba, A., & Gibaud, A. (2010). Multi-approach electron paramagnetic resonance investigations of UV-photoinduced Ti(3+) in titanium oxide-based gels. *Journal of Physical Chemistry B*, *114*, 4424–4431.
- Payne, M. C., Teter, M. P., Allan, D. C., Arias, T. A., & Joannopoulos, J. D. (1992). Iterative minimization techniques for Ab Initio total energy calculations: Molecular dynamics and conjugate gradients. *Reviews of Modern Physics*, *64*, 1045–1097.
- Payne, D. J., Robinson, M. D. M., Egdell, R. G., Walsh, A., McNulty, J., Smith, K. E., & Piper, L. F. J. (2011). The nature of electron lone pairs in BiVO₄. *Applied Physics Letters*, *98*, 212110.
- Perdew, J. P., & Zunger, A. (1981). Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B*, *23*, 5048–5079.
- Perdew, J. P., Chevary, J. A., Vosko, S. H., Jackson, K. A., Pederson, M. R., Singh, D. J., & Fiolhais, C. (1992). Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B*, *45*, 13244–13249.
- Perdew, J. P., Burke, K., & Ernzerhof, M. (1996a). Generalized gradient approximation made simple. *Physical Review Letters*, *77*, 3865–3868.
- Perdew, J. P., Ernzerhof, M., & Burke, K. (1996b). Rationale for mixing exact exchange with density functional approximations. *Journal of Chemical Physics*, *105*, 9982–9985.
- Persson, P., Gebhardt, J. C. M., & Luneel, S. (2003). The smallest possible nanocrystals of semiionic oxides. *Journal Physical Chemistry*, *B107*, 3336–3339.
- Pfrommer, B. G., Cate, M., Louie, A. G., & Cohen, M. L. (1997). Relaxation of crystals with the Quasi-Newton method. *Journal of Computational Physics*, *131*, 233–240.
- Qu, Z. W., & Kroes, G. J. (2006). Theoretical study of the electronic structure and stability of titanium dioxide clusters (TiO₂)_n with n = 1–9. *Journal of Physical Chemistry B*, *110*, 8998–9007.
- Qu, Z. W., & Kroes, G. J. (2007). Theoretical study of stable, defect-free (TiO₂)_n nanoparticles with n = 10–16. *Journal of Physical Chemistry C*, *111*, 16808–16817.
- Rajalingam, V., Velumani, S., & Kassiba, A. (2012). Mechanochemical synthesis of nanostructured BiVO₄ and investigations of related features. *Materials Chemistry and Physics*, *135*, 842–848.
- Reyes-Coronado, D., Rodríguez-Gattorno, G., Espinosa-Pesqueira, M. E., Cab, C., Coss, R., & Oskam, G. (2008). Phase-pure TiO₂ nanoparticles: anatase, brookite and rutile. *Nanotechnology*, *19*, 145605–145615.
- Ruiz Preciado, M. A., Kassiba, A., Morales-Acevedo, A., & Makowska-Janusik, M. (2015). Vibrational and electronic peculiarities of NiTiO₃ nanostructures inferred from first principle calculations. *RSC Advances*, *5*, 17396–17404.
- Sahoo, S. K., Pal, S., Sarkar, P., & Majumder, C. (2011). Size-dependent electronic structure of Rutile TiO₂ quantum dots. *Chemical Physics Letters*, *516*, 68–71.
- Salvador, P., Gutierrez, C., & Goodenough, J. B. (1982). Photoresponse of n-type semiconductor NiTiO₃. *Applied Physics Letters*, *40*, 188–190.
- Sayed, A., Khenata, R., Chahed, A., & Benhelal, O. (2013). Electronic and optical properties of layered RE₂Ti₂O₇ (RE=Ce and Pr) from first principles. *Journal of Applied Physics*, *113*, 173501.
- Schmidt, M. W., Baldrige, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S. J., Windus, T. L., Dupuis, M., & Montgomery, J. A.

- (1993). General atomic and molecular electronic structure system. *Journal of Computational Chemistry*, *14*, 1347–1363.
- Seidl, A., Görling, A., Vogl, P., Majewski, J. A., & Levy, M. (1996). Generalized Kohn-Sham schemes and the band-gap problem. *Physical Review B*, *53*, 3764–3774.
- Singh, R. K., & Tsuneda, T. (2013). Reaction energetics on long-range corrected density functional theory: Diels–Alder reactions. *Journal of Computational Chemistry*, *34*, 379–386.
- Singh, R. S., Ansari, T. H., Singh, R. A., & Wanklyn, B. M. (1995). Electrical conduction in NiTiO₃ single crystals. *Materials Chemistry and Physics*, *40*, 173–177.
- Slater, J. C. (1951). A simplification of the Hartree-Fock method. *Physical Review*, *81*, 385–390.
- Sleight, A. W., Chen, H. Y., Ferretti, A., & Cox, D. E. (1979). Crystal growth and structure of BiVO₄. *Materials Research Bulletin*, *14*, 1571–1581.
- Srivastava, G. P., & Weaire, D. (1987). The theory of the cohesive energies of solids. *Advances in Physics*, *26*, 463–517.
- Stausholm-Møller, J., Kristoffersen, H. H., Hinnemann, B., Madsen, G. K. H., & Hammer, B. (2010). DFT+U study of defects in bulk rutile TiO₂. *Journal of Chemical Physics*, *133*, 144708–144715.
- Stephenset, P. J., Devlin, F. J., Chabalowski, C. F., & Frisch, M. J. (1994). Ab Initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *Journal of Physical Chemistry*, *98*, 11623–11627.
- Stevens, W. J., Basch, H., & Krauss, M. (1984). Compact effective potentials and efficient shared-exponent basis sets for the first- and second-row atoms. *Journal of Chemical Physics*, *81*, 6026–6033.
- Stewart, J. J. P. (2007). Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular Modeling*, *13*, 1173–1213.
- Stewart, J. J. P. (2012). *MOPAC2012 Stewart computational chemistry*, Colorado Springs. [HTTP://OpenMOPAC.net](http://OpenMOPAC.net)
- Su, Y., & Balmer, M. L. (2000). Raman spectroscopic studies of silicotitanates. *Journal of Physical Chemistry B*, *104*, 8160.
- Tawada, Y., Tsuneda, T., Yanagisawa, S., Yanai, Y., & Hirao, K. (2004). A long-range-corrected time-dependent density functional theory. *Journal of Chemical Physics*, *120*, 8425–8433.
- Vanderbilt, D. (1990). Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Physical Review B*, *41*, 7892–7895.
- Venkatesan, R. (2014). Ph.D. Thesis dissertation, University of Maine Le Mans www.theses.fr/2014LEMA3004.pdf
- Venkatesan, R., Velumani, S., & Kassiba, A. (2012). Mechanochemical synthesis of nanostructured BiVO₄ and investigations of related features. *Materials Chemistry and Physics*, *135*, 842–848.
- Venkatesan, R., Velumani, S., Tabellout, M., Errien, N., & Kassiba, A. (2013). Dielectric behavior, conduction and EPR active centres in BiVO₄ nanoparticles. *Journal of Physics and Chemistry of Solids*, *74*, 1695–1702.
- Vijayalakshmi, R., & Rajendran, V. (2012). Effect of reaction temperature on size and optical properties of NiTiO₃ nanoparticles. *E-Journal of Chemistry*, *9*, 282–288.
- Vosko, S. H., Wilk, L., & Nusair, M. (1980). Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis. *Canadian Journal of Physics*, *58*, 1200–1211.
- Walsh, M. B., King, R. A., & Schaefer, H. F. (1999). The structures, electron affinities, and energetic stabilities of TiOn and (n=1–3). *Journal Chemical Physics*, *110*, 5224–5230.
- Walsh, A., Yan, Y., Huda, M. N., Al-Jassim, M. M., & Wei, S. H. (2009). Band edge electronic structure of BiVO₄: Elucidating the role of the Bi s and V d orbitals. *Chemistry of Materials*, *21*, 547–551.
- Wang, X., Li, G., Ding, J., Peng, H., & Chen, K. (2012). Facile synthesis and photocatalytic activity of monoclinic BiVO₄ micro/nanostructures with controllable morphologies. *Materials Research Bulletin*, *47*, 3814–3118.

- Wang, J. L., Li, Y. Q., Byon, Y. I., Mei, S. G., & Zhang, G. L. (2013). Synthesis and characterization of NiTiO₃ yellow nano pigment with high solar radiation reflection efficiency. *Powder Technology*, 235, 303–306.
- Wang, J. T. W., Ball, J. M., Barea, E. M., Abate, A., Alexander-Webber, J. A., Huang, J., Saliba, M., Mora-Sero, I., Bisquert, J., Snaith, H. J., & Nicholas, R. J. (2014). Low-temperature processed electron collection layers of Graphene/TiO₂ nanocomposites in thin film Perovskite solar cells. *Nano Letters*, 14, 724–730.
- Wilcoxon, J. P., Samara, G. A., & Provencio, P. N. (1999). Optical and electronic properties of Si nanoclusters synthesized in inverse micelles. *Physical Review B*, 60, 2704–2714.
- Llabrés i Xamena, F. X., Damin, A., Bordiga, S., & Zecchina, A. (2003). Healing of defects in ETS-10 by selective UV irradiation: A Raman study. *Chemical Communications* 1514–1515.
- Xie, M., Feng, Y., Fu, X., Luan, P., & Jing, L. (2015). Phosphate-bridged TiO₂–BiVO₄ nanocomposites with exceptional visible activities for photocatalytic water splitting. *Journal of Alloys and Compounds*, 631, 120–124.
- Yanai, T., Tew, D. P., & Handy, N. C. (2004). A new hybrid exchange–correlation functional using the Coulomb-Attenuating method (CAM-B3LYP). *Chemical Physics Letters*, 393, 51–57.
- Yin, W. J., Wei, S. H., Al-Jassim, M. M., Turner, J., & Yan, Y. (2011). Doping properties of monoclinic BiVO₄ studied by first-principles density-functional theory. *Physical Review B*, 83, 155102.
- Yu, J., & Kudo, A. (2006). Effects of structural variation on the photocatalytic performance of hydrothermally synthesized BiVO₄. *Advanced Functional Materials*, 16, 2163–2169.
- Zhang, Y.-F., Lin, W., Li, Y., Ding, K.-N., & Li, J.-Q. (2005). A theoretical study on the electronic structures of TiO₂: Effect of Hartree-Fock exchange. *Journal of Physical Chemistry B*, 109, 19270–19277.
- Zhang, X., Ai, Z., Jia, F., Zhang, L., Fan, X., & Zou, Z. (2007). Selective synthesis and visible-light photocatalytic activities of BiVO₄ with different crystalline phases. *Materials Chemistry and Physics*, 103, 162–167.
- Zhao, Z., Li, Z., & Zou, Z. (2011). Electronic structure and optical properties of monoclinic clinobisvanite BiVO₄. *Physical Chemistry Chemical Physics*, 13, 4746–4753.

Alexey A. Popov

Contents

Introduction	1032
Structures and Stability of Empty Fullerenes	1034
Definition of Fullerenes and Enumeration of Their Isomers	1034
The Isolated Pentagon Rule and Steric Strain	1036
The Isomers of IPR Fullerenes	1042
Bonding, Structures, and Stability of Endohedral Metallofullerenes	1051
Metal–Cage Bonding in Endohedral Metallofullerenes	1051
Isomerism in Endohedral Metallofullerenes: Stability of the Charged Carbon Cages ...	1056
Isomerism in Endohedral Metallofullerenes: Cage Form Factor	1061
Violation of the Isolated Pentagon Rule in Endohedral Metallofullerenes	1065
Structures and Stability of Fullerene Derivatives	1066
Addition of X_2 to C_{60} : Isomers of $C_{60}X_2$ and General Considerations	1066
Addition of H and F to C_{60} : Contiguous Addition, Benzene Rings, and Failures of AM1	1068
Addition of Bulky Groups to C_{60} : Bromination and Perfluoroalkylation	1073
Addition to C_{70} and Higher Fullerenes	1078
Bibliography	1082

Abstract

This chapter describes general principles in the stability and bonding of empty fullerenes, endohedral fullerenes, and exohedral derivatives of empty fullerenes. First, an overview of the structural properties of empty fullerenes is given. The problem of isomers' enumeration is described and the origin of the intrinsic steric strain of the fullerenes is discussed in terms of POAV (π -orbital vector analysis) leading to the isolated pentagon rule (IPR). Finally, theoretical studies of the isomers of fullerenes are discussed. In the second part of the chapter,

A.A. Popov (✉)

Leibniz-Institute for Solid State and Materials Research (IFW Dresden), Dresden, Germany

e-mail: A.Popov@ifw-dresden.de

bonding phenomena and molecular structures of endohedral metallofullerenes (EMFs) are reviewed. First, the bonding situation in EMFs is discussed in terms of ionic/covalent dichotomy. Then, the factors determining isomers of EMFs, including those favoring formation of non-IPR cage isomers, are reviewed. In the third part, general principles governing addition of atomic addends and trifluoromethyl radicals to fullerenes are analyzed.

Introduction

Carbon cluster research, originally the field of scientists interested in C_n species that might exist in deep space, “landed” on Earth with two discoveries that eventually changed the face of carbon science in particular and the nanoscience/nanotechnology in general. The first discovery was reported in 1985, when Kroto, Curl, Smalley, and coworkers made the Nobel-prize-winning suggestion that the high-mass peaks in the mass spectra of laser-evaporated graphite belonged to closed-cage structures now known as fullerenes (Kroto et al. 1985). These molecules remained the playthings of mass spectrometrists, the only scientists who could “observe” fullerenes by producing them in situ in infinitesimal amounts, until 1990 when Krätschmer and coworkers discovered that this “new form of carbon” can be produced in bulk (i.e., macroscopic) amounts by the arc burning of graphite (Krätschmer et al. 1990). C_{60} followed by C_{70} is the major fullerenes produced by arc burning; however, starting from C_{72} , classical fullerene cages are possible for each even number, and many of such fullerenes (known as *higher fullerenes*) are produced as well. Studies of higher fullerenes are complicated by the increasing number of possible isomers, even if the strict isolated pentagon rule (IPR; see below for more details) is applied. The fullerenes with the number of atoms less than 60 are also available, but still remain exotic objects. Thus, the world of fullerenes themselves is very rich, but they also provide uncountable possibilities for further chemical modifications.

One of the attractive properties of fullerenes intrinsic to their closed-cage structure is the possibility of using them as robust containers for other species. The first evidence that metal atoms can be put inside fullerenes was reported in 1985 and was based on mass spectrometry (Heath et al. 1985), while the first bulk samples of fullerenes with metal atoms inside (endohedral metallofullerenes; hereinafter EMFs) were obtained in 1991 by laser evaporation (Chai et al. 1991) or arc discharge (Alvarez et al. 1991) of the graphite rods mixed with lanthanum oxide. In the first decade after the discovery of macroscopic fullerene production, the field of EMFs remained very “hot.” Many metals were put inside fullerenes during this time, and many new molecules were reported (Shinohara 2000; Popov et al. 2013), but the yields of EMFs were very low (usually on the order of 1 % of the empty fullerenes in the arc-discharge soot). Until 1999, EMFs were mostly molecules with one to three metal atoms encapsulated in the carbon cage. In 1999 it was found that the presence of nitrogen gas in the arc-burning reactor resulted in metal nitride cluster fullerenes (NCFs) with the composition $M_3N@C_{2n}$ ($M = Sc,$

Y, Gd–Lu; $2n = 68–96$; the symbol “@” in the formula denotes how the EMF molecule is divided into the endohedral cluster, the composition of which is given before this symbol, and the carbon cage with optional exohedrally attached groups, which follow after this symbol) to be produced (Dunsch and Yang 2007; Stevenson et al. 1999). In 2001 it was discovered that some “conventional” EMFs actually had C_2 carbide units inside the cage. For example, “ $Sc_3@C_{82}$ ” and “ $Sc_2@C_{86}$ ” were shown to be $Sc_3C_2@C_{80}$ and $Sc_2C_2@C_{84}$, respectively (Iiduka et al. 2005; Wang et al. 2001). Judicious choice of the nitrogen source led the group in Dresden to the invention of the reactive atmosphere method, in which the use of ammonia in the arc-burning reactor resulted in NCFs as major fullerene products (Dunsch et al. 2003, 2004). The same method with the use of CH_4 resulted in $Sc_3CH@C_{80}$ (Krause et al. 2007). Oxides $Sc_4O_{2,3}@C_{80}$ were synthesized using copper nitrate as the source of oxygen (Stevenson et al. 2008), while the use of guanidinium thiocyanate afforded formation of sulfide clusterfullerenes $M_2S@C_{82}$ ($M = Sc, Y, Dy, Lu$) (Dunsch et al. 2010). The chemical route to open the empty fullerene cage, insert the small molecule inside (e.g., H_2), and then close the carbon cage was also recently reported (Komatsu et al. 2005). Finally, it is also possible to introduce one or two atoms of inert gases into the fullerene by high-pressure–high-temperature treatment (Saunders et al. 1993). An extended review of the field was published in 2013 (Popov et al. 2013).

Another important property of fullerenes from the chemist’s point of view is a very rich π -system based on the formal sp^2 hybridization state of all carbon atoms. In chemical reactions fullerenes behave as polyalkenes (original expectations of the “superaromatic” properties were not confirmed) and thus exhibit very rich addition chemistry. That is, numerous cycloaddition reactions and addition of the groups forming one single bond to the fullerene core (such as atoms, e.g., halogens or hydrogen, or radicals, e.g., alkyl groups) are reported (Hirsch and Brettreich 2005).

Development of fullerene science was always accompanied – and sometimes preceded – by theoretical studies. The principles governing stability of fullerenes were revealed in the late 1980s and early 1990s with the use of computational approaches, and such important data as relative energies of the fullerene isomers are still available only from the results of quantum-chemical calculations. Likewise, the most mysterious question about the fullerenes – why and how they are formed – is successfully addressed by theoreticians (Irlé et al. 2003, 2006). The rise of the fullerene era to a large extent coincided with a dramatic increase in the capabilities of computational chemistry – either from advances in hardware, software, or theory itself. If studies of fullerenes in the mid-1980s, even by semiempirical approaches, were possible only in special laboratories, presently (at 2015), accurate DFT calculations can be done by virtually any scientist using a standard office computer. Yet, the large size of fullerenes and their derivatives still imposes serious limitations: even now it is barely possible to perform post-Hartree–Fock *ab initio* calculations. Thus, DFT is the main method of choice in fullerene chemistry and will probably remain so in years to come. Although MP2 calculations are becoming more feasible (Darzynkiewicz and Scuseria 1997; Haser et al. 1991; Krapp and Frenking 2007), a significant increase of computational demands in many cases is not justified by

the reliability of the results when compared to DFT (the modeling of noncovalent interaction, the weak point of DFT, is, however, an important exception). Routine use of methods with higher scaling than MP2 is still very difficult at present.

In this chapter we will show how quantum-chemical calculations can assist in studies of fullerenes. In particular, we will discuss prediction and elucidation of the molecular structures of the empty fullerenes, endohedral metallofullerenes, and selected fullerene derivatives. Taking into the account the huge amount of literature published on this subject, any review of the field is necessarily incomplete. We will not try to cover the whole field; instead, we will focus on several representative examples.

Structures and Stability of Empty Fullerenes

Definition of Fullerenes and Enumeration of Their Isomers

IUPAC defines a fullerene as a “compound composed solely of an even number of carbon atoms which form a cage-like fused-ring polycyclic system with twelve five-membered rings and the rest six-membered rings.” In practice, all other closed-cage structures built from three-coordinate carbon atoms are also called fullerenes. The definition of fullerenes is based on the Euler theorem, which states that for a given polyhedron, the numbers of vertices (n), edges (e), and faces (f) are related by:

$$n + f = e + 2 \quad (1)$$

Each carbon atom in a fullerene is bonded to three other atoms; therefore, for a fullerene C_n , the number of edges is $e = 3n/2$, which yields the number of faces $f = n/2 + 2$. Since all arc-discharge produced fullerenes have only five- and six-membered faces (it is possible to obtain fullerenes with four- or seven-membered rings only by chemical modification), one can calculate the number of vertices and faces through the number of pentagons (p) and hexagons (h):

$$\begin{aligned} (5p + 6h) / 3 &= n \\ p + h &= n/2 + 2 \end{aligned} \quad (2)$$

Solution of Eq. 2 yields $p = 12$ (and hence the number in the definition given above) and $h = n/2 - 10$. Obviously, the smallest structure with 12 pentagons is C_{20} (it has no hexagons), and starting from $n = 24$ at least one fullerene structure is possible for each even n (note that it is not possible to build a fullerene for C_{22}). The structural consequences from Euler's theorem applied to fullerenes with pentagons, hexagons, and heptagons were discussed in detail by Fowler and coworkers (Fowler et al. 1996). The difference between the number of pentagonal and heptagonal faces is always 12. Therefore, addition of a heptagon increases the number of pentagons by one, while decreasing the number of hexagons by two.

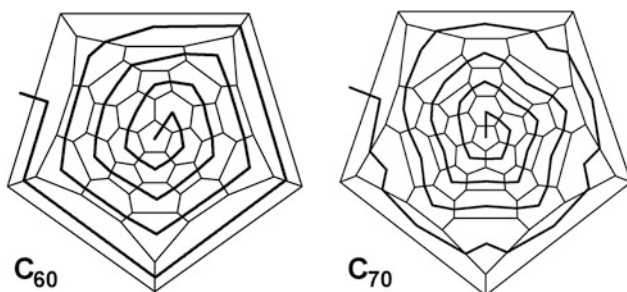


Fig. 1 Unwinding C_{60} and C_{70} into spirals (*thick black line*)

The number of the isomers is rapidly increasing with n , raising the problem of systematic enumeration of all possible isomers. The problem was solved for all practical fullerene sizes by the invention of the spiral algorithm (Manolopoulos et al. 1991; Fowler and Manolopoulos 1995). The algorithm is based on the spiral conjecture, which states that *The surface of a fullerene polyhedron may be unwound in a continuous spiral strip of edge-sharing pentagons and hexagons such that each new face in the spiral after the second share an edge with both (a) its immediate predecessor in the spiral and (b) the first face in the preceding spiral that still has an open edge* (Fowler and Manolopoulos 1995).

Figure 1 shows one of the ways in which C_{60} and C_{70} can be unwound into a spiral. The spiral can be coded by a sequence of “5” and “6” digits corresponding to the positions of the pentagons and hexagons. For instance, the spiral of the fullerene C_{60} shown in Fig. 1 can be written as:

$$566665656565656656565656566665 \quad (3)$$

Enumeration of all fullerene isomers is now straightforward: one should generate all possible spiral codes consistent with the given number of atoms (i.e., for C_n fullerene the length of the code is $f = n/2 + 10$, and positions of the pentagons can be used as running indices) and then check for each spiral if it can be wound up into a fullerene. The task, however, is not complete at this point since the fullerene molecule can be unwound into up to $6n$ spirals, and hence each fullerene will be generated many times. For instance, if C_{60} is unwound into a spiral starting from a hexagon, the following codes can be also obtained:

$$\begin{aligned} 656565666566566566566565656566 \\ 665656565665665665665665666565656 \end{aligned} \quad (4)$$

There is a simple way to solve the problem of checking the uniqueness of the given fullerene. One can consider the spiral codes given in Eqs. 3 and 4 as numbers, which can then simply be compared to each other. For the spirals of C_{60} under discussion,

(Eq. 3) < (Eq. 4) < (Eq. 4). The spiral with the smallest number is called a *canonical spiral*; all noncanonical spirals should be sorted out. Thus, in addition to checking whether a given spiral can be wound up into a fullerene, one should also check if it is canonical. This problem can be also solved straightforwardly: if it is found that the spiral forms a fullerene, this fullerene should then be unwound into all possible spirals and the canonical one should be found among them and compared to the testing spiral.

The algorithm outlined above has been proved to be very successful, and its use in enumeration of the fullerene isomers is now standard. Importantly, the sequential number of the given fullerene isomer (i.e., the number of the isomer in the list of all isomers generated for the given fullerene by the spiral algorithm) is its sufficient and unique identifier. Usually, in designation of the fullerene isomer, one also adds its symmetry (to be precise, the highest symmetry possible for this topology of atoms; the ground state structure of the molecule can have lower symmetry), and we will use this notation in this chapter.

To conclude this section, we note that the spiral codes in Eqs. 3 and 4 can be further simplified, taking into account that each fullerene has only 12 pentagons, and hence the spiral code can be uniquely defined by the numbers of the pentagon positions. In this notation, the spiral of C_{60} in Eq. 3 is written as:

$$1\ 7\ 9\ 11\ 13\ 15\ 18\ 20\ 22\ 24\ 26\ 32 \quad (5)$$

This notation is usually used in the literature, including the famous “An Atlas of Fullerenes” (Fowler and Manolopoulos 1995).

The Isolated Pentagon Rule and Steric Strain

The importance of exhaustive systematic enumeration of all isomers for a given fullerene becomes clear when the problem of structure elucidation of fullerenes is raised. Highly symmetric structures of C_{60} and C_{70} were proposed as a result of playing with the models. However, the success of this approach was to a large extent fortuitous, and this is hardly possible for other fullerenes. Of course, an unambiguous elucidation of the molecular structure might be possible through single-crystal X-ray diffraction studies. However, the quasi-spherical shape of fullerene molecules results in strong disorder, which in most cases precludes reliable elucidations of the isomeric structures. This problem can be solved by chemical functionalization of fullerenes (in particular, chlorination and perfluoroalkylation of higher fullerenes have been extensively studied during the last decade) and X-ray studies of their derivatives (Kareev et al. 2008a; Shustova et al. 2007; Simeonov et al. 2007; Troyanov and Tamm 2009; Troyanov and Kemnitz 2012; Boltalina et al. 2015), but this promising approach is not always applicable. Thus, since early 1990s and up to now, ^{13}C NMR spectroscopy has been the most important method in structural studies of newly isolated fullerenes. However, through the number of lines and their relative intensities, ^{13}C NMR spectrum gives at best only the symmetry of

Table 1 The total number of isomer of C₆₀–C₁₀₂ fullerenes and the number of fullerene C₉₀ isomers of different symmetry types compared to the number of corresponding IPR isomers

C _n	Total	IPR	C _n	Total	IPR	C ₉₀ -symmetry group	Total	IPR
60	1812	1	86	63761	19	C ₉₀ -C ₁	97936	16
70	8149	1	88	81738	35	C ₉₀ -C ₂	1266	16
72	11190	1	90	99918	46	C ₉₀ -C _s	655	6
74	14246	1	92	126409	86	C ₉₀ -C ₃	3	0
76	19151	2	94	153493	134	C ₉₀ -C _{2v}	50	7
78	24109	5	96	191839	187	C ₉₀ -D ₃	4	0
80	31924	7	98	231017	259	C ₉₀ -D ₅	1	0
82	39718	9	100	285914	450	C ₉₀ -D _{3h}	1	0
84	51592	24	102	445092	616	C ₉₀ -D _{5h}	2	1

the fullerene molecule, but not the exact isomeric pattern. Therefore, it is necessary that *all* isomers with NMR-determined symmetry are considered as possible structural guesses. Quantum-chemical calculations and studies by other spectroscopic techniques can be then used to determine the most appropriate structure.

Table 1 lists the total amount of possible isomers of fullerenes C_n for different *n*. These numbers clearly show that consideration of *all* isomers is hardly feasible for the low-symmetric structures. For instance, 99918 isomers are possible for C₉₀, and the information that an experimentally studied isomer has C₂ symmetry allows one to reduce the structural guesses to 1266 isomers, which is still not very feasible. On the other hand, the number of isomers formed in arc-discharge synthesis is usually rather low. Therefore, it seems possible and highly desirable to develop some simple rules to reduce the number of isomers that should be considered.

The most successful rule of this sort is the isolated pentagon rule (IPR) proposed by H. Kroto back in 1987 on the basis of general considerations (Kroto 1987) and by Schmalz and coworkers in 1988 on the basis of more careful considerations of the sources of strain in carbon clusters and Hückel calculations (Schmalz et al. 1988). The IPR rule states that the fullerene isomers with adjacent pentagons are less stable than the isomers in which pentagons are surrounded only by hexagons (i.e., pentagons are *isolated*). In subsequent works numerous isomers of fullerenes were studied and the number of the pentagon–pentagon edges was found to be a parameter exhibiting very good correlation to the stability of the fullerenes. Figure 2a shows the correlation between the number of pentagon–pentagon edges and the relative energies of all 1812 isomers of C₆₀ calculated at the QCFF/PI level of theory (Austin et al. 1995) (QCFF/PI is an inexpensive semiempirical computational approach combining classical mechanics for σ-bonds and the Pople–Pariser–Parr method for π-system). The correlation is close to the linear form, and the energy penalty for one pentagon–pentagon edge, 111 kJ/mol for C₆₀, can be estimated from the slope of the fitted line. In a detailed study of the C₃₀–C₆₀ fullerene isomers at the QCFF/PI level (Campbell et al. 1996), it was found that the penalty is increasing with the cage size (Fig. 2b), from 72 kJ/mol for C₃₀ to 105 kJ/mol for C₅₀. The study of 55 C₇₆ isomers of D₂ and higher symmetry at the QCFF/PI level (Austin et al. 1994) has shown that

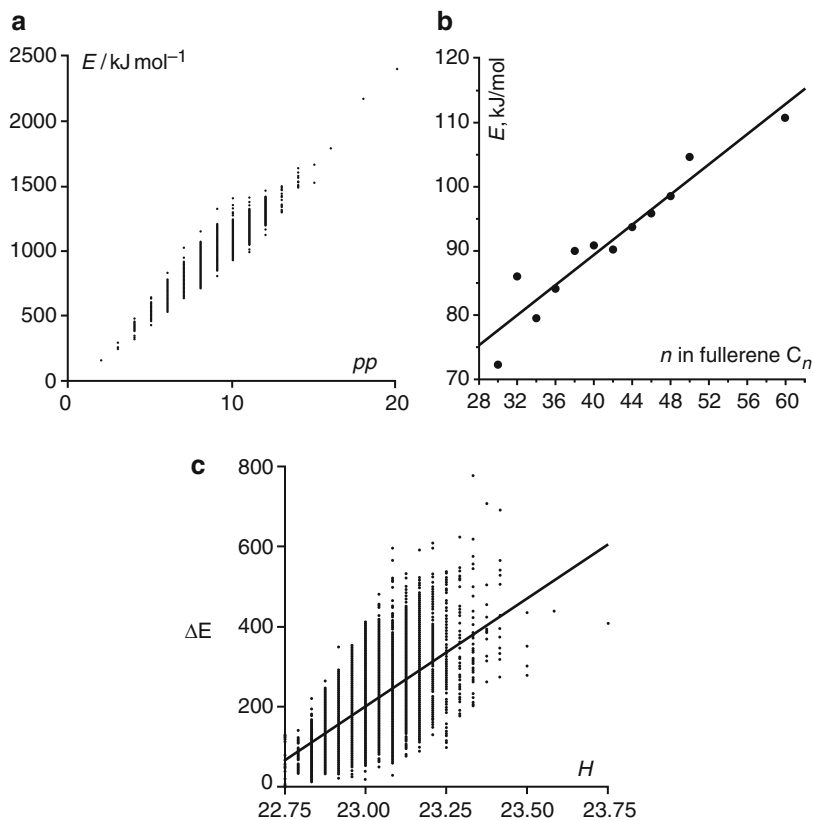


Fig. 2 (a) The relative energy of C_{60} isomers computed at the QCFF/PI level as a function of the number of pentagon–pentagon edges (Reproduced with permission from (Austin et al. 1995), © (1995) American Chemical Society); (b) energy penalty for the pentagon adjacency as a function of the fullerene size (Based on the data from Campbell et al. 1996); (c) relative energy of the IPR isomers of C_{116} computed at the QCFF/PI level as a function of the steric strain parameter H (Reproduced with permission from (Achiba et al. 1998a), © (1998) American Chemical Society)

the penalty of the pentagon adjacency can be estimated as 150 kJ/mol (but this value is probably overestimated since the non-IPR isomers with lower symmetry can be more stable). The study of all 40 isomers of C_{40} at various levels of theory, including DFT (B3LYP, BLYP, and LDA), the Hartree–Fock method, semiempirical methods (AM1, PM3, MNDO, QCFF/PI), the density-functional based tight-binding method (DFTB), and molecular mechanics (MM), has shown that for all methods except for MM, the penalty for C_{40} isomers falls in the narrow range of 82–100 kJ/mol (Albertazzi et al. 1999). In particular, B3LYP/STO-3G and QCFF/PI values are 92 and 91 kJ/mol, respectively, which confirms that the predictions on the broader range of isomers obtained at the QCFF/PI levels are also reliable. It should be noted, however, that the increase of the basis set to 6-31G* resulted in a 13 % increase

in the penalty at the BLYP level, showing that the basis set effects can be rather important.

The energy penalty for the pentagon adjacency is sufficiently large to suggest that all non-IPR isomers can be sorted out for higher fullerenes (relative energies of the isolable fullerene isomers are usually less than 100 kJ/mol; see below). Increase of the energy penalty with the cage size is another factor favoring the application of the IPR. Indeed, the rule is strictly followed by all experimentally available empty fullerenes (the notable exclusion is C_{72} , which is discussed below). For the experimentally accessible fullerenes, the IPR reduces the number of possible isomers to be considered from many thousands to tens and hundreds (Table 1). It is, however, possible to stabilize pentagon adjacencies by exohedral derivatization or by coordination to the metal atoms in EMFs (Tan et al. 2009), and hence larger sets of isomers have to be considered in these cases.

High energy penalty is the result of the combination of at least two factors. First, a pair of adjacent pentagons (pentalene) has an 8-electron π -system and is anti-aromatic. Another factor, which is probably more important, is the steric strain induced by adjacent pentagons. The formal hybridization state of carbon in fullerenes is sp^2 . Thus, carbon atoms should be planar (by this, we mean that it has three neighboring carbon atoms in the same plane). From the geometrical point of view, planarity is achieved when a carbon atom is located on the fusion of three hexagons. In this case, π -orbitals (p_z) of the neighboring atoms are exactly parallel and the optimal overlap of these orbitals is achieved. However, one cannot build fullerene only from hexagons – pentagons are necessary to make closed-cage structures. The carbon atoms in pentagons are not planar anymore (the small angle of 108° vs. 120° in hexagons has to be compensated by the out-of-plane shift of the central atom), which results in *rehybridization*, the effect easily explainable in terms of π -orbital axis vector (POAV) analysis. POAV is defined as the vector that forms equal angles (POAV angles $\theta_{\sigma\pi}$) to all three σ -bonds (it is assumed that σ -orbitals are aligned toward the neighboring atoms). Admixing of s -orbitals into the hybrid π -orbital (the essence of rehybridization) can be quantified from geometrical point of view as:

$$m = \frac{2\sin^2(\theta_{\sigma\pi} - \pi/2)}{1 - 3\sin^2(\theta_{\sigma\pi} - \pi/2)} \quad (6)$$

where m determines admixture of s -orbital to the hybrid π -orbital ($\pi = s^m p$; for perfect sp^2 hybridization, $\theta_{\sigma\pi} = 90^\circ$ and $m = 0$). When m is not equal to zero, π -orbitals of the neighboring atoms are not parallel anymore. Because of this misalignment, the π -orbital overlap is reduced, and therefore the π -bonding is weakened, the effect being the stronger the larger the value of m . Figure 3b shows possible junctions of pentagon and hexagons and hybrid π -orbital of the central atom computed for idealized polygonal bond angles. Misalignment of π -orbitals increases with the increase of the number of fused pentagons, and the IPR is a manifestation of this tendency.

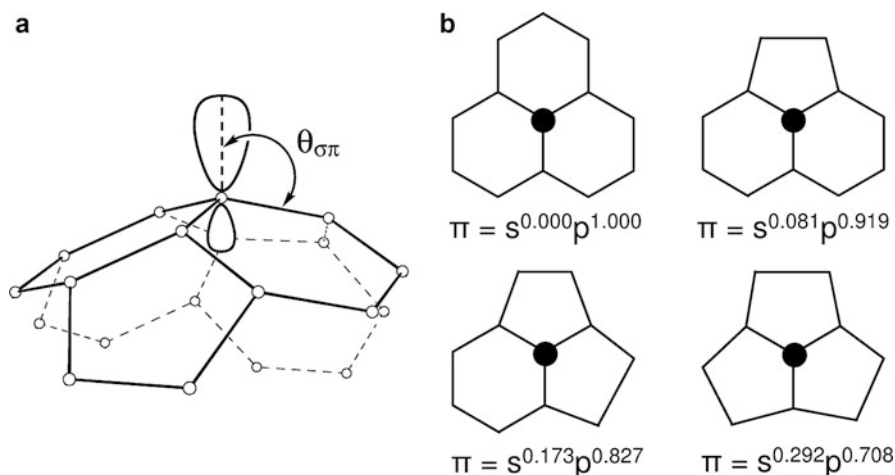


Fig. 3 (a) Definition of the $\theta_{\sigma\pi}$ angle in the POAV approach; (b) the component of the hybrid π -orbital of the central atom in different junctions of pentagons and hexagons computed for idealized polygonal bond angles (Based on the data from Fowler and Manolopoulos 1995)

It is obvious from Fig. 3 that all fullerenes are inherently strained. For instance, it was estimated that the steric strain constitutes approximately 80 % of the excess energy of C_{60} with respect to graphite (Haddon 1993). To quantify the steric strain, Raghavachari introduced the concept of hexagonal indices (Raghavachari 1992). The neighbor index k of a given hexagon is the number of its edges which are shared with other hexagons, and every fullerene isomer can be characterized by a set of indices h_k ($k = 0-6$), where h_k is the number of hexagons with neighbor index k . In the aforementioned study of 1812 isomers of C_{60} , it was shown that the parameter

$$H = \sum_k k^2 h_k \quad (7)$$

also gives good linear correlation with the relative energies of the isomers.

The concept of hexagonal indices was in fact invented to explain relative stabilities of the IPR isomers. As each hexagon in the IPR isomer is adjacent to at least three other hexagons, h_0 , h_1 , and h_2 indices are equal to 0, and the combination of four indices (h_3 , h_4 , h_5 , h_6) is sufficient to characterize hexagon adjacencies in a given fullerene isomer (Fowler and Manolopoulos 1995; Raghavachari 1992). Raghavachari suggested that to minimize the steric strain, the indices of all hexagons should be as close to each other as possible. In other words, the lowest strain energy can be achieved when the non-planarity is uniformly distributed over the whole carbon cage. Hence, the lowest strain is expected for those structures, in which all hexagons have the same neighborhood (all indices are equal), the condition fulfilled only for highly symmetric $C_{60}-I_h(1)$, $C_{80}-I_h(7)$, and $C_{80}-D_{5h}(6)$ in the whole range of IPR fullerenes C_{2n} with $2n < 120$ (Fowler and Manolopoulos 1995). More complex conditions were derived for other IPR fullerenes, namely:

Table 2 Optimal hexagon indices, the IPR isomers with optimal hexagon indices, and the lowest energy IPR C_{2n} and C_{2n}^{6-} isomers ($2n = 76-90$)

C_{2n}	(h_3, h_4, h_5, h_6)	The least-strain IPR isomers ^a	The lowest energy C_{2n} IPR isomers ^b	The lowest energy C_{2n}^{6-} IPR isomers ^c
C_{76}	(4,24,0,0)	$T_d(2)$	$D_2(1)$	$T_d(2)$
C_{78}	(2,27,0,0)	$D_{3h}(5)$	$C_{2v}(3)$	$D_{3h}(5)$
C_{80}	(0,30,0,0)	$D_{5h}(6), I_h(7)$	$D_{5d}(1), D_2(2)$	$D_{5h}(6), I_h(7)$
C_{82}	(0,29,2,0)	$C_{2v}(9)$	$C_2(3)$	$C_{2v}(9)$
C_{84}	(0,28,4,0)	$D_2(21), D_2(22), D_{2d}(23)$	$D_2(22), D_{2d}(23)$	$D_2(21)$
C_{86}	(0,27,6,0)	$D_3(19)$	$C_2(17)$	$D_3(19)$
C_{88}	(0,26,8,0)	$D_2(35)$	$C_s(17)$	$D_2(35)$
C_{90}	(0,25,10,0)	$C_2(40), C_2(41), C_2(43), C_2(44), C_2(45)$	$C_2(45)$	$C_2(43)$

^aBased on the data from (Fowler and Manolopoulos 1995)

^bSee Table 3 for more details

^cBased on the data from (Popov and Dunsch 2007)

$$(h_3, h_4, h_4, h_6) = \begin{cases} (80 - n, 3n/2 - 90, 0, 0); \rightarrow 60 \leq n \leq 80 \\ (0, 70 - n/2, n - 80, 0); \rightarrow 80 \leq n \leq 140 \\ (0, 0, 60, n/2 - 70); \rightarrow 140 \leq n \end{cases} \quad (8)$$

Table 2 lists optimal combinations of hexagonal indices as well as fullerene isomers in the range of C_{70} - C_{90} with such sets of indices. Importantly, C_{72} is the only fullerene that has no isomers satisfying these conditions.

As a numerical parameter to quantify the steric strain in IPR fullerenes, one can use the standard deviation of the index distribution (Fowler and Manolopoulos 1995):

$$\sigma_h = \sqrt{\langle k^2 \rangle - \langle k \rangle^2} \quad (9)$$

where

$$\langle k^i \rangle \Big|_{i=1,2} = \left(\sum_{k=3}^{k=6} k^i h_k \right) \cdot \left(\sum_{k=3}^{k=6} h_k \right)^{-1} \quad (10)$$

It was also shown that σ_h correlates well with the standard deviation of the square of POAV pyramidalization angle, $\sigma(\theta_p^2)$ ($\theta_p = \theta_{\sigma\pi} - 90$), as can be seen in Fig. 4a for a set of 24 IPR isomers of C_{84} (Boltalina et al. 2009). Therefore, σ_h and $\sigma(\theta_p^2)$ can be used interchangeably to quantify the steric strain, the latter parameter being more universal (for instance, it can be equally used for fullerene derivatives, in which some carbon atoms are in sp^3 hybridization state). Figure 4b also plots the relative energies (ΔE) of C_{84} computed at the B3LYP/6-31G* level of theory versus $\sigma(\theta_p^2)$.

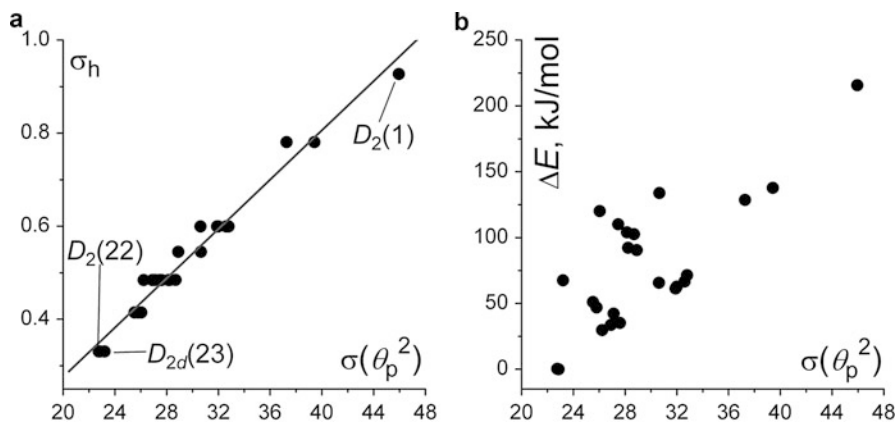


Fig. 4 (a) Correlation between $\sigma(\theta_p^2)$ and σ_h parameters for 24 IPR isomer of C_{84} (Based on the data from Boltalina et al. 2009); for a fitted line, $R^2 = 0.97$; (b) Correlation between $\sigma(\theta_p^2)$ and B3LYP/6-31G*-computed relative energies of 24 IPR isomer of C_{84} (Based on the data from Sun and Kertesz 2001b)

The lack of correlations between ΔE and $\sigma(\theta_p^2)$ for the isomers with medium values of these parameters indicates that for IPR fullerenes, the other factors (e.g., electronic) are at least equally important as the steric strain. As a result, the lowest energy isomers for C_{76} – C_{90} are usually not the isomers with the lowest steric strain (see Table 2), and steric strain analysis is not widely used now in the studies of higher fullerenes of medium size. If a larger set of isomers is considered (e.g., 6,063 IPR isomers C_{116}), then the steric strain indicators (such as H in Eq. 7) indeed show correlation with the relative energies, at least at the QCFF/PI level (Fig. 2c; Achiba et al. 1998a). However, the standard deviation of 75 kJ/mol is still too large, and such data can be of help only to sort out some high-energy isomer. Below we will show that more precise schemes have been developed based on the combination of topological motifs and their increments (Cioslowski et al. 2000).

The Isomers of IPR Fullerenes

With the advent of the IPR, the number of fullerene isomers to be considered was drastically reduced, and detailed computations of reasonable (i.e., IPR) isomers and comparison to the experimental results became feasible. In this section we will review the results of the search of the most stable isomers of fullerenes beyond C_{70} (starting from C_{70} , at least one IPR isomer is possible for each even number of atoms). It is necessary to point out that the term *stability* comprises some ambivalence. It can be used in terms of *thermodynamic* stability, which in the studies of isomers means low relative energy, or in terms of *kinetic* stability, which implies reactivity of the compound in the given experimental conditions (in the studies of fullerenes, it is usually determined by the HOMO–LUMO gap).

When predicting the possible isomeric structure of fullerenes, it is necessary to consider a combination of both thermodynamic and kinetic stability factors, the former determining the possibility of the formation of the given fullerene and the latter showing whether a given fullerene can be extracted from the soot and further processed by standard fullerene separation techniques. In fact, conclusions based on thermodynamic stability can be valid only when the reaction is close to the equilibrium (which is hard to prove for fullerenes). Moreover, for the equilibrium, Gibbs energies (rather than relative energies) should be taken into account, and it was shown by Slanina and coworkers that the equilibrium composition can be significantly altered at high temperatures (Slanina et al. 2004b). On the other hand, there are some examples (see below) showing that if thermodynamically stable fullerene has small HOMO–LUMO gap, it is possibly produced but remains in the insoluble part of the soot after extraction (presumably because of its polymerization). In this respect, a combination of the relative energy and the HOMO–LUMO gap still seems to be convenient parameters for the basic characterization of the fullerene isomers. Table 3 lists the data on the relative energies of some higher fullerene isomers obtained using B3LYP and PBE0 (also known as PBE1PBE) hybrid DFT functionals. B3LYP results are compiled from several works by different groups and are obtained mainly with the use of 6-31G or 6-31G* basis sets (Chen et al. 2001; Slanina et al. 2000; Sun and Kertesz 2000, 2001b, 2002; Sun 2003a, b; Zhao et al. 2003), while PBE0 values are results of single-point calculations with 6-311G* basis set and DFTB-optimized atomic coordinates (Shao et al. 2006, 2007). Table 3 also lists PBE0-obtained HOMO–LUMO gaps. Figure 5 shows the structures of the major isomers of C_{60} and C_{70} – C_{84} .

The first fullerene after C_{70} , C_{72} , is a “missing” fullerene since it has never been obtained in considerable amounts. The sole IPR isomers of C_{72} with D_{6d} molecular symmetry are predicted to have large HOMO–LUMO gap (1.42 eV at the PBE/TZ2P level), however, DFT calculations (Kobayashi et al. 1997; Kobayashi and Nagase 2002b; Slanina et al. 2004a) supported also by the measurements of electron affinity of C_{72} (Boltalina et al. 2000a) show that trace amounts of C_{72} present in the fullerene extract are most probably based on the non-IPR C_{2v} (11188) isomer, which is 46–48 kJ/mol more stable than the IPR D_{6d} (1) isomer at the DFT level of theory (Table 3). Thus, C_{72} is the only empty fullerene C_n ($n > 70$) which violates IPR.

C_{74} represents to some extent the opposite situation to C_{72} : according to DFT calculations, the sole IPR isomer of C_{74} , D_{3h} (1), has small HOMO–LUMO gap and probably has a triplet ground state (Kovalenko and Khamatgalimov 2003). This fullerene was also thought to be “missing” for a long time, until it was found that it is insoluble in common fullerene solvents (Karataev 1998), most probably because of polymerization. C_{74} as well as some other “insoluble” fullerenes can be solubilized by either electrochemical (Diener and Alford 1998) or chemical modification (Goryunkov et al. 2004b; Shustova et al. 2006, 2007). For instance, trifluoromethylation of “insoluble” fullerene mixture afforded $C_{74}(\text{CF}_3)_{12}$ as a major product and D_{3h} (1) carbon cage in this compound was unambiguously confirmed by single-crystal X-ray diffraction (Shustova et al. 2007).

Table 3 Relative energies (kJ/mol) and HOMO–LUMO gap (eV) in selected isomers of C₇₂–C₉₄ fullerenes as computed at the B3LYP and PBE0//DFTB levels of theory

C _n	Sym.	B3LYP	PBE0	Gap	Ref. ^a	C _n	Sym.	B3LYP	PBE0	Gap	Ref. ^a
C ₇₂	<i>D</i> _{6d} (1)	0.0	0.0	2.64	[b,c]	C ₈₆	<i>C</i>₂(17)	0.0	0.0	1.58	[b,d]
	<i>C</i> _{2v}	−48.1	−46.1	1.63			<i>C</i>₅(16)	25.9	23.6	1.97	
C ₇₄	<i>D</i>_{3h}(1)	0.0	0.0	0.74	[b,c]		<i>C</i> ₁ (12)	43.1	43.5	1.24	
C ₇₆	<i>D</i>₂(1)	0.0	0.0	2.10	[b,c]		<i>C</i> ₁ (11)	43.5	45.9	1.17	
	<i>T</i> _d (2)		91.2	0.71			<i>C</i> ₃ (18)	47.3	46.7	1.27	
C ₇₈	<i>C</i>_{2v}(3)	0.0	0.0	1.71	[b,c]	C ₈₈	<i>C</i>₅(17)	0.0	0.0	1.60	[e,d]
	<i>D</i>_{3h}(5)	18.8	22.7	1.55			<i>C</i>₂(33)	9.6	8.9	1.80	
	<i>C</i>_{2v}(2)	27.6	26.3	2.15			<i>C</i>₂(7)	10.9	5.1	1.63	
	<i>D</i>₃(1)	41.4	41.3	1.71			<i>C</i> ₁ (11)	43.5			
	<i>D</i> _{3h} (4)	102.9					<i>C</i> ₂ (20)	47.7	45.3	1.40	
C ₈₀	<i>D</i>₂(2)	0.0	0.0	1.39	[b,c]	C ₉₀	<i>C</i>₂(45)	0.0	0.0	1.74	[f,d]
	<i>D</i>_{5d}(1)	9.2	1.1	1.15			<i>C</i>_{2v}(46)	14.2	12.8	1.87	
	<i>C</i> _{2v} (3)	11.7	17.1	0.95			<i>C</i>₅(35)	15.1	12.0	1.91	
	<i>C</i>_{2v}(5)	27.7					<i>C</i>₁(30)	28.5	27.1	1.81	
	<i>D</i> ₃ (4)	37.2					<i>C</i>₂(28)	34.3	32.4	1.75	
C ₈₂	<i>C</i>₂(3)	0.0	0.0	1.69	[b,d]		<i>C</i> ₂ (40)	36.8	39.6	1.45	
	<i>C</i>₅(4)	16.2	17.8	1.63			<i>C</i>₁(32)	38.5			
	<i>C</i> ₅ (2)	27.4	27.3	1.69			<i>C</i> ₂ (18)	46.9	43.6	1.52	
	<i>C</i> ₂ (1)	32.1	33.2	1.28		C ₉₂	<i>D</i> ₃ (28)	51.4	0.0	2.20	[g,d]
	<i>C</i>₂(5)	34.1	38.5	1.33			<i>D</i> ₂ (84)	0.0	20.5	1.98	
C ₈₄	<i>C</i> ₅ (6)	51.1	57.0	1.14			<i>C</i> ₂ (26)	102.6	22.2	1.73	
	<i>D</i>_{2d}(23)	0.0	0.0	2.15	[h,d]		<i>C</i>₁(38)	51.1	22.2	1.89	
	<i>D</i>₂(22)	0.4	2.5	2.07			<i>D</i>₂(82)	84.9	24.1	1.76	
	<i>D</i>_{6h}(24)	29.7	30.3	2.46			<i>D</i>₂(81)	134.7	40.5	1.61	
	<i>C</i>₅(16)	33.7	34.5	1.96		C ₉₄	<i>C</i>₂(43)	15.5	0.0	1.91	[i,d]
	<i>C</i>₂(11)	35.2	36.4	1.70			<i>C</i>₅(42)	15.1	8.3	1.99	
	<i>D</i>_{3d}(19)	42.3					<i>C</i> ₅ (44)	35.1	25.2	1.74	
	<i>C</i> ₅ (15)	47.0	50.6	1.60			<i>C</i>₂(133)	0.0	29.4	1.66	
	<i>C</i> ₁ (12)	51.1					<i>C</i>₂(61)	24.7	33.4	1.63	
	<i>D</i>_{2d}(4)	61.3					<i>C</i>₁(37)	64.0	33.8	1.71	
	<i>C</i>₅(14)	62.7					<i>C</i>₁(34)	25.9	35.0	1.55	
<i>C</i>_{2v}(18)	65.4					<i>C</i> ₁ (15)	33.9	43.6	1.44		
<i>D</i>₂(5)	66.4					<i>C</i> ₁ (91)		44.7	1.50		

^aIn each pair of references, the first one is for B3LYP values, the second one is for PBE0/6-311G*//DFTB values; experimentally characterized fullerene isomers are listed in bold

^bB3LYP/6-31G* (Chen et al. 2001)

^cShao et al. 2007

^dShao et al. 2006

^eB3LYP/6-31G(Sun 2003a)

^fB3LYP/6-31G(Sun 2003b)

^gB3LYP/6-31G*//SAM1(Slanina et al. 2000)

^hB3LYP/6-31G* (Sun and Kertesz 2001b)

ⁱB3LYP/6-31G*//SAM1(Zhao et al. 2003)

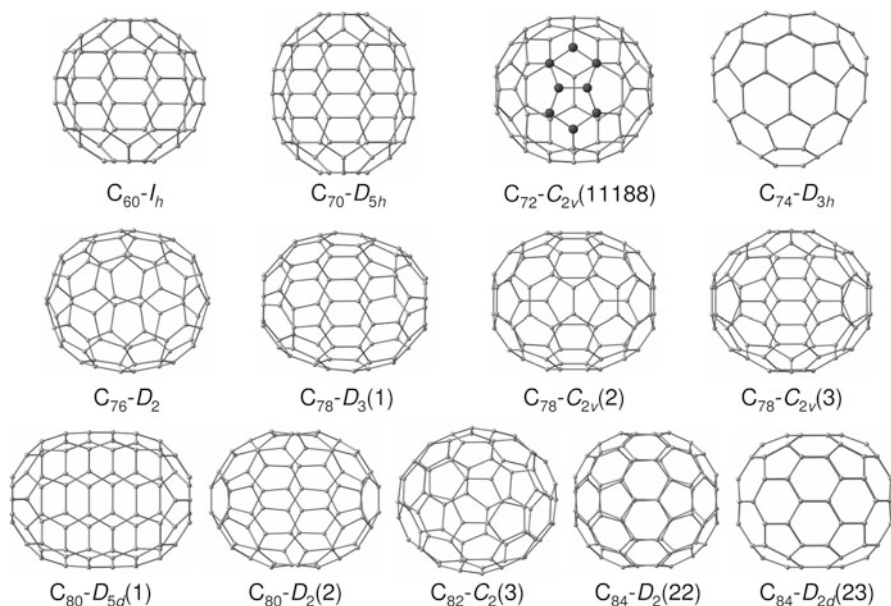


Fig. 5 Molecular structures of C_{60} , C_{70} , and selected higher fullerenes. The atoms in adjacent pentagon pair in C_{72} are highlighted

Identity of C_{76} as $D_2(1)$ isomer was unambiguously confirmed by ^{13}C NMR spectroscopy in the very early studies of the higher fullerenes (Ettl et al. 1991). According to DFT calculations, $D_2(1)$ is considerably more stable than the second IPR isomer, $T_d(2)$. Moreover, the latter is subject to Jahn–Teller distortion (the actual symmetry is D_2) and has a small HOMO–LUMO gap. Recently, it was shown that C_{76} – $T_d(2)$ is probably also formed in the fullerene synthesis but remains insoluble; however, it can be solubilized and characterized in the form of trifluoromethyl derivative (C_{76} – $T_d(2)$)(CF_3)₁₂ (Shustova et al. 2006).

C_{78} has five IPR isomers; four of them are quite stable (their ΔE covers the range of 41 kJ/mol) and are proved to be formed in the arc-discharge synthesis. Three isomers, $D_3(1)$, $C_{2v}(2)$, and $C_{2v}(3)$, are obtained in approximately comparable amounts (the actual ratio depends on the pressure used in the synthesis) and can be easily extracted from the soot as they have comparably high HOMO–LUMO gaps (Kikuchi et al. 1992; Wakabayashi et al. 1994). The isomer $D_{3h}(5)$, although it is the second in the stability order (Table 3), can be found in the fullerene extract in much smaller amounts (Simeonov et al. 2008). A possible reason is the low HOMO–LUMO gap, which results in the poor solubility. Indeed, this isomer was found among the “insoluble” fullerenes, its structure being unambiguously confirmed by the single-crystal X-ray structure of the trifluoromethyl derivative, (C_{78} – $D_{3h}(5)$)(CF_3)₁₂ (Shustova et al. 2007).

From seven IPR isomers of C_{80} , only the two most stable structures (Table 3), $D_2(1)$ and $D_{5d}(2)$, are obtained in the form of empty fullerenes (Hennrich et al.

1996; Wang et al. 2000b). DFT predicts that they are almost isoenergetic (the $D_2(1)$ isomer is somewhat more stable), but $D_{5d}(2)$ isomer has considerably smaller HOMO–LUMO gap. Besides, the higher symmetry of $D_{5d}(2)$ isomer also reduces its content if entropy factors at higher temperatures are taken into account. As a result, the yield of $D_2(1)$ is about 30 times higher (Wang et al. 2000b). Isolation of $(C_{80}-C_{2v}(5))(CF_3)_{12}$ after trifluoromethylation of the “insoluble” fullerene mixture has shown that $C_{80}-C_{2v}(5)$ is also produced in the arc-discharge synthesis (Shustova et al. 2006). $C_{2v}(5)$ isomer is rather stable but has small HOMO–LUMO gap ($\Delta E = 25$ kJ/mol, gap = 0.71 eV at the B3LYP/6-31G* level). Importantly, icosahedral $C_{80}-I_h(7)$ isomer, which should have the least steric strain among all IPR isomers of C_{80} , is in fact the least stable isomer ($\Delta E = 114$ kJ/mol at the B3LYP/6-31G* level of theory (Slanina et al. 2008). The reason is its electronic instability – in the ideal icosahedral symmetry, its HOMO is fourfold degenerate and is occupied by two electrons. BP/TZVP studies have shown that it has D_{2h} symmetry in the ground state, and its singlet and triplet states are isoenergetic (Furche and Ahlrichs 2001). However, this isomer becomes the most stable C_{80} isomer in the hexaanionic state, when its HOMO is fully occupied, which makes it a key structure for the endohedral metallofullerenes (see section “Bonding, Structures, and Stability of Endohedral Metallofullerenes” of this chapter).

C_{82} has seven IPR isomers, but only one of them, $C_2(3)$, was isolated in pure form (Kikuchi et al. 1992; Zalibera et al. 2007). $C_2(3)$ is the most stable isomer of C_{82} and has large HOMO–LUMO gap (Sun and Kertesz 2001a), so its high abundance is not surprising. Meanwhile, $C_s(4)$ isomer is also rather stable and has similar gap, but this isomer has never been isolated in a pure form. However, its presence in the fullerene soot was unambiguously confirmed by the characterization of its chloro-derivative $(C_{82}-C_s(4))Cl_{20}$ (Yang et al. 2013). The studies of the “insoluble” fullerene fraction have shown that $C_2(5)$ isomer is also formed in the arc-discharge synthesis, as verified by the isolation and characterization of its CF_3 derivative, $(C_{82}-C_2(5))(CF_3)_{12}$ (Shustova et al. 2006).

Starting from C_{84} , the increase in the number of possible structures becomes quite significant (see Table 1), which seriously complicates the studies, both experimentally and theoretically. This fact also increases the value of theoretical studies, since they can help in sorting out unrealistic isomers. The question of high importance in this situation is how to handle the rapidly increasing number of isomers. The obvious strategy of choice is to use several (usually two) computational approaches, first the “cheap” one (for instance, semiempirical method) for the whole set of isomers and then the more expensive but more reliable approach (usually DFT) for a reduced set of structures. It is important that the “cheap” method should be also sufficiently reliable so that the possible errors in the relative energies should be not very high, and a reasonable cutoff in the relative energies might be applied for the subsequent higher-level computations. Twenty-four IPR isomers of C_{84} form a convenient training set, since it is still feasible to perform a full DFT study for all of them. Figure 6 compares relative energies of 24 isomers obtained at the B3LYP/6-31G* level (Sun and Kertesz 2001b) to those obtained with three semiempirical approaches widely used in the fullerene chemistry, QCFF/PI, MNDO, and DFTB.

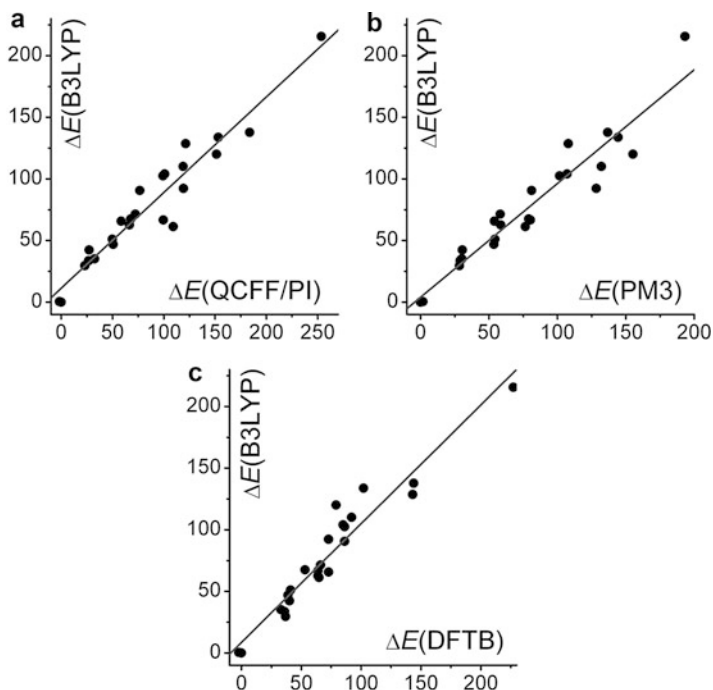


Fig. 6 Correlations between relative energies of 24 IPR isomer of C_{84} computed at the B3LYP/6-31G* level and those computed using QCFF/PI (a), PM3 (b), and DFTB (c), all values are in kJ/mol. R^2 values for the fitted lines are 0.925, 0.912, and 0.925, respectively

One can see that in all cases the correlation is quite reasonable ($R^2 \approx 0.92$), and the standard deviations (σ_E) are less than ~ 20 kJ/mol. Roughly estimating the energy cutoff for the isomers computed by means of “cheap” method as $3\sigma_E$, one can consider then only the isomers with $\Delta E < 3\sigma_E$ for the subsequent DFT study (the cutoff should be appropriately increased, of course, if one is interested in the whole set of isomers within the given energy window). Assessment study of three semiempirical approaches, MNDO, AM1, and PM3, in predicting relative energies of fullerene isomers in comparison to B3LYP/6-31G* was reported for a set of fullerenes including C_{30} – C_{50} , all IPR isomers of C_{72} – C_{86} , and selected isomers of C_{88} – C_{102} (Chen and Thiel 2003). Quite good correlations were found between the results of DFT and semiempirical calculations for higher fullerenes ($R^2 = 0.92$, 0.88, and 0.86, for MNDO, AM1, and PM3, respectively; standard deviations are 20–25 kJ/mol), while for small fullerenes with a large number of pentagon adjacencies, performance of semiempirical methods was significantly worse. In another work, AM1 and PM3 methods were compared to DFTB and B3LYP/6-31G* for all IPR isomers of C_{72} – C_{86} fullerenes and selected isomers of C_{28} – C_{36} (Zheng et al. 2005). At variance with the results in Chen and Thiel (2003), the authors have found that AM1 and PM3 performed rather bad ($R^2 = 0.77$ and 0.73,

respectively, for C_{72} – C_{86}), while DFTB-predicted values were close to B3LYP/6-31G* ($R^2 = 0.88$ and 0.93 for NCC and SCC-variants of DFTB, respectively). The significant difference in performance of AM1 and PM3 methods reported by two groups is quite remarkable. Close comparison of the relative energies from the two works shows that the difference in relative energies predicted for the *same* isomers by the *same* method can be as high as 50 kJ/mol and in one case ($C_{3v}(8)$ isomer of C_{82}) reaches 75 kJ/mol. At the same time, both works report the same values for the isomer of C_{84} and agree in that semiempirical methods give reliable predictions for this fullerene.

C_{84} is also a special case among the higher fullerenes since, on the one hand, it is the most abundant empty fullerene after C_{60} and C_{70} (at least, if all C_{84} isomers are taken together); on the other hand, it provides the broadest range of the isomers characterized for one fullerene. The lowest energy is predicted by many approaches for $D_2(22)$ and $D_{2d}(23)$ isomers, separated from the other isomers by the gap of ca 30 kJ/mol (Table 3). They are essentially isoenergetic (ΔE varies within few kJ/mol in favor of one or another isomer depending on the method) and have large HOMO–LUMO gaps. Indeed, these are the isomers produced in the largest amounts (the D_2 to D_{2d} ratio is ca 2:1 and varies depending on the conditions of synthesis) (Kikuchi et al. 1992). C_{84} also serves as a good example to emphasize that with the increase of the number of isomers, the density of their distribution in the energy scale is also increasing. The energies of at least 12 isomers fall in the range of $\Delta E = 30 - 71$ kJ/mol at the B3LYP/6-31G* level (Sun and Kertesz 2001b), and many of them can be isolated as “minor” isomers, including (in the order of their yield) $C_2(11)$, $C_s(16)$, $C_s(14)$, $D_{2d}(4)$, and $D_2(5)$ (Dennis et al. 1999). Isolation of $D_{6h}(24)$ and $D_{3d}(19)$ isomers from the arc-discharge soot of Gd-doped composite rods was also reported (note that these isomers were not formed in noticeable amounts with pure graphite rods) (Tagmatarchis et al. 1999). Perfluoroalkylation of the higher fullerene mixture resulted in the isolation and single-crystal X-ray characterization of the derivatives of several C_{84} isomers, including (C_{84} – $C_{2v}(18)$)(C_2F_5)₁₂, the carbon cage isomer of which has not been found in the previous studies (Tamm et al. 2009b).

For C_{86} , two isomers, with C_2 and C_s symmetry, could be isolated from the arc-discharge soot and characterized by ^{13}C NMR spectroscopy (Miyake et al. 2000). From 19 IPR isomer of C_{86} , 6 have C_2 symmetry and 3 have C_s symmetry. DFT calculations show that isomers $C_2(17)$ and $C_s(16)$ have the lowest energies and large HOMO–LUMO gaps and hence can be assigned to the experimentally observed structures (Sun and Kertesz 2002). The assignment was also confirmed by comparison of B3LYP/6-31G computed and experimentally measured ^{13}C NMR chemical shifts (Sun and Kertesz 2002) and by single-crystal X-ray diffraction study of the mixture of two C_{86} isomers (Wang et al. 2010).

In the most recent experimental study of C_{88} , three isomers of this fullerene were isolated and assigned by ^{13}C NMR spectroscopy to one C_s and two C_2 -symmetric structures (Miyake et al. 2000). Among 35 IPR isomers of C_{88} , 11 and 7 structures have C_s and C_2 symmetry, respectively. DFT calculations show that $C_s(17)$ has the lowest energy, followed by $C_2(7)$ and $C_2(33)$ isomers, all within the range of 10 kJ/mol (Table 3). These three structures have large HOMO–LUMO

gaps and are separated from the less-stable isomers by a gap of at least 30 kJ/mol. Therefore, they can be readily assigned to the experimentally observed structures, which are also confirmed by the comparison of the B3LYP/6-31G-computed ^{13}C NMR spectra to the experimental data (Sun 2003b). The presence of $C_2(7)$, $C_s(17)$, and $C_2(33)$ isomers in the arc-discharge produced fullerene mixture was confirmed by the isolation and single-crystal X-ray diffraction characterization of their chloro- and trifluoromethyl derivatives (Trojanov and Tamm 2009; Yang et al. 2012a; Wang et al. 2016).

Detailed experimental study of C_{90} has shown that at least five isomers are formed in the arc-discharge synthesis and that the major fraction is a mixture of C_2 - and C_{2v} -symmetric isomers (Achiba et al. 1995). The B3LYP/STO-3G calculations of all 46 IPR isomers with subsequent B3LYP/6-31G calculations of the isomers with $\Delta E < 100$ kJ/mol have shown that the lowest energy structures are $C_2(45)$ and $C_{2v}(46)$ (Sun 2003b). Analogous result was also obtained in the DFTB and PBE0/6-311G**/DFTB study (Shao et al. 2006). Therefore, $C_2(45)$ and $C_{2v}(46)$ are the most probable candidates for the main major isomers of C_{90} . Chlorination of the mixture of higher fullerenes (Kemnitz and Trojanov 2009) and several C_{90} fractions (Trojanov et al. 2011) resulted in a series of $C_{90}\text{Cl}_{2n}$ structures with six C_{90} isomers: $C_{2v}(46)$, $C_s(35)$, $C_1(30)$, $C_2(28)$, $C_1(32)$, and $C_s(34)$ cage isomers ($\Delta E = 14.2, 15.1, 28.5, 34.3, 38.5,$ and 92.0 kJ/mol, respectively, at the B3LYP/6-31G level). Likewise, trifluoromethylated derivatives were structurally characterized for $C_2(45)$ ($\Delta E = 0.0$ kJ/mol), $C_s(35)$, $C_1(30)$, and $C_1(32)$ isomers (Kareev et al. 2008b; Tamm and Trojanov 2014; Tamm and Trojanov 2015a). $D_{5h}(1)$ ($\Delta E = 76.2$ kJ/mol), $C_1(30)$, and $C_1(32)$ isomers were also isolated from the soot produced by vaporization of Sm_2O_3 /graphite rods and characterized by single-crystal X-ray diffraction (Yang et al. 2010, 2011).

Isolation and ^{13}C NMR characterization of five major isomers of C_{92} , of which two have C_2 symmetry and three have D_2 symmetry, have been reported (Achiba et al. 1998b); however, further spectroscopic details were not given. In another work, one C_2 -symmetric isomer and inseparable mixture of at least three other isomers were obtained from dysprosium arc-burned soot (Tagmatarchis et al. 2002). $C_{92}(\text{CF}_3)_{16}$ with either $D_2(81)$ or $D_2(82)$ carbon cage was identified among the other products of the trifluoromethylation of the higher fullerene mixture (Trojanov and Tamm 2009), whereas chlorination and trifluoromethylation of C_{92} fraction gave derivatives of the $C_1(38)$ isomer (Tamm and Trojanov 2015b). Results of DFT studies are more controversial for C_{92} than for smaller fullerenes discussed above (Table 3). Two detailed studies were reported, one based on the screening of the whole set of IPR isomers at the semiempirical SAM1 level with subsequent B3LYP/6-31G**/SAM1 calculations (Slanina et al. 2000), while another study included screening at the DFTB level followed by PBE0/6-311G**/DFTB calculations of the most stable isomers (Shao et al. 2006). As can be seen in Table 3, the two approaches resulted in substantially different relative energies and stability order of the isomers. While B3LYP//SAM1 approach favors $D_2(84)$ isomer as the most stable one, PBE0//DFTB predicts that $D_3(28)$ is 20 kJ/mol lower in energy than $D_2(84)$. Full optimization of $D_3(28)$ and $D_2(84)$ isomers at the B3LYP/6-311G*

and PBE0/6-311G* levels of theory in Shao et al. (2006) showed that the former is indeed more stable by 18.2 and 20.8 kJ/mol, respectively, in very close agreement with the results of PBE0//DFTB calculations. It appears thus that PBE0//DFTB values are more reliable than the results of B3LYP//AM1 calculations.

Experimental structural information on the empty fullerenes beyond C₉₂ is very scarce and usually limited to the mixtures of isomers. X-ray crystallographic studies of perfluoroalkylated and chlorinated higher fullerene mixtures or partially separated fractions characterized a series of isomer of C₉₄, C₉₆, C₁₀₀, and C₁₀₄ (Tamm et al. 2009a, 2015; Yang et al. 2012b, 2014; Fritz et al. 2014), but these data are inconclusive since the complete isomeric composition of the initial fullerene mixture remains unknown. Yet a lot of calculations were performed for larger fullerenes, and systematic information on the lowest energy isomers is available up to C₁₈₀.

The aforementioned PBE0/6-311G*//DFTB study was performed for all IPR fullerenes in the C₈₂–C₁₂₀ range as well as for all IPR and non-IPR isomers of C₃₈–C₈₀ fullerenes (Shao et al. 2006, 2007). For the most stable isomer, ¹³C NMR spectra were also predicted for comparison with experiment.

In another series of works, the most stable isomers of C₁₂₂ to C₁₈₀ were identified (Xu et al. 2006, 2008). To screen the huge number of the isomers, the authors used the empirical scheme of Cioslowski and coworkers (Cioslowski et al. 2000). According to the original scheme, formation enthalpies of fullerenes were factorized into contributions of 25 independent structural motifs, namely, six-member rings with their first and second neighbors, and the contributions of the motifs were fitted using the energies of 115 IPR fullerenes in the C₆₀–C₁₈₀ size range computed at the B3LYP/6-31G* level. In 2008, parameters were refitted using the energies of 536 fullerene molecules in the C₆₀–C₁₆₀ range (Xu et al. 2008). The serious advantage of this empirical method is that it is very inexpensive, and the energies can be estimated with high accuracy without quantum-chemical calculations. B3LYP/6-31G* calculations were then used by the authors to determine the lowest energy isomers (Xu et al. 2006, 2008). The same authors also performed benchmarking computations of C₁₁₆, C₁₁₈, and C₁₂₀ isomers to analyze performance of QCFF/PI, AM1, PM3, MNDO, and tight-binding (TB) methods in comparison to B3LYP/6-31G*//B3LYP/3-21G (Xu et al. 2007). It was found that, in contrast to the relative energies of the medium-size higher fullerenes such as C₈₄, for large fullerenes semiempirical methods performed badly and did not show any reasonable correlations with DFT-predicted relative energies. However, it should be noted that DFT is not a panacea itself. DFT has difficulties with description of the extended π -systems (Reimers et al. 2003), and hence it is possible that DFT is also not reliable for large fullerenes. Unfortunately, results of computations for larger fullerenes cannot be verified by experimental information at this moment.

In conclusion, numerous studies have shown that the most stable isomers of the middle-size higher fullerenes predicted by DFT are usually the major isomers in the experimentally produced fullerene mixtures. At the same time, it is also clear that experimental results cannot be fully described by considering only thermodynamic control (equilibrium), since some other *unknown* factors are also important in

determining the isomeric composition. For instance, it is not clear yet why some stable isomers are not formed in the arc-discharge synthesis; besides, the influence of the metals on the distribution of the arc-discharge produced fullerene isomers is not understood.

Bonding, Structures, and Stability of Endohedral Metallofullerenes

Metal–Cage Bonding in Endohedral Metallofullerenes

Since La@C_{82} was first isolated in 1991, the nature of the metal–fullerene bonding was of interest to researchers. Because of the radical nature of La@C_{82} as well as other $\text{M}^{\text{III}}\text{@C}_{82}$ ($\text{M} = \text{Y}, \text{Sc}$), the ESR spectroscopy appeared useful in revealing the main features of the electronic structure of EMFs. Small hyperfine splitting constant of the metal observed in the spectra indicated that the spin density in M@C_{82} is presumably localized on the fullerene cage. This, along with the high electron affinity of fullerenes, resulted in the concept of the endohedral metallofullerenes as non-dissociative salts (e.g., $\text{La}^{3+}\text{@C}_{82}^{3-}$), with the metal cation encapsulated in the negatively charged carbon cage. Likewise, ionic model is often used to describe electronic structure of endohedral fullerenes with more complex encapsulated species, such as dimetallofullerenes (e.g., $(\text{La}^{3+})_2\text{@C}_{80}^{6-}$), nitride clusterfullerenes (e.g., $(\text{Sc}_3\text{N})^{6+}\text{@C}_{80}^{6-}$), or carbide clusterfullerenes (e.g., $(\text{Sc}^{3+})_2\text{C}_2^{2-}\text{@C}_{82}^{4-}$).

The concept of the ionic metal–fullerene bonding was further developed by Kobayashi et al. (Kobayashi and Nagase 1998; Nagase et al. 1997). The authors analyzed the spatial distribution of electrostatic potential inside fullerene and showed that it has large negative values inside negatively charged carbon cages. This results in the strong stabilization of the metal cations if they are placed inside fullerene anions thus favoring formation of EMFs. Importantly, the minimum of the electrostatic potential in C_{82}^{3-} was found in the position where metal resides in M@C_{82} as predicted by calculations (Kobayashi and Nagase 1998; Lu et al. 2000) or shown by experimental X-ray diffraction studies (Takata et al. 2003). At the same time, it was also found that electrostatic potential inside C_{80}^{6-} has no distinguishable minima, and hence there are no distinct bonding sites for metal atoms or clusters in $\text{C}_{80}\text{-}I_h(7)$. This is indeed confirmed by several computational studies of $\text{La}_2\text{@C}_{80}\text{-}I_h(7)$ (Kobayashi et al. 1995; Shimotani et al. 2004; Zhang et al. 2007) and $\text{Sc}_3\text{N@C}_{80}\text{-}I_h(7)$ (Campanera et al. 2002; Gan and Yuan 2006; Kobayashi et al. 2001; Popov and Dunsch 2008; Yanov et al. 2006). For $\text{La}_2\text{@C}_{80}$ it is shown that there are two almost isoenergetic minima for positions of La atoms, with D_{2h} and D_{3d} molecular symmetries, and each minimum is multiplied by the group symmetry operations of I_h group. As a result, La_2 unit exhibits almost free rotation inside C_{80} cage, and the charge density of La atoms forms a pentagonal dodecahedron as revealed by synchrotron radiation powder diffraction study (Nishibori et al. 2001). Another important argument in favor of the ionic model is the carbon cage isomerism of EMFs, which shows a good correlation with

the relative stability of the appropriately charged empty cages, as will be discussed in detail below.

Although the ionic model explains some spectroscopic and structural properties of EMFs, there are numerous studies which indicate that this model is oversimplified. While the analysis of MO energy levels in the EMFs and the corresponding empty cages can be indeed interpreted as an electron transfer to the cage from the metal, experimental (Alvarez et al. 2002; Kessler et al. 1997) and theoretical (Campanera et al. 2002; Lu et al. 2000; Muthukumar and Larsson 2008; Nagase et al. 1997; Yang et al. 2008) studies clearly show a substantially non-zero population of *nd*-levels of the endohedral metal atoms. The mixing of the cluster and fullerene MOs as well as the corresponding change of the orbital energies in EMF when compared to the MOs of the empty cage is the most apparent for $\text{Sc}_3\text{N}@C_{78}$ (Fig. 7; Campanera et al. 2002; Krause et al. 2006). In many other EMFs, the metal contribution to individual π -MO orbitals of the cage is very small, yet the significant metal–cage interaction can be also revealed when the electron density distribution is analyzed (Liu et al. 2006; Popov and Dunsch 2008; Wu and Hagelberg 2008; Yang et al. 2008; Popov 2009). This can be best done with the use of promolecule deformation density approach, i.e., by visualizing the changes in the electron density distribution $\Delta\rho = [\rho(\text{Mol}) - \rho(\text{Ref})]$, where $\rho(\text{Mol})$ is the electron density of the molecule under study, while $\rho(\text{Ref})$ is the electron density of the reference system. A natural choice of the reference system in the studies of EMFs can be either noninteracting metal atoms or clusters (e.g., Sc_3N) and corresponding C_{2n} molecules or metal or cluster cations (e.g., Sc_3N^{6+}) and appropriately charged

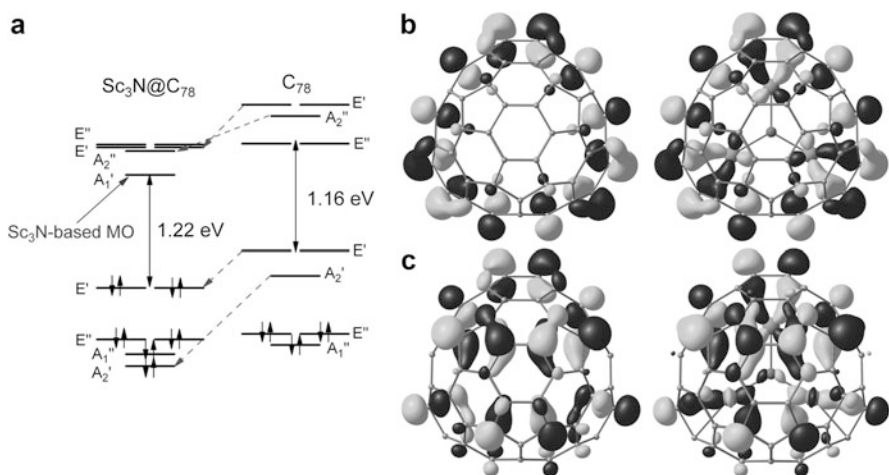


Fig. 7 (a) Frontier Kohn–Sham MO levels in $\text{Sc}_3\text{N}@C_{78}\text{-}D_{3h}(5)$ and $C_{78}\text{-}D_{3h}(5)$, the *dash arrows* show correspondence between the orbitals of the empty cage and $\text{Sc}_3\text{N}@C_{78}$; (b) LUMO (A_2') MO of C_{78} and corresponding occupied orbital in $\text{Sc}_3\text{N}@C_{78}$; (c) LUMO + 1 (E' , only one component is shown) in C_{78} and corresponding occupied orbital in $\text{Sc}_3\text{N}@C_{78}$ (Based on the data from Krause et al. 2006 and Popov 2009)

C_{2n} anions (e.g., C_{2n}^{6-}). In particular, when two sixfold charged ions were taken as a reference, a considerable concentration of the density at Sc atoms was found in all $Sc_3N@C_{2n}$ ($2n = 68, 78, 80$) EMFs (Popov and Dunsch 2008; Popov 2009), the effect referred to by some researchers as “back donation” (Liu et al. 2006; Wu and Hageberg 2008). The $\Delta\rho$ plots have also revealed that metal–cage bonds are not localized (such as, e.g., carbon–carbon bonds) but can be better described by “coordination” bonds such as those found in transition metal complexes (Cortés-Guzmán and Bader 2005).

The cluster–cage bonding in $Sc_3N@C_{78}$ and $Sc_3N@C_{80}$ was also analyzed using the energy decomposition method (Campanera et al. 2002). The authors have shown that after the Sc_3N^{6+} cation is encapsulated inside $C_{78,80}^{6-}$ cages, a strong orbital mixing and an electronic reorganization take place.

Computed atomic charges are also much smaller than the values expected for the purely ionic bonding. For instance, net Bader, Mulliken, and NBO charges of the Sc_3N cluster in $Sc_3N@C_{80}$ are +3.50, +2.88, and +2.99, respectively, at the B3LYP/6-311G* level of theory (Popov and Dunsch 2008); another group reported the net Sc_3N charges of +1.28 (Mulliken), +1.10 (Hirshfeld), and +0.94 (Voronoi) (Valencia et al. 2008) at the BP/TZVP level of theory with ZORA relativistic corrections. Note that the charges strongly depend both on the method of theory used to compute wavefunction and on the electron density partitioning, but all reported values are significantly smaller than +6 expected for the ionic $Sc_3N^{6+}@C_{60}^{6-}$.

In summary, from the molecular orbital point of view, the mechanism of the metal(cluster)–cage interactions in EMFs can be described by the formal transfer of the appropriate number of electrons from the endohedral species to the carbon cage with subsequent coordination of the metal cations by the cage as a “ligand” and reoccupation of the metal nd orbitals (see also (Liu et al. 2006) for a more detailed discussion). Note that in most cases there are no special localized orbitals that could be responsible for the cage-to-cluster electron transfer. Instead, this kind of interaction occurs through the overlap of many cage π -orbitals with nd orbitals of the metal. One of the consequences of this type of bonding is a spin-charge separation found in the radical anions of the EMFs with the metal-localized LUMO (such as $Sc_3N@C_{80}$, $La_2@C_{72,78,80}$, $Ti_2C_2@C_{78}$). In these anions, the spin density is mostly localized on the metal atoms; however, the changes of the metal charges compared to the neutral state of the EMF molecule are rather small; instead, the surplus charge is mostly delocalized over the carbon cage (Popov and Dunsch 2008).

Since the effect of the metal–cage bonding in EMFs is best revealed in the analysis of the electron density as a whole, rather than by studying the individual MOs, it is natural to use the method based on the analysis of the electron density for the quantitative description of the metal–cage interactions in EMFs. The quantum theory of atoms in molecules (QTAIM) developed by Bader is probably the most refined and well-established method for the analysis of the topology of the electron density frequently used for the revealing and quantifying of the bonding situation between the atoms (Bader 1990; Matta and Boyd 2007). Kobayashi and Nagase analyzed a topology of the electron density distribution in $Sc_2@C_{84}$, $Ca@C_{72}$, and

Sc₂@C₆₆ using QTAIM (Kobayashi and Nagase 1999, 2002a). Their studies have shown that although the bond critical points (BCPs: critical points of the electron density, in which the density has a minimum along one axis and has a maximum in the plane perpendicular to this direction) could be found between Sc or Ca and certain carbon atoms of the cage, the values of the electron density at BCPs (ρ_{bcp}) were very small (0.05 a.u. for Sc-C and 0.01–0.02 a.u. for Ca-C), and the Laplacian of the density ($\nabla^2\rho_{bcp}$) at metal–carbon BCPs was always positive. The authors concluded that such values for these descriptors are signatures of the highly ionic character of the metal–cage bonding. However, care should be taken when the bonding in transition metal compound is analyzed with QTAIM (Cortés-Guzmán and Bader 2005). It was revealed that for transition metals $\nabla^2\rho_{bcp}$ values are usually positive, while ρ_{bcp} is small because of the diffuse character of electron distribution (Macchi and Sironi 2003), and hence these values alone cannot be proper descriptors of the bonding situation. Therefore, more specific descriptors should be applied in the analysis of the bonds involving transition metals, and the analysis of the *energy density* appears more useful than the analysis of the electron density alone. While kinetic energy density, G , is positive everywhere, potential energy density, V , is negative everywhere, and their sum, the total energy density H , defines which of the energy components is dominating. Cremer and Kraka suggested that bonding between the atoms can be considered as covalent if the total energy density at BCP, H_{bcp} , is negative (Cremer and Kraka 1984).

The in-depth study of the metal–cage and intra-cluster bonding by DFT and QTAIM was recently reported for the four major classes of endohedral metallofullerenes, including monometallofullerenes, dimetallofullerenes, metal nitride clusterfullerenes, metal carbide clusterfullerenes, as well as Sc₃CH@C₈₀ and Sc₄O_x@C₈₀ ($x = 2, 3$) (Popov and Dunsch 2009). The analysis of the bond critical point indicators showed that both the intra-cluster and the metal–cage interactions in EMFs were characterized by the negative total energy density, which means that bonding in EMFs exhibits a high degree of covalency. Furthermore, from the point of the electron density distribution, the interior of the EMF molecules was found to be a complex topological object, with many critical points, unpredictable number of the bond paths, and large bond ellipticities, similar to the bonding situation in the complexes of transition metals with π -carbocyclic ligands (Farrugia et al. 2009). In this respect, the analysis of the metal–cage bonding based only on the properties of the bond paths was found to be insufficient, and the metal–cage delocalization indices, $\delta(M, C)$, were also analyzed (by definition, $\delta(M, C)$ is number of the electron pairs shared by the atoms M and C, which is very similar to the “bond order” in Lewis definition). It was found that in the majority of EMFs, $\delta(M, C)$ values do not exceed 0.25; however, when summed over all metal–cage interactions, values of $\delta(M, \text{cage})$ close to 2–3 were obtained. While QTAIM atomic charges of the metal atoms were found to be approximately two times smaller than their formal oxidation states, the total number of the electron pairs shared by the metal atoms with the EMF molecule, $\Delta(M)$, was found to be close to the typical valence of the given element (Table 4).

Table 4 QTAIM atomic charges (q) of the metal atoms and the clusters and the metal–cage and metal–cluster electron-pair sharing in selected EMFs ^a

EMF	$q(\text{M})$ QTAIM	$q(\text{cluster})$ QTAIM	$q(\text{cluster})$ formal	$\delta(\text{M},$ cluster)	$\delta(\text{M},$ cage)	$\Delta(\text{M})$
$C_2\text{-Ca@C}_{72}$	1.50		2		1.34	1.34
$C_s\text{-Ca@C}_{82}$	1.52		2		1.21	1.22
$C_s\text{-Sc@C}_{82}$	1.60		3		2.71	2.71
$D_3\text{-Sc}_3\text{N@C}_{68}$	1.74	3.45	6	0.81	1.89	2.70
$C_3\text{-Sc}_3\text{N@C}_{80}$	1.76	3.50	6	0.82	1.77	2.59
$D_2\text{-Sc}_2\text{C}_2\text{@C}_{84}$	1.72	2.01	4	0.63	2.00	2.63
$C_s\text{-Sc}_3\text{C}_2\text{@C}_{80}$	1.67, 1.67	3.30	6	0.86, 0.63	1.87, 2.13	2.73, 2.76
$C_3\text{-Sc}_3\text{CH@C}_{80}$	1.72	3.48	6	0.82	1.82	2.64
$C_1\text{-Sc}_4\text{O}_2\text{@C}_{80}$	1.41, 1.75	3.66	6	1.21, 1.32	1.74, 1.41	2.97, 2.65
$C_1\text{-Sc}_4\text{O}_3\text{@C}_{80}$	1.87, 1.82	3.47	6	1.43, 1.11	1.07, 1.48	2.52, 2.60
$C_2\text{-Ti}_2\text{C}_2\text{@C}_{78}$	1.67	2.17	6	0.82	3.29	4.11
$C_s\text{-Y@C}_{82}$	1.80		3		2.54	2.54
$C_s\text{-Y}_2\text{@C}_{82}$	1.35	2.69	4	0.63	2.25	2.88
$C_s\text{-Y}_2\text{@C}_{79}\text{N}$	1.52	3.04	5	0.37	2.37	2.74
$C_2\text{-Y}_3\text{N@C}_{78}$	1.90	3.90	6	0.83	1.76	2.61
$D_3\text{-Y}_3\text{N@C}_{80}$	1.90	3.87	6	0.85	1.77	2.62
$D_2\text{-Y}_3\text{N@C}_{88}$	1.93	3.95	6	0.79	1.65	2.44
$C_1\text{-Y}_2\text{C}_2\text{@C}_{82}$	1.89	2.40	4	0.66	1.90	2.56
$C_2\text{-La@C}_{72}$	1.91		3		2.72	2.72
$C_{2v}\text{-La@C}_{82}$	1.82		3		2.69	2.69
$D_{2h}\text{-La}_2\text{@C}_{80}$	1.65	3.29	6	0.08	2.95	3.04
$C_2\text{-La}_3\text{N@C}_{88}$	1.88	3.92	6	0.95	1.98	2.93
$C_2\text{-La}_3\text{N@C}_{96}$	1.89	3.99	6	0.90	1.89	2.80

^a $\delta(\text{M}, \text{cluster})$ is the number of the electron pairs shared by the metal with the cluster; $\delta(\text{M}, \text{cage})$ is the number of the electron pairs shared by the metal with the carbon cage; $\Delta(\text{M})$ is the total number of the electron pairs shared by M with all other atoms of the molecule; when EMF has inequivalent metal atoms, corresponding values are averaged, unless they are significantly different; in the latter cases, different values are listed in the table separated by comma. Based on the QTAIM-B3LYP/6-311G* calculations from (Popov and Dunsch 2009)

To conclude the section on the metal–cage bonding in EMFs, it appears useful to distinguish two ways to determine the charge on the atom in endohedral fullerene: (1) a “formal” charge and (2) the “actual” charge. The formal charge is integer and implies that the metal (cluster)–cage bonding is purely ionic. As such, formal charges do not describe the actual electron distribution in the endohedral fullerenes, but still appear to be useful in understanding the spectroscopic and structural properties of endohedral fullerenes. For instance, compounds with the same formal cage and metal charges exhibit very similar absorption and vibrational spectra. In

the next section, we will show that the formal charge of the cage largely determines the isomeric structure of a given endohedral fullerene.

The “actual” charge should represent the electron density distribution in the endohedral fullerenes, but unfortunately there are no unique ways to partition the electron density between the atoms. Depending on the definition and the method of theory used to evaluate the charges, they can vary in a large range in the same molecule as already noted for $\text{Sc}_3\text{N}@C_{80}$. However, whatever method is used, computed charges are always considerably smaller than the formal charges, which is a clear manifestation of the covalent contribution to the metal–cage interactions. At the same time, the charges are still considerably large to indicate that ionic contribution to the cluster–cage bonding is also important.

Isomerism in Endohedral Metallofullerenes: Stability of the Charged Carbon Cages

One of the specific features of the endohedral metallofullerenes is that their carbon cage isomers are usually different from those of the isolated empty fullerenes. This fact can be understood by taking into account the electron transfer from the encapsulated species as discussed in the previous section. Using the semiempirical QCFF/PI method, Fowler and Zerbetto have shown that charging dramatically changes relative stabilities of the fullerene isomers (Fowler and Zerbetto 1995). For instance, $I_h(7)$ isomer of C_{80} , being the least stable structure for (+2) and (0) charge states, is the lowest energy structure for (−4) and (−6). In agreement with this finding, it was shown that the most stable isomer of $\text{La}_2@C_{80}$ has $I_h(7)$ cage, as opposed to the D_2 -symmetric empty C_{80} (Kobayashi et al. 1995). Variation of the relative energies of C_{82} isomers in the negatively charged states was analyzed by Kobayashi and Nagase at the HF/3-21G//AM1 level to rationalize experimentally observed isomers of $\text{M}@C_{82}$ (Kobayashi and Nagase 1997, 1998). The authors have found that while $C_2(3)$ is the lowest energy isomer of the neutral C_{82} , the $C_{2v}(9)$ cage is the most stable for both C_{82}^{2-} and C_{82}^{3-} as well as for $\text{M}@C_{82}$ ($\text{M} = \text{Ca}, \text{Sc}, \text{Y}, \text{La}, \text{etc.}$) (Kobayashi and Nagase 1997, 1998). This finding is confirmed by experimental ^{13}C NMR studies of $\text{La}@C_{82}$ and $\text{Ce}@C_{82}$ anions (Akasaka et al. 2000; Tsuchiya et al. 2005; Wakahara et al. 2004). More recently, DFT calculations of C_{82}^{4-} isomers were reported (Valencia et al. 2008) and the $C_{3v}(8)$ isomer was shown to be the most stable tetraanion, in line with the finding of this cage in the most abundant isomers of $\text{Sc}_2\text{C}_2@C_{82}$ (Iiduka et al. 2006; Nishibori et al. 2006a) and $\text{Y}_2\text{C}_2@C_{82}$ (Inoue et al. 2004; Nishibori et al. 2006b). The data on the relative energies of all IPR C_{82} isomers in (0)–(−6) charge states obtained at the PBE/TZ2P level (Popov 2009) are summarized in Fig. 8a. DFT calculations show that $C_{2v}(9)$ is the most stable isomer for all charge states starting from (−2), however, there is a change in favor of $C_{3v}(8)$ for (−4) (Valencia et al. 2008). Note also that starting from (−3) charge state, the IPR isomers $C_{2v}(9)$, $C_{3v}(8)$, and $C_s(6)$ are considerably more stable than the other IPR isomers; as a result, only these isomers are found for $\text{M}^{\text{III}}@C_{82}$ and $\text{M}_2\text{C}_2@C_{82}$.

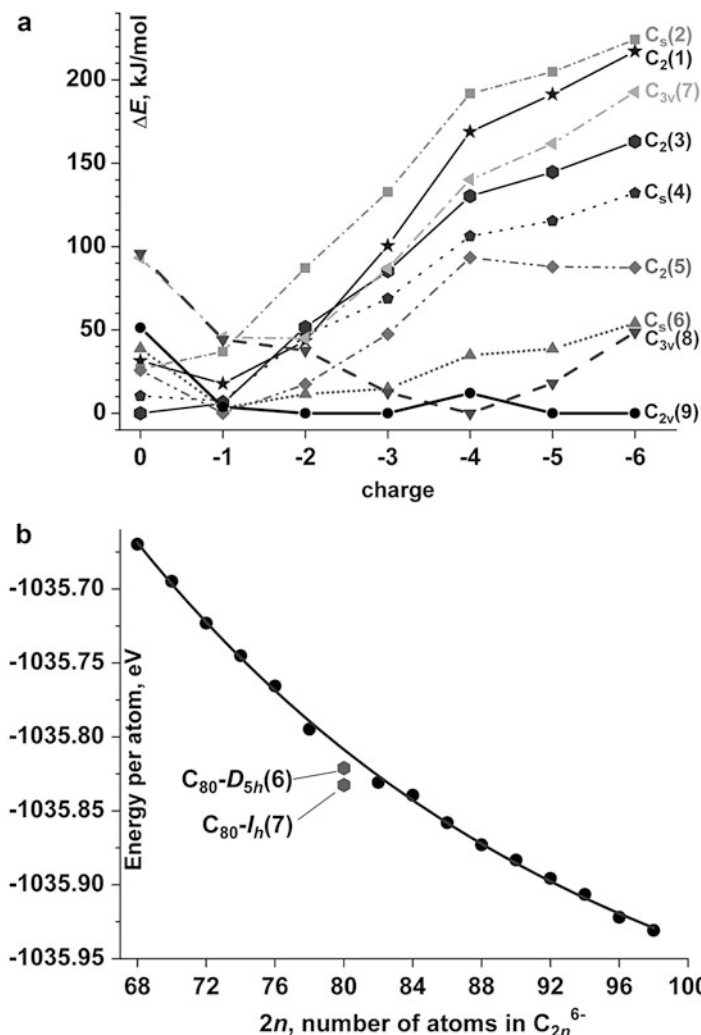


Fig. 8 (a) Relative energies IPR isomers of C_{82} computed at the PBE/TZ2P level in different charge states (Based on the data from Popov 2009); (b) PBE/TZ2P computed energies of the most stable C_{2n}^{6-} hexaanions on the per-atomic basis plotted as a function of the cage size (Based on the data from Popov and Dunsch 2007)

The data on representative fullerenes discussed above show that (1) the relative energies of the fullerene isomers are dramatically affected by the charge and (2) the stability of the isomers of appropriately charged empty fullerenes provides a good estimation to the relative stability of the cage isomers of endohedral fullerenes. Thus, though ionic model is not realistic for the description of the electron density distribution in endohedral metallofullerenes, it can be useful for understanding the carbon cage isomerism of endohedral fullerenes.

In the last decade, the higher yields, increased stability, and broad variety of the formed structures attracted special attention to the family of nitride clusterfullerenes. Their electronic structure may be conceived as a result of a formal sixfold electron transfer from the $M_3^{III}N$ cluster to the fullerene. Predicting of the lowest energy isomers of $M_3N@C_{2n}$ endohedral fullerenes is a difficult task since, in addition to the carbon cage isomerism, different orientation of the M_3N cluster inside the cage can give structures with substantially different relative energies; besides, DFT calculations of the molecules with transition metals are more complex than those of the empty fullerenes. In this respect, attempts were made to predict the possible structures of $M_3N@C_{2n}$ molecules based on the calculation for the empty fullerenes.

The simple use of this formal sixfold electron transfer in $M_3N@C_{2n}$ to predict the most suitable cage isomers capable of encapsulating a nitride cluster was proposed by Poblet and coworkers. It was supposed that only the fullerenes with a considerable gap between LUMO + 2 and LUMO + 3 (which become HOMO and LUMO, respectively, in the C_{2n}^{6-} hexaanion and presumably in $M_3N@C_{2n}$) should be considered as suitable hosts for nitride clusters (Campanera et al. 2005). With the application of this criterion, only C_{60} , $C_{78}-D_{3h}(5)$, $C_{80}-I_h(7)$, and $C_{80}-D_{5h}(6)$ were found to be suitable cage isomers among all IPR fullerenes in the C_{60} – C_{84} range. Indeed, besides C_{60} , which is too small to host a Sc_3N cluster, only these and no other IPR cage isomers were found for $Sc_3N@C_{2n}$ clusterfullerenes. Later this work was expanded by considering also all IPR isomers in the C_{86} – C_{100} range (Valencia et al. 2007), and the isomers with the largest orbital gaps found for C_{86} and C_{88} hexaanions ($D_3(19)$ and $D_2(35)$, respectively) were those proved to exist in $Tb_3N@C_{86}$ and $Tb_3N@C_{88}$ by single-crystal X-ray diffraction studies (Zuo et al. 2007).

A different approach to the search of the suitable cage isomers for nitride clusterfullerenes is based on the stability of the hexaanions (Popov and Dunsch 2007; Popov et al. 2007c; Yang et al. 2007a, b). In the view of the correlation found between the relative energies of the empty fullerenes in appropriate charge state and the relative energies of endohedral fullerenes, it was suggested that the suitable cage isomers of $M_3N@C_{2n}$ should be found among the most stable isomers of C_{2n}^{6-} . The growing number of endohedral fullerenes violating IPR shows that non-IPR isomers should also be included in the search. Importantly, good correlation was found between the relative energies of the C_{2n}^{6-} isomers computed at the AM1 and DFT levels of theory (Yang et al. 2007b). Since optimization of the structure of the fullerene at AM1 levels normally takes only about 1–2 min with the standard office computer, this fact dramatically facilitates the search of the stable isomers. That is, screening of *all* possible isomers can be done at AM1 level, and only few lowest energy C_{2n}^{6-} isomers (and corresponding $M_3N@C_{2n}$ isomers) can then be studied at the higher – and more computationally demanding – level of theory. By screening through the large number of the C_{70}^{6-} , C_{76}^{6-} , and C_{78}^{6-} isomers, the molecular structures of $Sc_3N@C_{70}$ (Yang et al. 2007b), $M_3N@C_{78}$ ($M = Dy, Tm$) (Popov et al. 2007c), and $DySc_2N@C_{76}$ (Yang et al. 2007a) were proposed, all violating the IPR. The structural guesses were then confirmed by comparison of the experimental and DFT-computed vibrational spectra.

The full search of the most favorable cage isomers of $\text{Sc}_3\text{N}@C_{2n}$ and $\text{Y}_3\text{N}@C_{2n}$ ($2n = 68-98$) at the PBE/TZ2P level was reported in 2007 (Popov and Dunsch 2007). For $C_{68} - C_{88}$, both IPR and non-IPR isomers were considered, which however resulted in the huge number of possible structures (for instance, there are totally 81,738 for C_{88}). Taking into account that metal atoms in all known non-IPR fullerenes are coordinated to the adjacent pentagon pairs, the authors suggested that only the isomers with such pairs should be considered (while the structures with three or more adjacent pentagons could be omitted). This assumption considerably reduced the number of the isomers (e.g., 16,717 isomers were studied for C_{88}), making the computations more feasible.

Table 5 lists the lowest energy isomers of C_{2n}^{6-} found in that work with their relative energies and HOMO–LUMO gaps computed at the PBE/TZ2P level; experimentally available structurally characterized EMFs are listed in the last column. Importantly, the isomers shown to exist by single-crystal X-ray studies can be found among the lowest energy isomers of C_{2n}^{6-} . $C_{78}\text{-}C_2(22,010)$ and $C_{82}\text{-}C_5(39,663)$ isomers were predicted as possible structures of $\text{M}_3\text{N}@C_{78}$ and $\text{M}_3\text{N}@C_{82}$ (Popov and Dunsch 2007; Popov et al. 2007c) before their structures were confirmed by single-crystal X-ray studies of $\text{Gd}_3\text{N}@C_{78}$ and $\text{Gd}_3\text{N}@C_{82}$ (Beavers et al. 2009; Mercado et al. 2008). Even better correlation with experimental data is found if the relative energies of $\text{Sc}_3\text{N}@C_{2n}$ and $\text{Y}_3\text{N}@C_{2n}$ isomers are considered (Popov and Dunsch 2007). It is remarkable that C_{2n}^{6-} isomers with high thermodynamic stability (i.e., low relative energy) quite often exhibit high kinetic stability (i.e., large HOMO–LUMO gap). As a result, the IPR isomers proposed by the group of Poblet (Campanera et al. 2005; Valencia et al. 2007) in most cases coincide with predictions based on the relative stability of the hexaanions (Popov and Dunsch 2007).

The broad range of the fullerene sizes studied (Popov and Dunsch 2007) enabled the authors to follow the general trends in their stabilities. To compare the energies of the fullerenes of different size, the absolute energies were normalized to the number of atoms in the given fullerene. Figure 8b plots the normalized energies of the most stable C_{2n}^{6-} isomers versus the number of atoms. The smooth decrease of the energy with the increase of the fullerene size is the result of a combination of two factors: (1) the decrease of the curvature of the cage, which results in smaller curvature-induced strain, and (2) the decrease of the on-site Coulomb repulsions of six surplus electrons in C_{2n}^{6-} with the increase of the cage size. There is, however, a significant deviation from the smooth curve found for $C_{80}\text{-}I_h(7)$ and $C_{80}\text{-}D_{5h}(6)$ isomers. These isomers are 185 and 98 kJ/mol more stable than they might be if they were like all other fullerenes (i.e., like those which obey the smooth decay in the normalized energy). The enhanced stability of the two C_{80}^{6-} isomers explains the increased yield of $\text{M}_3\text{N}@C_{80}$ compared to all other cage sizes. Besides, it explains why the $C_{80}\text{-}I_h(7)$ isomer is always the most preferable structure in other clusterfullerenes with formal sixfold electron transfer to the cage, such as $\text{La}_2@C_{80}$, $\text{Sc}_4\text{O}_{2,3}@C_{80}$ (Stevenson et al. 2008), and $\text{Sc}_4\text{C}_2@C_{80}$ (Wang et al. 2009).

Table 5 The lowest energy C_{2n}^{6-} ($2n = 68-88$) isomers with their relative energies (kJ/mol) and HOMO–LUMO gaps (eV) computed at the PBE/TZ2P level of theory and the experimentally available EMFs with corresponding carbon cages (Based on the data from Popov and Dunsch 2007)

C_{2n}	Isomer		ΔE	Gap	Exp. EMFs
C_{68}	$D_3(6140)$	Non-IPR	0.0	1.20	$Sc_3N@C_{68}$
C_{70}	$C_{2v}(7854)$	Non-IPR	0.0	1.24	$Sc_3N@C_{70}$
C_{72}	$D_2(10611)$	Non-IPR	0.0	1.12	$M_2@C_{72}$ ($M = La, Ce$)
C_{74}	$C_2(13295)$	Non-IPR	0.0	1.22	Not known
C_{76}	$C_s(17490)$	Non-IPR	0.0	1.12	$DySc_2N@C_{76}, La_2@C_{76}$
C_{78}	$D_{3h}(5)$	IPR	0.0	1.21	$M_2@C_{78}$ ($M = La, Ce$), $Ti_2C_2@C_{78}, Sc_3N@C_{78}$
	$C_2(22010)$	Non-IPR	59.1	1.60	$M_3N@C_{78}$ ($M = Y, Dy, Tm, Gd$), $Sc_3NC@C_{78}$
C_{80}	$I_h(7)$	IPR	0.0	1.83	$M_2@C_{80}$ ($M = La, Ce$), $M_3N@C_{80}$ ($M = Sc, Y, Gd-Lu$), $Sc_{3,4}C_2@C_{80}$, $Sc_4O_{2,3}@C_{80}$
	$D_{5h}(6)$	IPR	88.2	1.51	$Ce_2@C_{80}, M_3N@C_{80}$ ($M = Sc, Y, RE$)
	$C_s(\text{hept})^a$	Heptagon	171.0	1.27	$LaSc_2N@C_{80}$
C_{82}	$C_{2v}(9)^b$	IPR	0.0	0.79	$Sc_3N@C_{82}$
	$C_{2v}(39705)^b$	Non-IPR	30.1	1.32	Not known
	$C_s(39663)^b$	Non-IPR	61.2	1.42	$M_3N@C_{82}$ ($M = Y, Gd, Tm, Dy, Tb$)
C_{84}	$D_2(21)$	IPR	0.0	0.80	Not known
	$C_s(51365)$	Non-IPR	1.8	1.34	$M_3N@C_{84}$ ($M = Y, Gd, Tm, Dy, Tb$)
C_{86}	$D_3(19)$	IPR	0.0	1.51	$M_3N@C_{86}$ ($M = Y, Gd, Tm, Dy, Tb$)
C_{88}	$D_2(35)$	IPR	0.0	0.97	$M_3N@C_{88}$ ($M = Y, Ce, Nd, Gd, Tm, Dy, Tb$)

^aFor $LaSc_2N@C_{80}$, the relative energies for isomers $I_h(7)$, $D_{5h}(6)$, and heptagon-containing $C_s(\text{hept})$ are 0, 61, and 27 kJ/mol, respectively, at the PBE/TZ2P level (Zhang et al. 2015)

^bFor $Y_3N@C_{82}$, the relative energies for isomers $C_{2v}(9)$, $C_{2v}(39,705)$, and $C_s(39,663)$ are 29.6, 0.0, and 32.6 kJ/mol, respectively, at the PBE/TZ2P level

The exclusive stability of the $I_h(7)$ and $D_{5h}(6)$ isomers of C_{80}^{6-} can be rationalized using the concept of hexagon indices (Raghavachari 1992) discussed in section “Structures and Stability of Empty Fullerenes.” The lowest strain condition (all indices are equal) is fulfilled only for $C_{60}-I_h(1)$, $C_{80}-I_h(7)$, and $C_{80}-D_{5h}(6)$ in the whole range fullerenes C_{2n} with $2n < 120$ (Fowler and Manolopoulos 1995). Thus, the exceptional stability of two C_{80}^{6-} isomers can be explained by the favorable distribution of the pentagons, which leads to the least steric strain. Moreover, the list of the least-strained IPR isomers (Table 2) perfectly corresponds to the lowest energy IPR isomers of $C_{76}^{6-}-C_{90}^{6-}$ (Popov and Dunsch 2007). For larger cages

these conditions are less instructive, because many of the IPR isomers satisfy them. Interestingly, the relative energies of the IPR C_{2n}^{6-} isomers follow the rationalization of the stability based on the steric strain much better than the relative energies of the uncharged IPR fullerenes, which usually violate the requirement of the minimized strain (the only exclusions are C_{84} and C_{90} ; see Table 2). Very recently, this phenomenon was clarified by Poblet and coworkers (Rodríguez-Fortea et al. 2010). The authors have suggested that the relative stabilities of the isomers of multiply-charged fullerene anions are largely determined by the on-site Coulomb repulsion. Since the negative charge in anions is mostly localized on the pentagons because of their non-planarity, the lowest energy structures should be those in which the maximal separation (and hence the most uniform distribution) of the pentagons is achieved. The authors have constructed the inverse pentagon separation index to show that its correlation with the relative stabilities of the isomers is improving with the increase of the charge. Thus, minimization of the on-site Coulomb repulsion and minimization of the steric strain are achieved by fulfillment of the same condition, namely, the most uniform distribution of the pentagons. If for the non-charged fullerenes, the steric strain factor can be outweighed by other factors such as stability of the π -system (and hence the lowest strain condition is usually violated), with the increase of the negative charge on the cage, the Coulomb repulsion sooner or later outweighs all other factors. Therefore, to predict the lowest energy isomers of the EMFs with high formal charge of the cage (such as $M_3N@C_{2n}$), one can apply the same arguments as were developed in early 1990s to find the isomers of empty fullerenes with the lowest steric strain. Another group has also shown that the most suitable host cages have the highest aromaticity, and so aromaticity criterion can be used to explain preference of certain isomers (García-Borràs et al. 2013).

Isomerism in Endohedral Metallofullerenes: Cage Form Factor

The carbon cage stability is very important in determining the molecular structure of endohedral fullerenes; however, the isolated structures are not always based on the lowest energy isomers of the hollow anions. In addition to the high stability, the fullerene cage should also provide a suitable shape for enclosed metal atoms or clusters; that is, there should be enough inner space in the cage for the endohedral species. If then a fullerene has adjacent pentagon pairs, they should be located in such a way that their coordination by the endohedral metal atoms is possible. This factor can be especially important for the nitride clusterfullerenes since the shape of the cluster itself imposes significant limitations on the possible location of the binding sites.

Figure 9a plots the relative energies of the 20 lowest energy isomers of C_{72}^{6-} and corresponding $Sc_3N@C_{72}$ isomers calculated at the PBE/TZ2P level. The significant difference between the two sets of data is observed. While the most stable isomer of C_{72}^{6-} is $D_2(10,611)$ (Slanina et al. 2006; Popov and Dunsch 2007), this cage isomer is destabilized for $Sc_3N@C_{72}$, and the lowest energy is found for $Sc_3N@C_{72}$

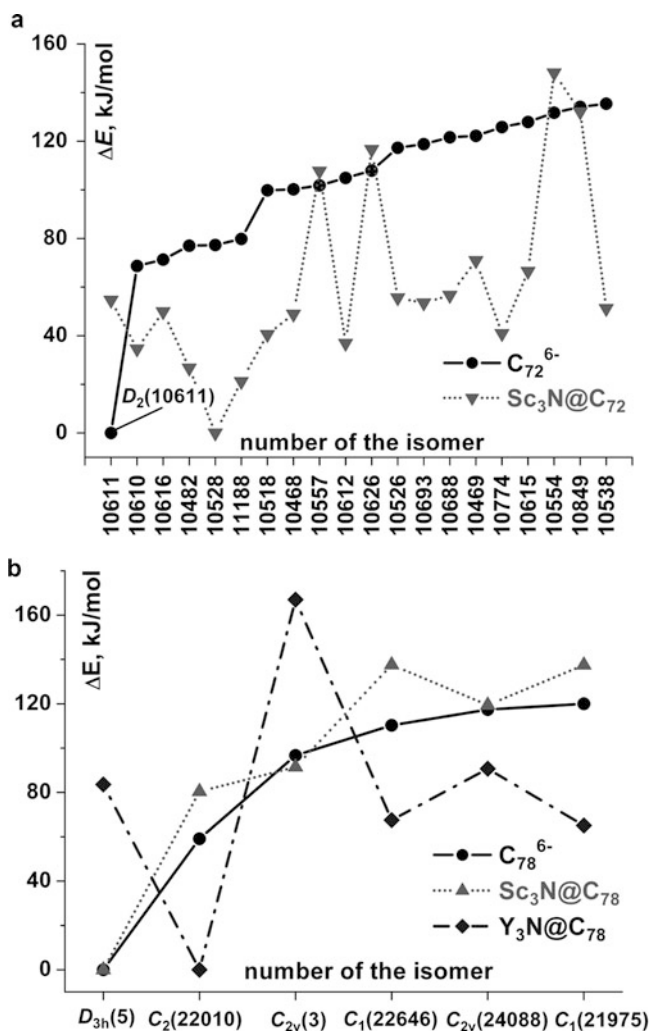


Fig. 9 (a) Relative energies of C_{72}^{6-} and corresponding $Sc_3N@C_{72}$ isomer computed at the PBE/TZ2P level (Based on the data from Popov and Dunsch 2007); (b) relative energies of C_{78}^{6-} and corresponding $Sc_3N@C_{78}$ and $Y_3N@C_{78}$ isomer computed at the PBE/TZ2P level (Based on the data from Popov et al. 2007c)

with $C_s(10,528)$ cage. The explanation of this phenomenon becomes clear when the structures of the cage isomers are analyzed. $C_{72}-D_2(10,611)$ has elongated shape with two adjacent pentagon pairs on the opposite poles of the cage (Fig. 10). It is impossible to coordinate both pentagon pairs at once with the Sc atoms of the triangular Sc_3N cluster, and hence one pentagon pair in $Sc_3N@C_{72}-D_2(10,611)$ is not stabilized by the metal, which explains the high relative energy found for

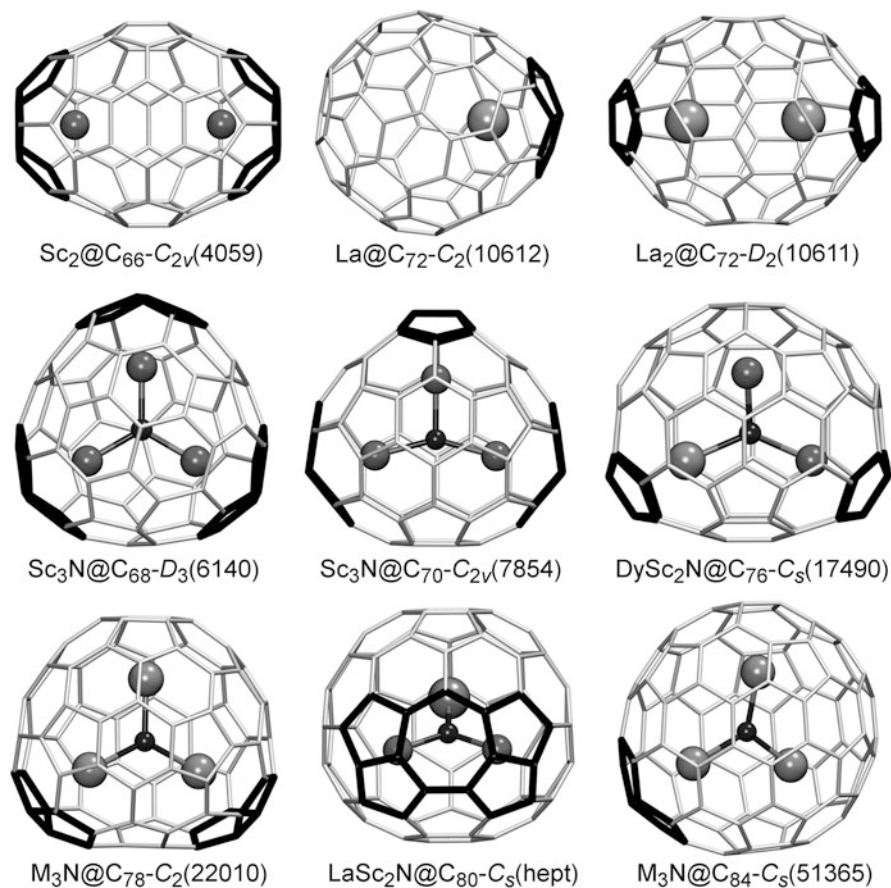


Fig. 10 Molecular structures of selected non-IPR endohedral metallofullerenes. Carbon cages are shown in *gray* except for the adjacent pentagon pairs, which are shown in *black*. Note that in $\text{Sc}_2@C_{66}-C_{2v}(4059)$ has two fragments with three fused pentagons in each. In $\text{M}_3\text{N}@C_{78}$, and $\text{M}_3\text{N}@C_{84}$, M can be Y, Gd, Dy, Tm, Tb, and possibly some other lanthanides

this isomer by theory. The lowest energy C_{72}^{6-} isomer with favorable location of adjacent pentagon pairs is $C_s(10,528)$, and it corresponds to the most stable isomer of $\text{Sc}_3\text{N}@C_{72}$. However, this isomer is based on the relatively unstable carbon cage and, besides, it has a small HOMO–LUMO gap (0.39 eV at the PBE/TZ2P level). As a result, the $\text{Sc}_3\text{N}@C_{72}$ nitride clusterfullerene has never been observed experimentally. Similar situation was also found for C_{74}^{6-} and $\text{Sc}_3\text{N}@C_{74}$, which also has never been observed. Since the bimetallic cluster is more flexible, the steric factor is less important for dimetallofullerenes, and $\text{La}_2@C_{72}$ is based on the lowest energy $D_2(10,611)$ cage isomer (Kato et al. 2003; Lu et al. 2008).

A different situation was found for the clusterfullerenes based on C_{78} (Popov et al. 2007c). Figure 9b compares relative energies of the lowest energy isomers of

C_{78}^{6-} and $M_3N@C_{78}$ ($M = Sc, Y$). The most stable C_{78}^{6-} isomer is IPR $D_{3h}(5)$, which is followed by the non-IPR $C_2(22,010)$ isomer with two adjacent pentagon pairs. The relative energies of $Sc_3N@C_{78}$ isomers closely follow those of C_{78}^{6-} , which agrees with the formation of the $Sc_3N@C_{78}-D_{3h}(5)$ isomer (Olmstead et al. 2001). $C_{78}-D_{3h}(5)$ was also proved for $M_2@C_{78}$ ($M = La, Ce$) and $Ti_2C_2@C_{78}$ (Cao et al. 2004, 2008; Tan and Lu 2005; Sato et al. 2006; Yumura et al. 2005; Yamada et al. 2008). However, with the increase of the cluster size from Sc_3N to Y_3N , the relative energies of the $M_3N@C_{78}$ isomers are changed dramatically. The isomer based on the $D_{3h}(5)$ cage is significantly destabilized, and the lowest energy is now found for the $Y_3N@C_{78}-C_2(22,010)$ (84 kJ/mol more stable than $Y_3N@C_{78}-D_{3h}(5)$). The dramatic change in the stability of the isomers can be explained by insufficient size of the $C_{78}-D_{3h}(5)$ cage for encapsulation of the clusters larger than Sc_3N . As a result, Y_3N cluster is forced to be pyramidal in $Y_3N@C_{78}-D_{3h}(5)$ (nitrogen atoms is displaced from the plane of the metal atoms by 0.554 Å), which is an indication of the strong strain exerted by the cage. On the contrary, the inner space in $C_{78}-C_2(22,010)$ isomer is sufficient to retain the planar shape of the Y_3N cluster (Fig. 10). Results of these calculations were supported by experimental spectroscopic studies of $Tm_3N@C_{78}$ and the major isomer of $Dy_3N@C_{78}$ (Popov et al. 2007c) and by single-crystal X-ray diffraction studies of $Gd_3N@C_{78}$ (Beavers et al. 2009).

$M_3N@C_{80}-I_h(7)$ represents another important example of the cluster size influence on the stability of the clusterfullerenes (Popov and Dunsch 2007). The lowest energy isomers for C_{80}^{6-} and $Sc_3N@C_{80}$ are $I_h(7)$ and $D_{5h}(6)$, and there is a gap of approximately 80 kJ/mol between these two isomers and the gap of approximately 120 kJ/mol between $D_{5h}(6)$ and all other, less-stable isomers. Stability order of C_{80}^{6-} and $Sc_3N@C_{80}$ isomers is rather close, but relative energies of unstable $Sc_3N@C_{80}$ isomers are 20–40 kJ/mol smaller than those in C_{80}^{6-} . These data agree with formation of $Sc_3N@C_{80}-I_h(7)$ as the most abundant product of the $Sc_3N@C_{2n}$ clusterfullerene synthesis, followed by $Sc_3N@C_{80}-D_{5h}(6)$ as the second most abundant structure. With the increase of the cluster size to the Y_3N , the $I_h(7)$ isomer is still the most stable one, and the energy difference between $I_h(7)$ and $D_{5h}(6)$ isomers is almost the same as in the $Sc_3N@C_{80}$ series, but all other isomers of $Y_3N@C_{80}$ are significantly stabilized with respect to the $I_h(7)$ isomer. For instance, relative energies of the $D_{5h}(6)$ and non-IPR $C_1(28,325)$ isomers are quite close (70.2 and 90.0 kJ/mol, respectively). Similar to the case of $M_3N@C_{78}$, this can be explained by the relatively small size of almost spherical $C_{80}-I_h(7)$ and $C_{80}-D_{5h}(6)$ cages, which leads to the increase of the strain when large M_3N clusters are encapsulated. As a result, $M_3N@C_{80}-I_h(7)$ is still the most abundant nitride clusterfullerene for the clusters up to the size of Gd_3N , but overall yields are suppressed compared to the clusterfullerenes with smaller clusters (Krause and Dunsch 2005). For the larger cluster size ($M = Nd, Pr, Ce, La$). Echegoyen and coworkers have shown that the larger cages (in particular, C_{88}) are produced in higher yield than that of $M_3N@C_{80}$ (Chaur et al. 2008).

Violation of the Isolated Pentagon Rule in Endohedral Metallofullerenes

The specific feature of endohedral fullerenes is that the isolated pentagon rule, which imposes a strict limitation on the possible isomers of empty fullerenes, is often violated when a metal or a cluster is encapsulated inside the cage. Such possibility was first suggested by Kobayashi et al. in theoretical studies of $\text{Ca}@C_{72}$ isomers (Kobayashi et al. 1997; Nagase et al. 1999), and the first experimental evidence was provided in 2000 when $\text{Sc}_2@C_{66}-C_{2v}(4,059)$ (Wang et al. 2000a; Yamada et al. 2014) and $\text{Sc}_3\text{N}@C_{68}-D_3(6,140)$ (Stevenson et al. 2000) were isolated. Since that time, a lot of non-IPR endohedral fullerenes have been reported (Fig. 10), including $\text{Sc}_3\text{N}@C_{70}-C_{2v}(7,854)$ (Yang et al. 2007b), $\text{Sc}_2\text{S}@C_{70}-C_2(7,892)$ (Chen et al. 2013), $\text{Sc}_2\text{S}@C_{72}-C_s(10,528)$ (Chen et al. 2012), $\text{Sc}_2\text{C}_2@C_{72}-C_s(10,518)$ (Feng et al. 2013), $\text{La}@C_{72}-C_2(10,612)$ (Wakahara et al. 2006), $\text{La}_2@C_{72}-D_2(10,611)$ (Kato et al. 2003; Lu et al. 2008), $\text{DySc}_2\text{N}@C_{76}-C_s(17,490)$ (Yang et al. 2007a), $\text{La}_2@C_{76}-C_s(17,490)$ (Suzuki et al. 2013), $\text{Sm}@C_{76}-C_{2v}(19,138)$ (Hao et al. 2015), $\text{M}_3\text{N}@C_{78}-C_2(22,010)$ ($\text{M} = \text{Dy, Tm, Gd}$) (Beavers et al. 2009; Popov et al. 2007c), $\text{Gd}_3\text{N}@C_{82}-C_s(39,663)$ (Mercado et al. 2008), and $\text{Y}_2\text{C}_2@C_{84}-C_1(51,383)$ (Zhang et al. 2013), and $\text{M}_3\text{N}@C_{84}-C_s(51,365)$ ($\text{M} = \text{Tb, Gd, Tm}$) (Beavers et al. 2006; Zuo et al. 2008). The state of the art in the studies of non-IPR fullerenes has been reviewed (Tan et al. 2009). Recently, $\text{LaSc}_2\text{N}@C_{80}-C_s(\text{hept})$ with one heptagon and 13 pentagons (two pentalene units) has been synthesized (Zhang et al. 2015).

The tendency to violate the IPR in endohedral fullerenes can be explained by the changes of the relative energies of the fullerene isomers with the increase of the cage charge. It was shown already in 1995 that for C_{60} , the non-IPR $C_{2v}(1,809)$ isomer with two pairs of adjacent pentagons is substantially less stable than the IPR isomer in the neutrally charge state, but with the increase of the negative charge on the cage, the difference in the relative energies is diminishing, reaching almost zero for (-4) and inverting for (-6) states (Fowler and Zerbetto 1995). Semiempirical and DFT calculations for the whole set of 6,332 isomers of C_{68} in 0, (-2) , (-4) , and (-6) charge states were reported (Chen et al. 2008). This study has shown that the lowest energy isomers of C_{68}^{4-} and C_{68}^{6-} are $C_{2v}(6,073)$ and $D_3(6,140)$, in agreement with isolation of these isomers for $\text{Sc}_2\text{C}_2@C_{68}$ (Shi et al. 2006) and $\text{Sc}_3\text{N}@C_{68}$ (Olmstead et al. 2003; Stevenson et al. 2000), respectively. By plotting the relative energies of the isomers of the same charge versus the number of pentagon adjacencies, the authors have found that the penalty for a pentagon adjacency is decreasing from 92 kJ/mol in C_{68} to 70 kJ/mol in C_{68}^{6-} at the AM1 level, confirming that the increase of the charge stabilizes pentagon adjacencies.

An explanation of the stabilization of the pentagon adjacencies in metallofullerenes was proposed by Slanina et al. (2006). The authors noted that the pentalene (a unit of two fused pentagons) is an 8π anti-aromatic system in the neutrally charged state, but becomes 10π aromatic in the dianionic state (Zywietz et al. 1998). In a similar fashion, when the metal in the non-IPR endohedral

fullerenes is coordinated to the pentagon pair, it donates two electrons to the pentalene unit to make it aromatic. This reasoning agrees with the fact that the maximum number of pentagon pairs found in endohedral fullerenes is equal to the half of the formal charge of the cage: three pairs for (6⁻) charge as in Sc₃N@C₆₈-D₃(6,140) or Sc₃N@C₇₀-C_{2v}(7,854), two pairs for (4⁻) state as in Sc₂S@C₇₂-C_s(10,528) (Chen et al. 2012) or Sc₂C₂@C₇₂-C_s(10,518) (Feng et al. 2013), and one pair for (3⁻) and (2⁻) states as in La@C₇₂-C₂(10,612) or Sm@C₇₆-C_{2v}(19,138).

A correlation of the total number of adjacent pentagon pairs in the lowest energy isomers of C_{2n}⁶⁻ (2n = 68–88) with the cage size was revealed in Popov and Dunsch (2007). The authors found a pronounced tendency to decrease the number of pentagon pairs in the molecules with the increase of the cage size: while three pairs are common for C₆₈⁶⁻ and C₇₀⁶⁻ isomers, two and one pairs are preferable for C₇₂⁶⁻-C₇₈⁶⁻ and C₈₀⁶⁻-C₈₆⁶⁻, respectively. At the same time, relative energies of the non-IPR isomers are increasing with respect to the IPR isomers, and starting from C₈₆ formation of the non-IPR fullerenes is not expected (at least, for the cages with the formal charge of 6⁻ and smaller). Indeed, the largest cage reported to date for any non-IPR fullerene is C₈₄-C_s(51,365) (Beavers et al. 2006; Zuo et al. 2008). This tendency agrees with the increase of energy penalty for the adjacent pentagon pair in the empty fullerenes with increase of the cage size as discussed in the section “The Isolated Pentagon Rule and Steric Strain.” With the increase of the cage size, a more uniform distribution of the pentagon-induced strain over the fullerene is possible, and hence, localization of such a strain in pentagon adjacencies should become more unfavorable than for the smaller cages. Besides, for the charged cages, the influence of the charging of the fullerene should be diminished with the growth of the cage size, and hence its stabilizing role for the pentagon adjacencies is leveled down for larger fullerenes.

Structures and Stability of Fullerene Derivatives

Addition of X₂ to C₆₀: Isomers of C₆₀X₂ and General Considerations

The C₆₀ molecule has two types of C–C bonds, hexagon/hexagon (hex/hex) edges, and pentagon/hexagon (pent/hex) edges. Determination of their bond lengths by different experimental and theoretical methods gives the values of ca 1.40 and 1.45 Å, respectively. As might be expected, these values fall in the range of the bond lengths typical for conjugated systems; yet, they also show that there are formal *double* and *single* bonds in C₆₀, and it is more appropriate to call it polyalkene rather than superaromatic molecule. The chemical properties of fullerenes are to a large extent determined by the presence of the conjugated double bonds, and vast majority of fullerene reactions are *addition* reactions. We will not cover here the cycloaddition, mainly because quantum-chemical approaches are not so helpful in predicting and elucidation molecular structures of the reaction products. Thus, we will focus on the addition of atoms or groups that form single bonds to the fullerene

core (i.e., atoms or radicals, “X” hereafter). Such reactions, on the one hand, often occur at rather high temperatures allowing close-to-equilibrium distribution of the products, and on the other hand, addition of such groups allows multiple possibilities from the point of the distribution of the groups on the surface of the fullerene. For instance, there are 23 possible isomers of $C_{60}X_2$; 4,190 isomers of $C_{60}X_4$; and 418,470 isomers of $C_{60}X_6$, and the number of possible isomers is growing rapidly with the further increase of the number of the added groups (Balasubramanian 1991). It is therefore desirable to find the principles which govern addition of the addends to the fullerenes, the task to be naturally addressed by theoreticians. It should be also noted that equilibrium or close-to-equilibrium distribution of the products in the addition reaction (i.e., thermodynamic control) is a necessary prerequisite for the conclusions based on the relative energies of isomers to be valid. Otherwise, kinetic factors can be equally or even more important than the relative stability of the isomers, and theoretical predictions of the products are severely complicated.

All possible isomers which can be obtained by the addition of two hydrogen atoms to C_{60} (addition of the odd number of groups leads to the highly reactive radicals and therefore is not considered here) were studied at the PM3 level of theory by Dixon and coworkers (Matsuzawa et al. 1992). An AM1 study of the relative stabilities of $C_{60}X_2$ isomers for $X = H, F, Cl, Br, CH_3$, and $t-C_4H_9$ was reported in 2003 (Clare and Kepert 2003a). Let us consider several possible isomers of $C_{60}X_2$ (Fig. 11a). The simplest addition pathway is when two X groups are added to one double bond of C_{60} ; the rest of the π -system of the fullerene remains unchanged (in other words, this is 1,2-addition to a cyclohexatriene). If then two groups are attached to one hexagon in *para* position (1,4-addition to a cyclohexatriene), the π -system of the fullerene has to be adjusted by relocating one double bond to the pent/hex edge. All other variants of addition require more pronounced changes in the π -system of C_{60} and lead to the larger number of double bonds in pentagons

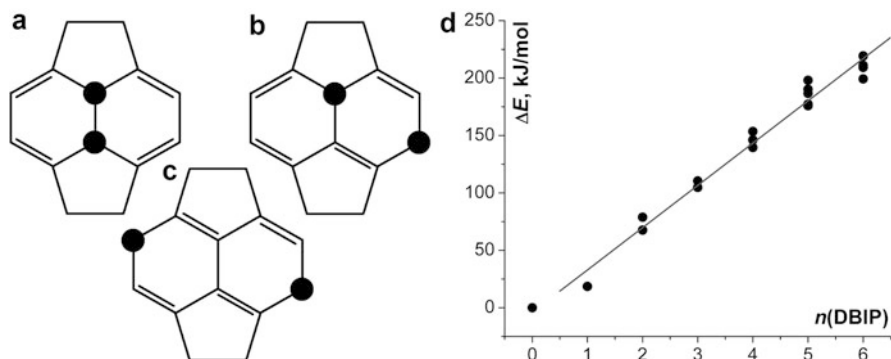


Fig. 11 (a–c) Fragments of the lowest energy isomers of $C_{60}H_2$: (a) 1,2-addition to hexatriene fragment; (b) 1,4-addition to hexatriene fragment; (c) 2,6-addition to naphthalene fragment; (d) relative energies of $C_{60}H_2$ isomers computed at the AM1 level as a function of the number of double bonds in pentagons (DBIPs) (Based on the data from Clare and Kepert 2003a)

(DBIPs). It was found that the number of DBIPs is an important parameter showing perfect correlation with the stability of the isomers of $C_{60}X_2$ (Fig. 11b): the larger the number of DBIPs, the lower the stability of the given isomer (Matsuzawa et al. 1992). From the slope of the fitted line, the penalty for each DBIP in $C_{60}H_2$ can be estimated as 36 and 37 kJ/mol at the PM3 and AM1 levels, respectively. The penalty is virtually independent on the nature of X, being 39 kJ/mol for $C_{60}F_2$ and 37 kJ/mol for $C_{60}(t-C_4H_9)_2$ at the AM1 level (Clare and Kepert 2003a). Obviously, “1,2”-addition should be the most preferable from the energetic point of view, and 1,2- $C_{60}X_2$ is indeed the most stable isomer for X = H and F (Dixon et al. 1992; Clare and Kepert 2003a) in good agreement with the experimental structures of $C_{60}H_2$ (Henderson and Cahill 1993) and $C_{60}F_2$ (Boltalina et al. 2000b). However, in the 1,2- $C_{60}X_2$, the X groups are experiencing eclipsing interactions. With the increase of the size of the groups, the influence of the steric factor is increasing and can balance the destabilizing effect of DBIP. Indeed, either semiempirical or DFT calculations show that for the bulky groups (such as Br or CF_3), 1,4-addition is more preferable (Dixon et al. 1992; Clare and Kepert 2003a; Goryunkov et al. 2003). Furthermore, for even larger groups such as $t-C_4H_9$, the isomer with two DBIPs and $t-C_4H_9$ groups on different hexagons (Fig. 11d) is the most stable at the B3LYP/6-31G**/AM1 level (at the same time, AM1 method predicts that 1,4- $C_{60}(C_4H_9)_2$ is the lowest energy isomer) (Clare and Kepert 2003a).

The studies of $C_{60}X_2$ isomers allow one to conclude that the structures of the products of the multiple addition of X groups should be a result of the interplay of at least two factors: (1) destabilizing double bonds in pentagons and (2) destabilizing eclipsing interactions of bulky groups. It can be expected that for the relatively small atoms such as H or F, multiple 1,2-addition is preferable, while for the bulky groups, multiple 1,4-addition is to be expected. Detailed studies of the multiple additions of various groups (in particular, H and Br) to fullerenes performed by Clare and Kepert in 1990s using AM1 method confirmed these suggestions (Clare and Kepert 1993, 1994a, b, c, 1995a, b, 1999a, b, 2002a; Kepert and Clare 1996). In brief, the methodology of the authors included a search of the most stable isomers of $C_{60}X_n$, sorting out majority of the unstable structures, and then a search of the most stable isomers of $C_{60}X_{n+2}$ based on the several most stable isomers of $C_{60}X_n$. Gradual increase of n allowed Clare and Kepert to cover broad range of compositions, reveal some general principles of the multiple addition to C_{60} and other fullerenes, and predict several addition patterns which were indeed found later in the experimental studies. In the remaining parts of this chapter, we will discuss several examples of the addition patterns to C_{60} and C_{70} .

Addition of H and F to C_{60} : Contiguous Addition, Benzene Rings, and Failures of AM1

Early on semiempirical calculations of $C_{60}H_2$ and $C_{60}F_2$ have shown that 1,2-isomers are 18–19 kJ/mol lower in energy than 1,4-isomers, and then all other structures are substantially less stable (Dixon et al. 1992; Matsuzawa et al. 1992;

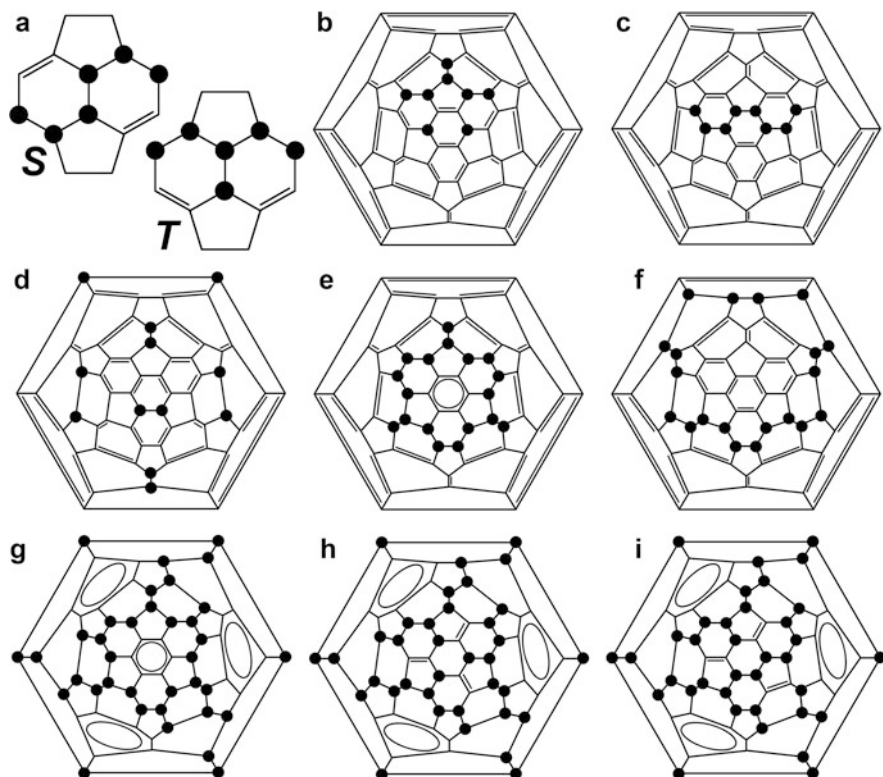


Fig. 12 Addition patterns typical for small groups (H, F). (a) *S* and *T* addition motifs; (b) lowest energy C_5 -isomer of $C_{60}F_8$; (c) C_3 - $C_{60}F_8$ with *S*-motif; (d) T_h - $C_{60}X_{12}$ with isolated CX–CX edges; (e) crown C_{3v} - $C_{60}X_{18}$ with isolate benzene ring; (f) D_{5d} - $C_{60}X_{20}$ with *S*-type loop around equator; (g) T - $C_{60}X_{36}$; (h) C_3 - $C_{60}X_{36}$; (i) C_1 - $C_{60}X_{36}$

Clare and Kepert 1993). Based on this fact, in their investigation of the multiple addition of hydrogen atoms to C_{60} , Clare and Kepert considered only 1,2-addition (addition to hex/hex edges). In the AM1 study (Clare and Kepert 1993), they have found that the most energetically favorable would be consequent addition of H_2 molecules to hex/hex edges in a pathway, which eventually leads to T_h -symmetric $C_{60}H_{12}$ with all CH–CH edges separated from each other (Fig. 12d). While this addition pattern has never been experimentally verified for $C_{60}H_{12}$ or $C_{60}F_{12}$, it was indeed proved for many hexakis-cycloadducts such as $C_{60}(C(COOEt)_2)_6$ (Lamparth et al. 1995; Hirsch and Vostrowsky 2001).

Further addition to hex/hex edges in T_h - $C_{60}H_{12}$ inevitably requires formation of longer $CH-(CH)_x-CH$ strings. It was, however, found that formation of C_6H_6 is energetically unfavorable, and $-(CH)_x-$ strings should be built up from the edge-sharing C_6H_4 hexagons. Two motifs of the string growth, *S* and *T* (Fig. 12a), have been indicated (Clare and Kepert 1994b).

Among many addition patterns studied by Clare and Kepert for $C_{60}H_n$ ($n \leq 20$), two are of particular interest. For $C_{60}H_{18}$, C_{3v} -symmetric “crown” structure built from combination of *S* and *T* motifs (Fig. 12e) was shown to be one of the most stable isomers (Clare and Kepert 1994b). In this molecule, $-(CH)_x-$ string forms a loop, which isolates one hexagon from the rest of the π -system. The isolated hexagon is highly aromatic, and the carbon cage is flattened in its neighborhood. Discovered first in 1994 in the theoretical study, in 1996 this addition pattern was proved by 1H NMR spectroscopy for the experimentally available $C_{60}H_{18}$ (Darwish et al. 1996), and later the same pattern was found in $C_{60}F_{18}$ (Neretin et al. 2000). Consecutive extension of the *S*-motif leads to the closure of the $-(CH)_{20}-$ string in a loop around the fullerene equator with formation of the D_{5d} -symmetric $C_{60}H_{20}$ (Fig. 12f) (Clare and Kepert 1994a). This addition pattern was found experimentally in $C_{60}F_{20}$ (Boltalina et al. 2001), whose structure was recently confirmed by a single-crystal X-ray diffraction study (Shustova et al. 2010).

The crown motif with isolated benzene-like ring proposed for $C_{60}H_{18}$ was then found common for highly hydrogenated C_{60} (Clare and Kepert 1994a, b, 1999a). In particular, $C_{60}H_{36}$ was studied in detail because a compound with this composition was among the first C_{60} derivatives ever synthesized by different methods (Hauffer et al. 1990; Attalla et al. 1993; Ruchardt et al. 1993), but its structure remained unclear for more than a decade. For $C_{60}H_{36}$, combination of several crown motifs can lead up to four isolated benzene ring as in the *T*-symmetric isomer (Fig. 12g), but in the AM1 studies, some isomers with two or three benzene rings were found to be equally stable (Clare and Kepert 1999a). Almost identical results were obtained at the AM1 level for isomers of $C_{60}F_{36}$ (Clare and Kepert 1999a), whose structure also could not be unambiguously elucidated in the first synthetic studies. It was also found in the studies of $C_{60}X_{36}$ and $C_{60}X_{48}$ that double bonds in pentagons are not destabilizing anymore at high degrees of functionalization (Clare and Kepert 1997, 1999a, 2002a). For instance, the most stable isomers of $C_{60}F_{48}$ have all their six double bonds in pentagons (Clare and Kepert 1997).

In early 2000s, when routine DFT calculations have become feasible for the fullerene derivatives, Clare and Kepert reconsidered some of their earlier results by comparing AM1 to B3LYP/6-31G**/AM1. One of the main conclusions of these studies was the understanding that one should treat results of AM1 calculations with caution.

First, it was found that AM1 underestimated the energy difference between 1,2- and 1,4-isomers of $C_{60}X_2$ by ca 10 kJ/mol (at least, when compared to the results of DFT calculations) and the error accumulated at higher degrees of addition. As a result, while AM1 calculations predicted the $C_{60}H_6$, $C_{60}H_{12}$, and $C_{60}H_{18}$ structures with, respectively, one, two, and three *skew pentagonal pyramidal* (SPP) motifs (Fig. 13d) to be the most stable isomers for their compositions (Clare and Kepert 1996), DFT calculations showed that these addition motifs are energetically unfavorable. For instance, $C_{60}H_{18}$ with three SPP motifs is 73.8 kJ/mol more stable than the “crown” C_{3v} - $C_{60}H_{18}$ isomer at the AM1 level, but the latter is 256.8 kJ/mol more stable at the B3LYP/6-31G**/AM1 level (Clare and Kepert 2003b). This

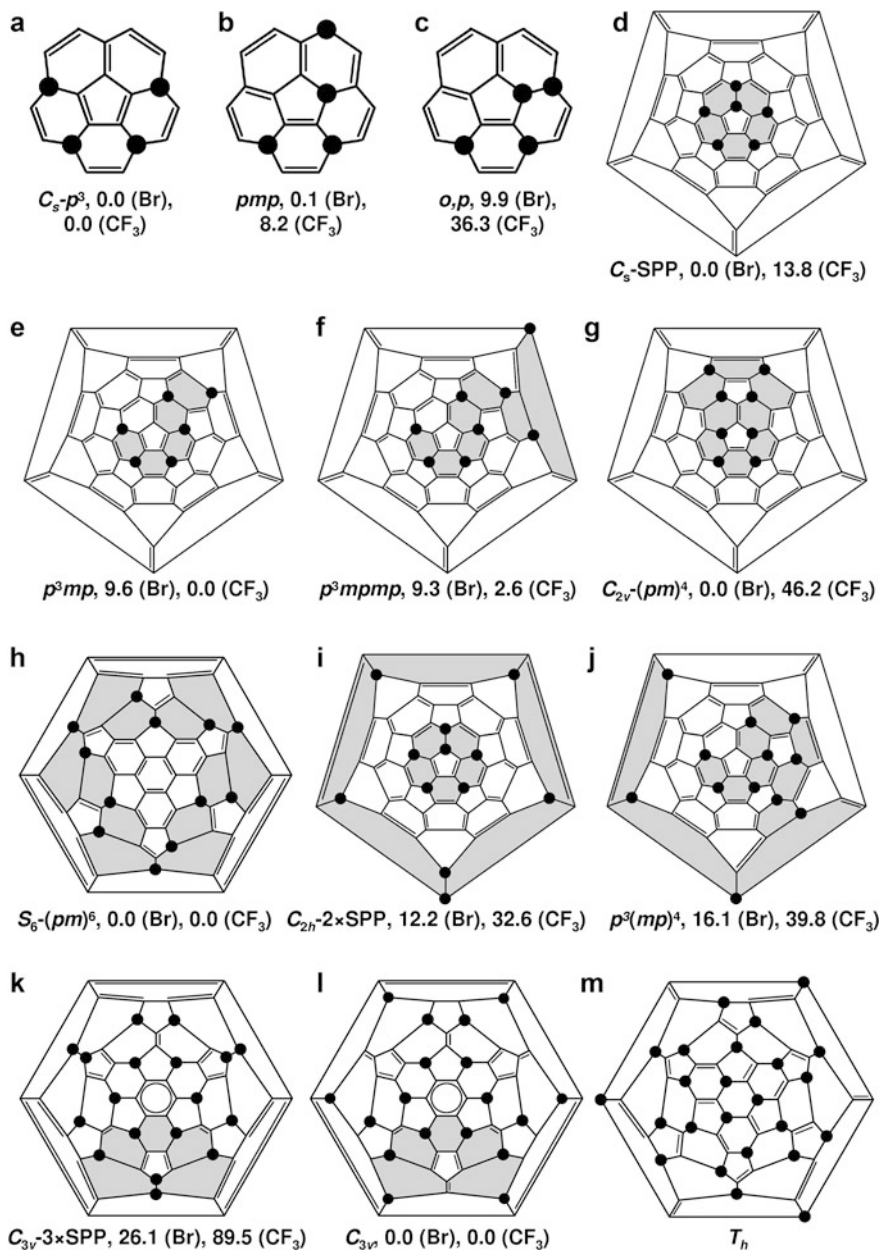


Fig. 13 $C_{60}X_n$ addition patterns typical for bulky groups (for $C_{60}Br_n$ and $C_{60}(CF_3)_n$, PBE/TZ2P-computed relative energies (kJ/mol) are listed). (a–c) isomers of $C_{60}X_4$; (d–e) isomers of $C_{60}X_6$; (d–e) isomers of $C_{60}X_8$; (h–j) isomers of $C_{60}X_{12}$; (k–l) isomers of $C_{60}X_{18}$; (m) T_h - $C_{60}X_{24}$. Symmetry group of the isomer is indicated only when it is different from C_1 . Ribbons of edge-shared C_6X_2 -hexagons are shaded (except for $C_{60}X_{18}$, here shaded are one SPP and one isolated fulvene moiety, respectively)

agrees with the fact that hydro- or fluorofullerenes with SPP motifs have never been observed experimentally.

Second, DFT studies have shown that for the low degree of additions, the most stable are those isomers, which have $-(CH)_x-$ strings, rather than isolated CH–CH fragments (Clare et al. 2003a). At the AM1 level, the most stable isomer of $C_{60}H_4$ is 13.4 kJ/mol level lower in energy than the C_s -symmetric isomer with C_6H_4 fragment, while at the B3LYP/6-31G* level, the latter is lower in energy by 15.3 kJ/mol. Likewise, the isomer of $C_{60}H_6$ with the *S*-motif is 31.5 kJ/mol lower in energy at the B3LYP/6-31G*//AM1 level than the isomer with three isolated CH–CH fragments, while at the AM1 level, the latter is more stable by 24 kJ/mol. For the isomers of $C_{60}H_8$, analogous energy differences increase to 59.7 and 31.1 kJ/mol, respectively. The preference of the contiguous addition at the early stages of fluorination was proved experimentally by ^{19}F NMR spectroscopy (Boltalina et al. 2002).

It is instructive to show that although contiguous addition appears to be the rule supported by both experimental and DFT studies, the other possibilities should not be completely abandoned. In the experimental studies of low C_{60} fluorides, C_s -symmetric isomer of $C_{60}F_8$ was isolated (Boltalina et al. 2002). C_s symmetry is not compatible with the *S*-type string, and the authors proposed a contiguous addition pattern with *T*-motif and two additional fluorine atoms attached at both ends of the string. However, later calculations have shown that such a structure would be unstable. Sandall and Fowler performed a broad search of all possible structures of C_s symmetry compatible with NMR data (Sandall and Fowler 2003). After calculations of more than 500 isomers using MNDO, AM1, and PM3 methods and – for selected structures – Harree-Fock method with STO-3G and 6-31G* basis sets, the authors have found that the most suitable isomer indeed has *T*-motif, but two additional fluorine atoms are added in a noncontiguous way with formation of 1,4- C_6F_2 hexagon (Fig. 12b). In a recent study, it was shown that two isomers of $C_{60}F_8$ are actually formed in the synthesis, aforementioned isomer with C_s symmetry and another one, also with C_s symmetry (Goryunkov et al. 2006b). Exhaustive DFT PBE/TZ2P calculations performed by the authors confirmed assignment of the C_s -isomer to the structure proposed by Sandall and Fowler and showed that another C_s -isomer is based on the *S*-motif. At the PBE/TZ2P level, *S*- C_s -isomer of $C_{60}F_8$ is 5.5 kJ/mol lower in energy.

The third failure of semiempirical methods revealed in the comparative AM1/DFT studies is that AM1 strongly underestimates stabilizing factor of the isolated benzene rings. Comparison of the relative energies of the isomers of $C_{60}H_{36}$ with different number of benzene rings computed at the AM1 and B3LYP/6-31G*//AM1 levels of theory shows that AM1-predicted relative energies should be adjusted by ca 60 kJ/mol per benzene ring to match the DFT-predicted values (Clare and Kepert 2002b). For instance, at the AM1 level, the relative energies of S_6 (2 rings), C_3 (3 rings), and *T* (4 rings) isomers of $C_{60}H_{36}$ are 0.0, 19.4, and 52.0 kJ/mol, while at the B3LYP/6-31G*//AM1 level, the values change to 0.0, –40.7, and –77.5 kJ/mol, respectively (Clare and Kepert 2002a). In a similar study of $C_{60}H_{36}$ isomers at the PM3 and PBE/TZ2P levels of theory, it was found that

PM3-computed relative energies should be corrected by approximately 67 kJ/mol per benzene ring to match the DFT results (Popov et al. 2004). For the isomers of $C_{60}F_{36}$, AM1-predicted relative energies should be corrected by approximately 30 kJ/mol per ring to match B3LYP/6-31G**/AM1 energies (Clare and Kepert 2002b). At the same time, relative energies of the isomers with the same number of benzene ring are reliably predicted by semiempirical approaches (Clare and Kepert 2002b; Popov et al. 2006). Thus, semiempirical methods were unable to predict the most stable isomers of $C_{60}H_{36}$ and $C_{60}F_{36}$. At the same time, DFT calculations correctly predict that *T*, C_3 , and C_1 isomers shown in Fig. 12g–i are the most stable isomers, in perfect agreement with the experimental results, which show that both $C_{60}H_{36}$ and $C_{60}F_{36}$ are obtained as mixtures of these three isomers (Boltalina et al. 1998; Gakh and Tuinman 2001; Gakh et al. 2003; Avent et al. 2002; Hitchcock and Taylor 2002; Popov et al. 2006). It should be noted that formation of isolated benzene rings is a very important stabilizing factor in the multiple addition to fullerenes, and it has been observed for many addends and many fullerenes. The failure of semiempirical methods in adequate prediction of the relative energies of isomers with such fragments is therefore a very serious problem and makes these methods hardly applicable when the broad search of the possible isomers is necessary. Good results demonstrated by DFTB in the studies of the fullerenes of different size show that this method might be useful for fullerene derivatives as well, but to our knowledge, benchmark calculations for fullerene derivatives with DFTB are not available yet.

Addition of Bulky Groups to C_{60} : Bromination and Perfluoroalkylation

In contrast to H and F, bromine atom is sufficiently large to make eclipsing interactions strongly repulsive, so that 1,4-addition of two Br atoms is more energetically preferable than 1,2-addition either at the AM1 ($\Delta E = 5.1$ kJ/mol) or DFT levels of theory (2.4 kJ/mol at B3LYP/6-31G**/AM1, 2.1 kJ/mol at LDA/DNP//AM1, and 14.0 kJ/mol at the PBE/TZ2P levels) (Dixon et al. 1992; Clare and Kepert 1995a, 2001). At the same time, the difference in the energy of 1,2- and 1,4-isomers is not very large, and hence the possibility of 1,2-addition should be also considered. Indeed, in the AM1 study of $C_{60}Br_4$ isomers, it was found that two lowest energy isomers are results of the consecutive 1,4-additions, while the third most stable isomer was a result of a combined 1,2- and 1,4-additions (Fig. 13a–c; Clare and Kepert 1995a). Regression analysis of the relative energies of $C_{60}Br_4$ isomers as a function of the number of different structural motifs (such as C_6Br , C_6Br_2 , C_6Br_3 , etc. hexagons) revealed that the most stable isomer of bromofullerenes might be obtained through formation of ribbons (also referred to as “strings”) of edge-sharing C_6Br_2 hexagons, rather than by addition of pairs of bromine atoms to the distant parts of C_{60} . Thus, the most stable isomer has *para-para-para* string (p^3 hereafter), while the second most stable isomer has *para-meta-para* string (*pmp* hereafter) (Fig. 13a, b). The relative energy of *pmp* isomer at the PBE/TZ2P (AM1) level

of theory with respect to the p^3 isomer is 0.1 (6.7) kJ/mol, while the next in stability order with one *ortho*- and one *para*-C₆Br₂ hexagons is 9.9 (11.0) kJ/mol less stable (the DFT calculations reported hereafter for bromofullerenes at the PBE/TZ2P level with effective-core potential basis set for the Br atoms were performed specifically for this chapter).

Further addition of Br atoms to the three most stable isomer of C₆₀Br₄ confirmed the preference of the string formation (Clare and Kepert 1995a). However, the most stable isomer at the AM1 level has a special C_s-symmetric addition pattern titled by some authors as “skew pentagonal pyramid” (SPP, Fig. 13d). This addition pattern includes one pair of Br atoms bonded to neighboring carbon atoms and is 21.8 kJ/mol more stable than the second most stable isomer with a p^3mp ribbon (Fig. 13e). At the PBE/TZ2P level, the energy difference between SPP and p^3mp isomers is reduced to 9.6 kJ/mol. Experimentally isolated C₆₀Br₆ indeed has the SPP pattern (Birkett et al. 1992), and this addition motif was also found for some other addends, including Cl (Birkett et al. 1993) and CF₃ (Kareev et al. 2006). It should be noted that formation of SPP-C₆₀X₆ derivatives for relatively large X is also favorable kinetically if the radical addition mechanism is adopted (Rogers and Fowler 1999).

Next two added Br atoms continue the tendency of the ribbon formation. At the AM1 level, the most stable isomer of C₆₀Br₈ has p^3mpmp ribbon (Fig. 13f), while DFT (PBE/TZ2P) shows that the isomer in which the ribbon is closed to a C_{2v}-symmetric loop ((*pm*)⁴, Fig. 13g) is 9.3 kJ/mol more stable. The latter motif is indeed observed in the experimentally isolated C₆₀Br₈ (Birkett et al. 1992). AM1 calculations have also shown that the isomers with the SPP motif and an additional *para*-C₆Br₂ hexagon on the opposite part of C₆₀ are also equally stable (Clare and Kepert 1995a).

While there are no experimentally isolated C₆₀Br_x compounds between C₆₀Br₈ and C₆₀Br₂₄, theoretical studies of the intermediate compositions were still very useful for predicting the general trends in the addition of bulky groups (Clare and Kepert 1995a, b). In particular, a remarkable prediction was done for C₆₀Br₁₂. The authors have shown that the most stable isomers in SPP and ribbon series are the C_{2h}-isomer with two SPP motifs on the opposite sides of C₆₀ and the S₆-symmetric isomer in which the ribbon is closed to a loop around the equator of the carbon cage (Fig. 13h, i). At the AM1 level, C_{2h} isomer is 6.8 kJ/mol more stable, while PBE/TZ2P favors the S₆-isomer by 12.2 kJ/mol. Though not available experimentally for C₆₀Br₁₂, both addition patterns have been found for C₆₀(CF₃)₁₂ derivatives (Popov et al. 2007b; Troyanov et al. 2006). In further addition of Br atoms, the competition between SPP motifs and ribbons is continued up to C₆₀Br₁₈, for which the isomer with three SPP moieties was predicted to be the lowest energy structure (Clare and Kepert 1995b). For even larger number of Br atoms, SPP motif does not provide stable structures anymore, and the most stable isomer of C₆₀Br₂₄ is T_h-symmetric structure in which all Br atoms are bonded to the non-neighboring carbon atoms (24 is the largest number of addends for which such a restriction is possible) (Clare and Kepert 1995a; Fowler et al. 1999). This isomer corresponds to the experimentally available structure of C₆₀Br₂₄ (Tebbe et al. 1992).

Significant progress achieved in the last years in the perfluoroalkylation of fullerenes (Boltalina et al. 2015) initiated exhaustive theoretical studies of $C_{2n}(CF_3)_x$ isomers with the aim to find the most stable structure(s) for each composition. Since the studies were performed almost a decade after the AM1-based analysis of $C_{60}Br_x$ addition patterns by Clare and Kepert, they have strongly benefited from the serious progress in hardware. Instead of the maximum of about 100 structures reported in the early exploratory studies, up to thousands and tens of thousand isomers were routinely studied for CF_3 derivatives, and results of the AM1 studies were then verified by DFT calculations. Besides, the isolation of $C_{60}(CF_3)_n$ derivatives with each even n from 2 to 18 (in many cases each composition is also presented by several isomers) enabled more detailed comparison between experimental and theoretical results.

Both AM1 and DFT (PBE/TZ2P) unambiguously show that 1,4-isomer of $C_{60}(CF_3)_2$ is much more stable than the 1,2-isomer, the DFT relative energy of the latter being 34.7 kJ/mol (Goryunkov et al. 2004a). The energy difference is more pronounced than for the $C_{60}Br_2$, which clearly shows that CF_3 is bulkier than bromine. Therefore, in the compounds with larger number of CF_3 groups, 1,2-addition is to be avoided. Indeed, the most stable isomers of $C_{60}(CF_3)_4$ are p^3 ($\Delta E = 0.0$ kJ/mol at the PBE/TZ2P level) and pmp ($\Delta E = 8.2$ kJ/mol) structures, just like for $C_{60}Br_4$; however the o,p - $C_{60}(CF_3)_4$ isomer, the third stable for $C_{60}Br_4$, is 36.3 kJ/mol less stable than p^3 - $C_{60}(CF_3)_4$ (for $C_{60}Br_4$ its relative energy is 9.9 kJ/mol) (Goryunkov et al. 2004a). The pmp is the major isolable isomer of $C_{60}(CF_3)_4$, while the more stable p^3 isomer is very reactive (Whitaker et al. 2013) and is usually isolated in the form of its epoxide $C_{60}(CF_3)_4O$ (Goryunkov et al. 2003; Kareev et al. 2006).

The trend revealed for $C_{60}(CF_3)_4$ in comparison to $C_{60}Br_4$ is further emphasized for six added groups. For CF_3 , the ribbon addition motifs are noticeably more preferable than the SPP motif. The most stable isomer of $C_{60}(CF_3)_6$ has p^3mp addition pattern, while the SPP isomer is 13.8 kJ/mol less stable (for $C_{60}Br_6$, SPP is 9.6 kJ/mol lower in energy) (Dorozhkin et al. 2007). The SPP- $C_{60}(CF_3)_6$ still can be isolated, but its yield is ca 20 times lower than the yield of the major, p^3mp isomer (Kareev et al. 2006).

Another difference between CF_3 and Br which becomes apparent for further additions is that CF_3 groups are sufficiently bulky to experience repulsive interaction when two groups are in one pentagon, even though they are not bonded to the neighboring carbon atoms. Thus, the stable ribbon addition pattern of $C_{60}Br_x$ with C_5Br_2 pentagons are destabilized for CF_3 groups (although some stable isomer still can have them). For instance, the most stable C_{2v} -symmetric addition pattern of $C_{60}Br_8$ (9.3 kJ/mol more stable than p^3mpmp) is 46.2 kJ/mol less stable than the p^3mpmp isomer of $C_{60}(CF_3)_8$ at the PBE/TZ2P level (Goryunkov et al. 2007). The p^3mpmp isomer itself is the second most stable ($\Delta E = 2.6$ kJ/mol, PBE/TZ2P), while the lowest energy was predicted for one of the isomers with p^3mp,p addition pattern (i.e., addition pattern of the lowest energy p^3mp - $C_{60}(CF_3)_6$ isomer with an additional isolated $para$ - $C_6(CF_3)_2$ hexagon). Theoretical results agree well with

experimental isolation of one p^3mpmp , three p^3mp,p , and one $pmpmpmp$ isomers of $C_{60}(CF_3)_8$ (Goryunkov et al. 2007; Popov et al. 2007b).

Fortuitously, performance of the AM1 method for addition of bulky groups is much better than for the consecutive 1,2-additions of H and F, which encourages the use of AM1 for the prescreening of thousand isomers of $C_{60}(CF_3)_n$. Although there is rather poor correlation between AM1 and PBE/TZ2P values (Fig. 14a), the trend in the energies of the most stable isomers is correctly predicted by AM1. For instance, when all isomers of $C_{60}(CF_3)_{10}$ obtained by a combination 1,4-additions were studied at the AM1 level and then a cutoff of 40 kJ/mol was applied to the AM1 results for the consequent PBE/TZ2P calculations, the most stable isomers within the range of ca 15 kJ/mol could be identified (Popov et al. 2007b). Figure 14 shows addition patterns and relative energies of the eight most stable isomers of $C_{60}(CF_3)_{10}$. Except for one isomer, which has two symmetry-related p^3m^2 loops, all other structures have ribbon addition patterns similar to those of $C_{60}(CF_3)_8$. From the eight theoretically predicted structures in Figure 14, six isomers of $C_{60}(CF_3)_{10}$ are characterized experimentally by single-crystal X-ray diffraction and/or ^{19}F NMR spectroscopy. The splitting of the ^{19}F NMR lines due to the through-space interaction of the fluorine atoms of CF_3 groups in the same hexagon allows experimental identification of the lengths of the ribbon, and when combined with results of DFT calculations of the relative energies, these data allow unambiguous identification of the addition pattern. Molecular structures of several isomers of $C_{60}(CF_3)_{10}$ were elucidated this way by a combination of NMR and DFT, and some of them were later confirmed by X-ray studies (Popov et al. 2007b).

Two key addition motifs predicted by Clare and Kepert for $C_{60}Br_{12}$, double-SPP pattern and S_6 -symmetric loop (Fig. 13h, i), have been identified experimentally for $C_{60}(CF_3)_{12}$ (Popov et al. 2007b; Troyanov et al. 2006). Extensive AM1 and DFT calculations showed that the S_6 -loop isomer is at least 19 kJ/mol more stable than all other isomer of $C_{60}(CF_3)_{12}$, while the C_{2h} -($2 \times$ SPP) isomer is 32.6 kJ/mol less stable (Kareev et al. 2007). The isomer of $C_{60}(CF_3)_{12}$ with ribbon motif and two $C_5(CF_3)_2$ pentagons (Fig. 13j) has been found to be 39.8 kJ/mol less stable than the S_6 -isomer, showing that kinetic factors can be also important in the isomeric distribution of $C_{60}(CF_3)_x$ derivatives even at 500 °C (Kareev et al. 2007).

When addition of more than 12 CF_3 groups is considered, one has to consider that fullerenes have only 12 pentagons, and hence it is impossible to avoid formation of destabilizing $C_5(CF_3)_2$ pentagons. Therefore, when CF_3 groups are too crowded on the fullerene surface, results of 1,2-addition can be not as destabilizing as at the lower addition stages. Exhaustive AM1 and DFT calculations have shown that for $C_{60}(CF_3)_{14}$, the lowest energy isomer with one SPP fragment is only 0.4 kJ/mol less stable than the most stable ribbon isomer with two $C_5(CF_3)_2$ pentagons; however, experimentally characterized isomers of $C_{60}(CF_3)_{14}$ do not have SPP fragments (Omelyanyuk et al. 2007). Likewise, for $C_{60}(CF_3)_{16}$, SPP and ribbon isomers are equally stable; moreover, it was found that stable addition patterns with CF_3 groups on adjacent carbon atoms can be realized without formation of the SPP moiety, and one of the experimentally characterized $C_{60}(CF_3)_{16}$ isomers has o - $C_6(CF_3)_2$

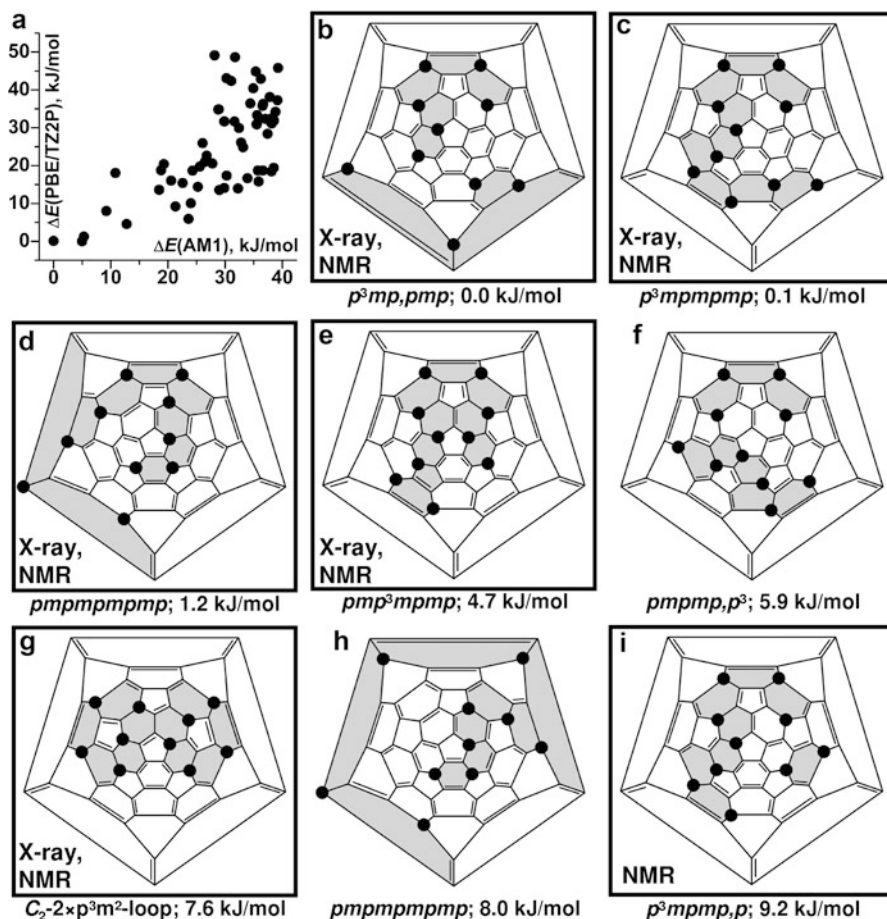


Fig. 14 (a) Correlation between relative energies of $\text{C}_{60}(\text{CF}_3)_{10}$ isomers computed at the AM1 and PBE/TZP levels of theory; (b–i) Schlegel diagrams of the eight lowest energy isomers of $\text{C}_{60}(\text{CF}_3)_{10}$ (PBE/TZ2P-computed relative energies are also listed, based on the data from Popov et al. 2007a). Experimentally characterized compounds are framed in black (“X-ray” means single-crystal X-ray diffraction; “NMR” means ^{19}F NMR spectroscopy). Ribbons of edge-shared $\text{C}_6(\text{CF}_3)_2$ -hexagons are shaded

hexagon (Trojanov et al. 2007). For $\text{C}_{60}(\text{CF}_3)_{18}$, theoretical studies have shown that the most stable isomer has C_{3v} -symmetry, two isolated benzene rings, three isolated fulvene fragments, nine $\text{C}_5(\text{CF}_3)_2$ pentagons, and no CF_3 groups on adjacent carbon atoms (Fig. 13j; Trojanov et al. 2007). DFT calculations show that this isomer is 89.5 kJ/mol more stable than the isomer with triple SPP motif, which also has one benzene ring (Fig. 13k). Moreover, even for $\text{C}_{60}\text{Br}_{18}$ this isomer is 26.1 kJ/mol more stable than the structure with three SPP moieties. Recently, the isomer of $\text{C}_{60}(\text{CF}_3)_{18}$ with this addition pattern has indeed been isolated and characterized

by single-crystal X-ray diffraction (Samokhvalova et al. 2009). Up to now, 18 is the largest number of CF_3 groups in structurally characterized $\text{C}_{60}(\text{CF}_3)_x$ derivatives; the T_h -symmetric $\text{C}_{60}\text{X}_{24}$ motif known for $\text{X} = \text{Br}$ and Cl could not be reached yet for CF_3 derivatives.

Addition to C_{70} and Higher Fullerenes

I_h - C_{60} has only two types of C–C bonds, which can be straightforwardly classified according to their bond lengths as “single” and “double” bonds. For C_{70} and higher fullerenes, the symmetry in most cases is much lower, and the lengths of C–C bonds are more uniformly distributed. It is therefore almost impossible to make an unambiguous classification of the bond types, and the relevant chemical properties (such as addition pathways) are much harder to predict.

The structure of C_{70} can be described as a combination of two C_{60} -like hemispheres on the poles separated by the belt of ten carbon atoms and five benzene-like rings around the equator (Fig. 15a). The bonds between the belt carbon atoms are significantly elongated compared to C_{60} values; besides, these atoms are located on triple hexagon junctions. Clare and Kepert reported AM1 study of the addition patterns of C_{70}X_n derivatives ($n = 2$ –12, $\text{X} = \text{H}, \text{F}, \text{Br}, \text{C}_6\text{H}_5$) (Clare and Kepert 1999b). For C_{70}H_2 , the authors have found that 1,2-addition is energetically preferable in the pole region, while the lowest energy 1,4-addition occurs at the equator. The AM1-computed energy difference between the most stable 1,2- and 1,4-isomers is only 4.4 kJ/mol, which can be compared to 18 kJ/mol for C_{60}H_2 . Such a considerable difference between the relative energies of C_{70}H_2 and C_{60}H_2 isomer can be explained by the fact that the 1,4-addition at the equator of C_{70} does not generate destabilizing double bond in any pentagon. For C_{70}H_4 , the isomer with p^3 -string at the equator (Fig. 15) is in fact 9.1 kJ/mol lower in energy than the isomer obtained by two 1,2-additions. The studies of C_{70}H_n with $n = 6$ –10 further emphasized the preference of consecutive 1,4-additions with formation of the p^{n-1} -ribbon along the equator of the carbon cage. The most stable C_{70}H_6 and C_{70}H_8 isomers in these series are, respectively, 28.0 and 54.3 kJ/mol lower in energy than the isomers obtained by 1,2-addition. Note that p^5 and p^7 -strings at the equator of C_{70} can be realized as either C_s or C_2 -symmetric isomers. For C_{70}H_6 , C_2 - p^5 isomer is 76.6 kJ/mol more stable than the C_s - p^5 pattern (56.6 kJ/mol at the PBE/TZ2P level (Dorozhkin et al. 2006a), while for C_{70}H_8 , the C_s - p^7 isomer is 34.7 kJ/mol (21.6 kJ/mol at the PBE/TZ2P level) more stable than C_2 - p^7 . The most stable isomer at the next addition step is obtained by the closure of the p^7 ribbon via addition of two atoms to pent/hex junction with formation of the C_s - p^9o -loop (Fig. 15), which cuts the π -system of the fullerene into two noncommunicating parts. This structure is 26.9 kJ/mol more stable than the C_2 - p^9 isomer, obtained by the growth of the equatorial ribbon via one more 1,4-addition step. Computations for the bulkier Br and C_6H_5 groups have also shown that formation of the ribbon around the cage equator is the most referable addition motif in C_{70} (Clare and Kepert 1999b). Even for C_6H_5 , the sterically strained C_s - p^9o -loop addition pattern

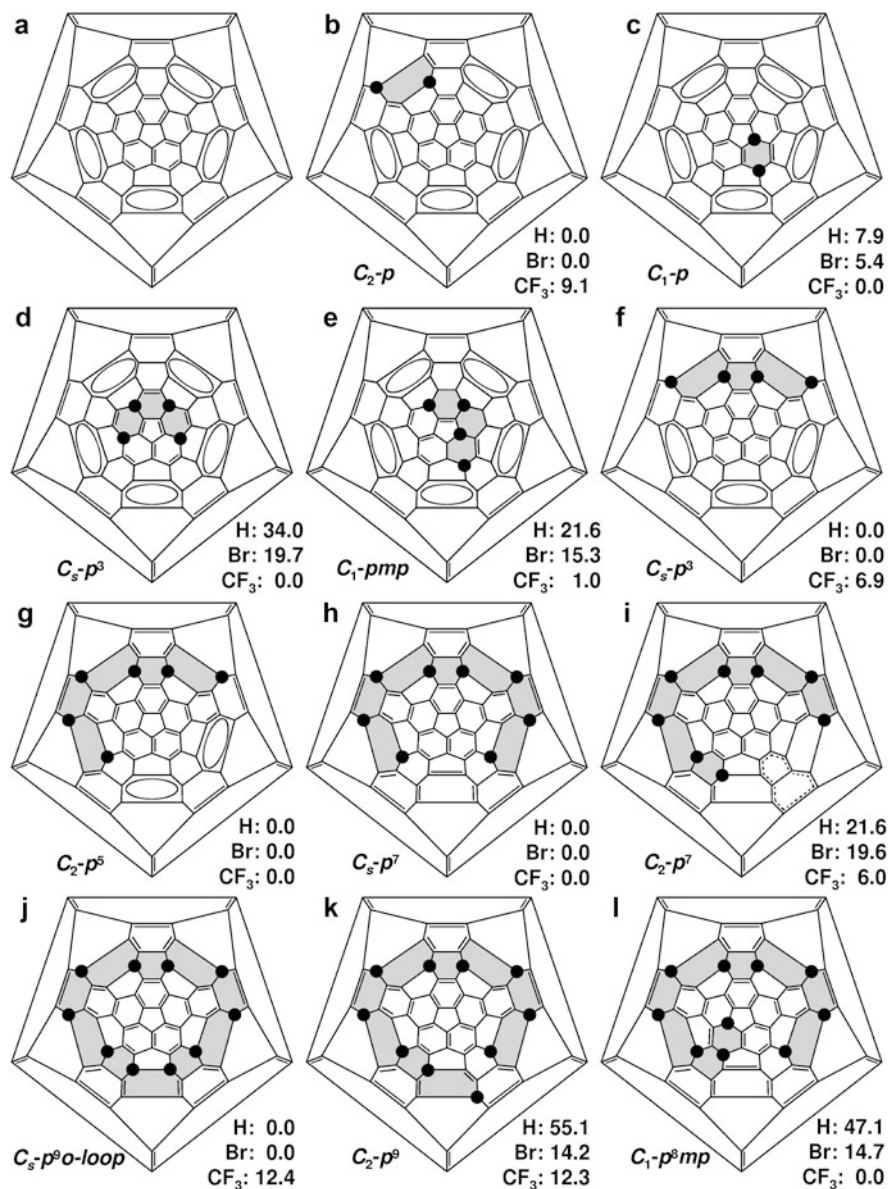


Fig. 15 (a) Schlegel diagram of C_{70} ; (b–i) addition patterns of selected $C_{70}X_n$ ($n = 2-10$) derivatives and the relative energies of their isomer (kJ/mol, PBE/TZ2P level of theory) for $X = H, Br,$ and CF_3 ; ribbons of edge-shared C_6X_2 -hexagons are shaded. Based on the data from (Dorozhkin et al. 2006a) and additional computations performed for this chapter

is more stable for $C_{70}X_{10}$ than C_2-p^9 isomer. Results of these calculations agree very well with experimental data. Except for $C_{70}(CF_3)_{10}$ discussed below, the majority of experimentally available $C_{70}X_{10}$ compounds ($X = H, Cl, Br, Ph, CH_3$) have C_s-p^9o -loop (Birkett et al. 1995; Avent et al. 1997; Spielmann et al. 2000; Al-Matar et al. 2002; Troyanov et al. 2003), and only bulky OO^tBu groups form the decakis adduct with C_2-p^9 addition pattern (Xiao et al. 2005). Some $C_{70}X_n$ compounds with less than ten added groups have also been reported, and most of them also have equatorial string addition patterns. Thus, both computational and experimental data confirm that the most energetically stable addition motif for C_{70} is formation of the string from *para*- C_6X_2 hexagons around the cage equator, and even for hydrogen this motif might be more preferable than the consecutive 1,2-additions.

Trifluoromethylation of C_{70} deserves separate attention since, like for C_{60} , this reaction was studied recently in great details, and many compounds from $C_{70}(CF_3)_2$ to $C_{70}(CF_3)_{20}$ (each composition being presented by several isomers) have been isolated and structurally characterized (Goryunkov et al. 2005, 2006a; Kareev et al. 2005; Dorozhkin et al. 2006a, b; Ignat'eva et al. 2006; Popov et al. 2007a; Boltalina et al. 2015). The experimental studies were usually combined with and sometimes preceded by detailed theoretical analysis at the AM1 and PBE/TZ2P level.

It was found that the two lowest energy isomers of $C_{70}(CF_3)_2$ have a *p*- $C_6(CF_3)_2$ hexagon on the pole of the molecule, and these structures are assigned to the experimentally available isomer (Dorozhkin et al. 2006a; Popov et al. 2007a). The most stable isomer with equatorial addition is 9.1 kJ/mol higher in energy at the PBE/TZ2P level of theory; meanwhile, the same method shows that the isomer of $C_{70}H_2$ with equatorial addition is 7.9 kJ/mol more stable than the isomer with *p*- C_6H_2 hexagon on the pole. The reason for such a difference is most probably a much higher curvature of C_{70} cage at the poles as compared to the almost flat equator. Therefore, the distance between two CF_3 groups is longer for *p*- $C_6(CF_3)_2$ hexagon on the pole than for the hexagon on the equator; hence the repulsion between the groups is weaker on the pole. For $C_{70}H_2$, repulsion between the atoms is a less important factor, and the equatorial addition is more preferable.

Addition on the pole is more preferable for $C_{70}(CF_3)_4$ as well. Similar to $C_{60}(CF_3)_4$, the most stable isomers have C_s-p^3 and C_1-pmp ribbons with the energy difference of only 1.0 kJ/mol (PBE/TZ2P) (Dorozhkin et al. 2006a). Both can be isolated experimentally; however the p^3 isomer is kinetically unstable, and the isolable form is its epoxide $C_{70}(CF_3)_4O$ (Popov et al. 2007a). DFT calculations have also shown that there are at least 15 other isomers of $C_{70}(CF_3)_4$ within the range of 20 kJ/mol, including the isomer with C_s-p^3 string at the equator ($\Delta E = 6.9$ kJ/mol, PBE/TZ2P), but no other structure has been characterized experimentally. Note that the model calculations of $C_{70}H_4$ and $C_{70}Br_4$ isomers at the PBE/TZ2P level have also shown that the equatorial addition is more energetically preferable for H and even Br (relative energy of the pole C_s-p^3 isomer is 19.7 and 34.0 kJ/mol for Br and H, respectively, versus the equatorial C_s-p^3 isomer).

Starting from $C_{70}(CF_3)_6$, CF_3 addition switches to the equatorial motif. Like for many other groups (H, Cl, Br, Ph), the lowest energy isomer has a C_2-p^5 ribbon at the equator, and this is indeed the most abundant experimentally characterized isomer (Dorozhkin et al. 2006a). For $C_{70}(CF_3)_8$, the two most stable isomers have C_s-p^7 and C_2-p^7 ($\Delta E = 6.0$ kJ/mol, PBE/TZ2P) ribbons, and both isomers could be isolated and characterized by single-crystal X-ray diffraction (Goryunkov et al. 2005; Mutig et al. 2008).

For $C_{70}(CF_3)_{10}$, the large size of CF_3 groups again makes the lowest energy addition pattern different from that for other addends. Repulsion of CF_3 groups on adjacent carbon atoms destabilizes the C_s-p^9 *o*-loop isomer of $C_{70}(CF_3)_{10}$, and the most stable isomer is obtained if the C_s-p^7 equatorial ribbon is continued by the *mp* fragment which is crawling to the pole again, as in the early stages of addition, and forming as a result C_1-p^7mp ribbon (Kareev et al. 2005). The C_2-p^9 string is 12.3 kJ/mol less stable followed by C_s-p^9 *o*-loop (12.4 kJ/mol). For comparison, the C_1-p^7mp isomer of $C_{70}Br_{10}$ is 14.7 kJ/mol less stable than the C_s-p^9 *o*-loop. In accordance with its DFT-predicted high thermodynamic stability, the C_1-p^7mp is the most abundant $C_{70}(CF_3)_{10}$ isomer, and it can be obtained with unprecedented high yield (Popov et al. 2007a). C_2-p^9 as well as three other less-stable minor isomers could be also structurally characterized, but their yield is much lower; the C_s-p^9 *o*-loop has never been found (Popov et al. 2007a).

The stable addition pattern of the major $C_{70}(CF_3)_{10}$ isomer is found in all except one lowest energy and experimentally isolated addition patterns of $C_{70}(CF_3)_{12-18}$. For instance, four isolated isomers of $C_{70}(CF_3)_{12}$ have p^7mp,p patterns, while $C_{70}(CF_3)_{16}$ and $C_{70}(CF_3)_{18}$ have SPP fragment on the pole with p^7mp and p^7mpmp ribbons, respectively (Avdoshenko et al. 2006; Goryunkov et al. 2006a; Ignat'eva et al. 2006).

One of the important findings revealed in the studies of the multiple addition to C_{70} is the fact that the isomers in which addends are bonded to the carbon atoms on the triple hexagon junctions (THJs) are very unstable. This rule is fulfilled for all C_{70} derivatives with low and medium addition rate, and from all known derivatives of C_{70} , the only violations of this rule are $C_{70}F_{38}$ (Hitchcock et al. 2005) and $C_{70}Cl_{28}$ (Troyanov et al. 2005). Moreover, it was found that this is also a strict rule for higher fullerenes (Troyanov and Kemnitz 2005; Shustova et al. 2006, 2007; Kareev et al. 2008b) when 12 or less groups are added. The energy penalty for addition of one CF_3 group to THJ position was estimated as 60 kJ/mol (Boltalina et al. 2015). Thus, when considering the possible isomers of $C_{2n}X_m$ ($m \leq 12$), one can exclude addition to triple hexagon junctions. This fact can dramatically reduce the number of possible isomers. For instance, in the experimental study of the trifluoromethylation of the “insoluble” small HOMO–LUMO gap fullerenes, the conditions were optimized to produce $C_{2n}(CF_3)_{12}$ as major products (Shustova et al. 2006). ^{19}F NMR spectroscopy then enabled identification of the symmetry of the derivatives and the relative location of the groups (i.e., number and length of the ribbons). Note that the cage isomers of the studied fullerenes were unknown.

As a result, million isomers were possible for these derivatives. However, when the rule about the triple hexagon junctions was taken into account and combined with the condition that CF_3 groups cannot share the same pentagon (12 is the largest number of the groups added to any fullerene which can be distributed with one group per pentagon), the number of possible isomers was reduced to only a few structures for all possible cage isomers, and the DFT studies of all compatible structures could easily identify the experimental structures and determine their carbon cage isomers (Shustova et al. 2006). Note that for higher degree of addition (more than 12 groups), the avoidance of THJ position by addends is not a strict rule anymore, and CF_3 -bearing THJ carbons have been found in several derivatives of higher fullerenes ($\text{C}_{76}(\text{CF}_3)_{14-18}$, $\text{C}_{94}(\text{CF}_3)_{20}$). Interestingly, in contrast to the empty fullerenes, multiple addition of CF_3 groups to THJ carbon atoms was found for $\text{Sc}_3\text{N}@C_{80}$ (Shustova et al. 2009, 2011) already at early stages of addition proving that the endohedral cluster and electron transfer to the carbon cage dramatically change chemical properties of fullerenes.

Bibliography

- Achiba, Y., Kikuchi, K., Aihara, Y., Wakabayashi, T., Miyake, Y., & Kainosho, M. (1995). *Higher fullerenes – Structure and properties. Science and technology of fullerene materials*. Pittsburg: Materials Research Society symposium proceedings.
- Achiba, Y., Fowler, P. W., Mitchell, D., & Zerbetto, F. (1998a). Structural predictions for the C_{116} molecule. *Journal of Physical Chemistry A*, 102(34), 6835–6841.
- Achiba, Y., Miyake, Y., Ishiwatari, H., Kainosho, M., & Kikuchi, K. (1998b). *NMR characterization of higher fullerenes up to C94*. Boston: Materials Research Society 1998 Fall Meeting.
- Akasaka, T., Wakahara, T., Nagase, S., Kobayashi, K., Waelchli, M., Yamamoto, K., Kondo, M., Shirakura, S., Okubo, S., Maeda, Y., Kato, T., Kako, M., Nakadaira, Y., Nagahata, R., Gao, X., Van Caemelbecke, E., & Kadish, K. M. (2000). $\text{La}@C_{82}$ anion. An unusually stable metallofullerene. *Journal of the American Chemical Society*, 122(38), 9316–9317.
- Albertazzi, E., Domene, C., Fowler, P. W., Heine, T., Seifert, G., Van Alsenoy, C., & Zerbetto, F. (1999). Pentagon adjacency as a determinant of fullerene stability. *Physical Chemistry Chemical Physics*, 1(12), 2913–2918.
- Al-Matar, H., Sada, A. K. A., Avent, A. G., Taylor, R., & Wei, X. W. (2002). Methylation of [70]fullerene. *Journal of the Chemical Society-Perkin Transactions*, 2, 1251–1256.
- Alvarez, M. M., Gillan, E. G., Holczer, K., Kaner, R. B., Min, K. S., & Whetten, R. L. (1991). La_2C_{80} – A soluble dimetallofullerene. *Journal of Physical Chemistry*, 95(26), 10561–10563.
- Alvarez, L., Pichler, T., Georgi, P., Schwieger, T., Peisert, H., Dunsch, L., Hu, Z., Knupfer, M., Fink, J., Bressler, P., Mast, M., & Golden, M. S. (2002). Electronic structure of pristine and intercalated $\text{Sc}_3\text{N}@C_{80}$ metallofullerene. *Physical Review B*, 66(3), 035107.
- Attalla, M. I., Vassallo, A. M., Tattam, B. N., & Hanna, J. V. (1993). Preparation of hydrofullerenes by hydrogen radical-induced hydrogenation. *Journal of Physical Chemistry*, 97(24), 6329–6331.
- Austin, S. J., Fowler, P. W., Orlandi, G., Manolopoulos, D. E., & Zerbetto, F. (1994). Relative stabilities of C_{76} isomers – A numerical test of the fullerene isolated-pentagon rule. *Chemical Physics Letters*, 226(1–2), 219–225.
- Austin, S. J., Fowler, P. W., Manolopoulos, D. E., Orlandi, G., & Zerbetto, F. (1995). Structural motifs and the stability of fullerenes. *Journal of Physical Chemistry*, 99(20), 8076–8081.
- Avdoshenko, S. M., Goryunkov, A. A., Ioffe, I. N., Ignat'eva, D. V., Sidorov, L. N., Pattison, P., Kemnitz, E., & Troyanov, S. I. (2006). Preparation, crystallographic characterization and theoretical study of $\text{C}_{70}(\text{CF}_3)_{16}$ and $\text{C}_{70}(\text{CF}_3)_{18}$. *Chemical Communications*, 23, 2463–2465.

- Avent, A. G., Benito, A. M., Birkett, P. R., Darwish, A. D., Hitchcock, P. B., Kroto, H. W., Locke, I. W., Meidine, M. F., O'Donovan, B. F., Prassides, K., Taylor, R., Walton, D. R. M., & van Wijnkoop, M. (1997). The structure of fullerene compounds. *Journal of Molecular Structure*, *437*, 1–9.
- Avent, A. G., Clare, B. W., Hitchcock, P. B., Kepert, D. L., & Taylor, R. (2002). C₆₀F₃₆: There is a third isomer and it has C₁ symmetry. *Chemical Communications*, *2006*(20), 2370–2371.
- Bader, R. F. W. (1990). *Atoms in molecules – A quantum theory*. Oxford: Oxford University Press.
- Balasubramanian, K. (1991). Enumeration of isomers of polysubstituted C₆₀ and application to NMR. *Chemical Physics Letters*, *182*(3–4), 257–262.
- Beavers, C. M., Zuo, T. M., Duchamp, J. C., Harich, K., Dorn, H. C., Olmstead, M. M., & Balch, A. L. (2006). Tb₃N@C₈₄: An improbable, egg-shaped endohedral fullerene that violates the isolated pentagon rule. *Journal of the American Chemical Society*, *128*(35), 11352–11353.
- Beavers, C. M., Chaur, M. N., Olmstead, M. M., Echegoyen, L., & Balch, A. L. (2009). Large metal ions in a relatively small fullerene cage: The structure of Gd₃N@C₂(22010)-C₇₈ departs from the isolated pentagon rule. *Journal of the American Chemical Society*, *131*(32), 11519–11524.
- Birkett, P. R., Hitchcock, P. B., Kroto, H. W., Taylor, R., & Walton, D. R. M. (1992). Preparation and characterization of C₆₀Br₆ and C₆₀Br₈. *Nature*, *357*(6378), 479–481.
- Birkett, P. R., Avent, A. G., Darwish, A. D., Kroto, H. W., Taylor, R., & Walton, D. R. M. (1993). Preparation and ¹³C NMR spectroscopic characterization of C₆₀Cl₆. *Journal of the Chemical Society, Chemical Communications*, *1993*(15), 1230–1232.
- Birkett, P. R., Avent, A. G., Darwish, A. D., Kroto, H. W., Taylor, R., & Walton, D. R. M. (1995). Formation and characterization of C₇₀Cl₁₀. *Journal of the Chemical Society, Chemical Communications*, *1995*(6), 683–684.
- Boltalina, O. V., Street, J. M., & Taylor, R. (1998). C₆₀F₃₆ consists of two isomers having T and C₃ symmetry. *Journal of the Chemical Society, Perkin Transactions*, *2*(3), 649–653.
- Boltalina, O. V., Ioffe, I. N., Sidorov, L. N., Seifert, G., & Vietze, K. (2000a). Ionization energy of fullerenes. *Journal of the American Chemical Society*, *122*(40), 9745–9749.
- Boltalina, O. V., Lukonin, A. Y., Street, J. M., & Taylor, R. (2000b). C₆₀F₂ exists! *Chemical Communications*, *2000*(17), 1601–1602.
- Boltalina, O. V., Markov, V. Y., Troshin, P. A., Darwish, A. D., Street, J. M., & Taylor, R. (2001). C₆₀F₂₀: “Saturnene”, an extraordinary squashed fullerene. *Angewandte Chemie-International Edition*, *40*(4), 787–789.
- Boltalina, O. V., Darwish, A. D., Street, J. M., Taylor, R., & Wei, X. W. (2002). Isolation and characterisation of C₆₀F₄, C₆₀F₆, C₆₀F₈, C₆₀F₇CF₃ and C₆₀F₂O, the smallest oxahomofullerene; the mechanism of fluorine addition to fullerenes. *Journal of the Chemical Society-Perkin Transactions*, *2*(2), 251–256.
- Boltalina, O. V., Popov, A. A., & Strauss, S. H. (2009). Physicochemical properties and the unusual structure of fullerenes: Single-crystal X-ray structures of fullerenes and their derivatives. In H. Dodziuk (Ed.), *Strained hydrocarbons: Beyond the Van't Hoff and Le Bel hypothesis* (pp. 225–238). Weinheim: Wiley-VCH.
- Boltalina, O. V., Popov, A. A., Kuvychko, I. V., Shustova, N. B., & Strauss, S. H. (2015). Perfluoroalkylfullerenes. *Chemical Reviews*, *115*(2), 1051–1105.
- Campanera, J. M., Bo, C., Olmstead, M. M., Balch, A. L., & Poblet, J. M. (2002). Bonding within the endohedral fullerenes Sc₃N@C₇₈ and Sc₃N@C₈₀ as determined by density functional calculations and reexamination of the crystal structure of Sc₃N@C₇₈-Co(OEP)-1.5(C₆H₆)-0.3(CHCl₃). *Journal of Physical Chemistry A*, *106*(51), 12356–12364.
- Campanera, J. M., Bo, C., & Poblet, J. M. (2005). General rule for the stabilization of fullerene cages encapsulating trimetallic nitride templates. *Angewandte Chemie-International Edition*, *44*(44), 7230–7233.
- Campbell, E. E. B., Fowler, P. W., Mitchell, D., & Zerbetto, F. (1996). Increasing cost of pentagon adjacency for larger fullerenes. *Chemical Physics Letters*, *250*(5–6), 544–548.
- Cao, B. P., Wakahara, T., Tsuchiya, T., Kondo, M., Maeda, Y., Rahman, G. M. A., Akasaka, T., Kobayashi, K., Nagase, S., & Yamamoto, K. (2004). Isolation, characterization, and theoretical study of La₂@C₇₈. *Journal of the American Chemical Society*, *126*(30), 9164–9165.

- Cao, B., Nikawa, H., Nakahodo, T., Tsuchiya, T., Maeda, Y., Akasaka, T., Sawa, H., Slanina, Z., Mizorogi, N., & Nagase, S. (2008). Addition of adamantylidene to $\text{La}_2@C_{78}$: Isolation and single-crystal X-ray structural determination of the monoadducts. *Journal of the American Chemical Society*, 130, 983–989.
- Chai, Y., Guo, T., Jin, C. M., Hauffer, R. E., Chibante, L. P. F., Fure, J., Wang, L. H., Alford, J. M., & Smalley, R. E. (1991). Fullerenes with metals inside. *Journal of Physical Chemistry*, 95(20), 7564–7568.
- Chaur, M. N., Melin, F., Elliott, B., Kumbhar, A., Athans, A. J., & Echegoyen, L. (2008). New $M_3N@C_{2n}$ endohedral metallofullerene families ($M = \text{Nd, Pr, Ce}$; $n = 40\text{--}53$): Expanding the preferential templating of the C_{88} cage and approaching the C_{96} Cage. *Chemistry – A European Journal*, 14(15), 4594–4599.
- Chen, Z. F., & Thiel, W. (2003). Performance of semiempirical methods in fullerene chemistry: Relative energies and nucleus-independent chemical shifts. *Chemical Physics Letters*, 367(1–2), 15–25.
- Chen, Z. F., Cioslowski, J., Rao, N., Moncrieff, D., Buhl, M., Hirsch, A., & Thiel, W. (2001). Endohedral chemical shifts in higher fullerenes with 72–86 carbon atoms. *Theoretical Chemistry Accounts*, 106(5), 364–368.
- Chen, D. L., Tian, W. Q., Feng, J. K., & Sun, C. C. (2008). C_{68} fullerene isomers, anions, and their metallofullerenes: Charge-stabilizing different isomers. *ChemPhysChem*, 9, 454–461.
- Chen, N., Beavers, C. M., Mulet-Gas, M., Rodriguez-Forteza, A., Munoz, E. J., Li, Y.-Y., Olmstead, M. M., Balch, A. L., Poblet, J. M., & Echegoyen, L. (2012). $\text{Sc}_2\text{S}@C_s(10528)\text{-}C_{72}$: A dimetallic sulfide endohedral fullerene with a non-IPR cage. *Journal of the American Chemical Society*, 134(18), 7851–7860.
- Chen, N., Mulet-Gas, M., Li, Y.-Y., Stene, R. E., Atherton, C. W., Rodriguez-Forteza, A., Poblet, J. M., & Echegoyen, L. (2013). $\text{Sc}_2\text{S}@C_2(7892)\text{-}C_{70}$: A metallic sulfide cluster inside a non-IPR C_{70} cage. *Chemical Science*, 4(1), 180–186.
- Cioslowski, J., Rao, N., & Moncrieff, D. (2000). Standard enthalpies of formation of fullerenes and their dependence on structural motifs. *Journal of the American Chemical Society*, 122(34), 8265–8270.
- Clare, B. W., & Kepert, D. L. (1993). Structures and stabilities of hydrofullerenes $C_{60}H_n$. *Theochem-Journal of Molecular Structure*, 281, 45–52.
- Clare, B. W., & Kepert, D. L. (1994a). An analysis of the 63 Possible isomers of $C_{60}H_{36}$ containing a threefold axis – A new structure for $C_{60}H_{20}$. *Theochem-Journal of Molecular Structure*, 315, 71–83.
- Clare, B. W., & Kepert, D. L. (1994b). Structures and stabilities of hydrofullerenes – Completion of a tetrahedral fused quadruple crown structure and a double crown structure at $C_{60}H_{36}$. *Theochem-Journal of Molecular Structure*, 304, 181–189.
- Clare, B. W., & Kepert, D. L. (1994c). Structures and stabilities of hydrofullerenes. Completion of crown structures at $C_{60}H_{18}$ and $C_{60}H_{24}$. *Theochem-Journal of Molecular Structure*, 303, 1–9.
- Clare, B. W., & Kepert, D. L. (1995a). Stereochemical patterns in bromofullerenes $C_{60}Br_2$ to $C_{60}Br_{12}$. *Theochem-Journal of Molecular Structure*, 340, 125–142.
- Clare, B. W., & Kepert, D. L. (1995b). Stereochemical patterns in bromofullerenes, $C_{60}Br_{12}$ to $C_{60}Br_{24}$. *Theochem-Journal of Molecular Structure*, 358, 79–94.
- Clare, B. W., & Kepert, D. L. (1996). Fullerene hydrides based on skew pentagonal pyramidal arrangements of hydrogen atoms. *Theochem-Journal of Molecular Structure*, 363, 179–190.
- Clare, B. W., & Kepert, D. L. (1997). An analysis of the 94 possible isomers of $C_{60}F_{48}$ containing a three-fold axis. *Journal of Molecular Structure-Theochem*, 389, 97–103.
- Clare, B. W., & Kepert, D. L. (1999a). The structures of $C_{60}F_{36}$ and new possible structures for $C_{60}H_{36}$. *Journal of Molecular Structure-Theochem*, 466, 177–186.
- Clare, B. W., & Kepert, D. L. (1999b). The structures of $C_{70}X_n$, $X = \text{H, F, Br}$, C_6H_5 and $n = 2\text{--}12$. *Theochem-Journal of Molecular Structure*, 491, 249–264.
- Clare, B. W., & Kepert, D. L. (2001). Structures and stabilities of adducts of carbenes and fullerenes, $C_{60}(\text{CR}_2)_n$, $R = \text{H, CH}_3$, C_4H_9 and $n = 1\text{--}6$. *Journal of Molecular Structure-Theochem*, 548, 61–91.

- Clare, B. W., & Kepert, D. L. (2002a). Structures of $C_{60}H_n$ and $C_{60}F_n$, $n = 36-60$. *Journal of Molecular Structure-Theochem*, 589, 209–227.
- Clare, B. W., & Kepert, D. L. (2002b). Structures, stabilities and isomerism in $C_{60}H_{36}$ and $C_{60}F_{36}$. A comparison of the AM1 Hamiltonian and density functional techniques. *Journal of Molecular Structure-Theochem*, 589, 195–207.
- Clare, B. W., & Kepert, D. L. (2003a). Early stages in the addition to C_{60} to form $C_{60}X_n$, $X = H, F, Cl, Br, CH_3, C_4H_9$. *Journal of Molecular Structure-Theochem*, 621(3), 211–231.
- Clare, B. W., & Kepert, D. L. (2003b). Structures, stabilities and isomerism in $C_{60}H_n$, $n = 2-36$. A comparison of the AM1 Hamiltonian and density functional techniques. *Journal of Molecular Structure-Theochem*, 622(3), 185–202.
- Cortés-Guzmán, F., & Bader, R. F. W. (2005). Complementarity of QTAIM and MO theory in the study of bonding in donor–acceptor complexes. *Coordination Chemistry Reviews*, 249(5–6), 633–662.
- Cremer, D., & Kraka, E. (1984). Chemical bonds without bonding electron density – Does the difference electron-density analysis suffice for a description of the chemical bond? *Angewandte Chemie-International Edition in English*, 23(8), 627–628.
- Darwish, A. D., Avent, A. G., Taylor, R., & Walton, D. R. M. (1996). Structural characterisation of $C_{60}H_{18}$; a C_{3v} symmetry crown. *Journal of the Chemical Society-Perkin Transactions*, 2(10), 2051–2054.
- Darzynkiewicz, R. B., & Scuseria, G. E. (1997). Noble gas endohedral complexes of C_{60} buckminsterfullerene. *Journal of Physical Chemistry A*, 101(38), 7141–7144.
- Dennis, T. J. S., Kai, T., Asato, K., Tomiyama, T., & Shinohara, H. (1999). Isolation and characterization by ^{13}C NMR spectroscopy of [84]fullerene minor isomers. *Journal of Physical Chemistry A*, 103(44), 8747–8752.
- Diener, M. D., & Alford, J. M. (1998). Isolation and properties of small-bandgap fullerenes. *Nature*, 393(6686), 668–671.
- Dixon, D. A., Matsuzawa, N., Fukunaga, T., & Tebbe, F. N. (1992). Patterns for addition to C_{60} . *Journal of Physical Chemistry*, 96(15), 6107–6110.
- Dorozhkin, E. I., Ignat'eva, D. V., Tamm, N. B., Goryunkov, A. A., Khavrel, P. A., Ioffe, I. N., Popov, A. A., Kuvychko, I. V., Streletskiy, A. V., Markov, V. Y., Spandl, J., Strauss, S. H., & Boltalina, O. V. (2006a). Synthesis, characterization, and theoretical study of stable isomers of $C_{70}(CF_3)_n$ ($n = 2, 4, 6, 8, 10$). *Chemistry - A European Journal*, 12(14), 3876–3889.
- Dorozhkin, E. I., Ignat'eva, D. V., Tamm, N. B., Vasilyuk, N. V., Goryunkov, A. A., Avdoshenko, S. M., Ioffe, I. N., Sidorov, L. N., Pattison, P., Kemnitz, E., & Troyanov, S. I. (2006b). Structure of 1,4,10,19,25,41- $C_{70}(CF_3)_6$, isomer with unique arrangement of addends. *Journal of Fluorine Chemistry*, 127(10), 1344–1348.
- Dorozhkin, E. I., Goryunkov, A. A., Ioffe, I. N., Avdoshenko, S. M., Markov, V. Y., Tamm, N. B., Ignat'eva, D. V., Sidorov, L. N., & Troyanov, S. I. (2007). Synthesis, structure, and theoretical study of lower trifluoromethyl derivatives of [60]fullerene. *European Journal of Organic Chemistry*, 2007(30), 5082–5094.
- Dunsch, L., & Yang, S. (2007). Metal nitride cluster fullerenes: Their current state and future prospects. *Small*, 3(8), 1298–1320.
- Dunsch, L., Georgi, P., Krause, M., & Wang, C. R. (2003). New clusters in endohedral fullerenes: The metal nitrides. *Synthetic Metals*, 135(1–3), 761–762.
- Dunsch, L., Krause, M., Noack, J., & Georgi, P. (2004). Endohedral nitride cluster fullerenes – Formation and spectroscopic analysis of $L_{3-x}M_xN@C_{2n}$ ($0 \leq x \leq 3$; $n = 39, 40$). *Journal of Physics and Chemistry of Solids*, 65(2–3), 309–315.
- Dunsch, L., Yang, S., Zhang, L., Svitova, A., Oswald, S., & Popov, A. A. (2010). Metal sulfide in a C_{82} fullerene cage: A new form of endohedral clusterfullerenes. *Journal of the American Chemical Society*, 132(15), 5413–5421.
- Ettl, R., Chao, I., Diederich, F., & Whetten, R. L. (1991). Isolation of C_{76} , a chiral (D_2) allotrope of carbon. *Nature*, 353(6340), 149–153.
- Farrugia, L. J., Evans, C., Lentz, D., & Roemer, M. (2009). The QTAIM approach to chemical bonding between transition metals and carbocyclic rings: A combined experimental and

- theoretical study of $(\eta^5\text{-C}_5\text{H}_5)\text{Mn}(\text{CO})_3$, $(\eta^6\text{-C}_6\text{H}_6)\text{Cr}(\text{Co})_3$, and $(E)\text{-}\{(\eta^5\text{-C}_5\text{H}_4)\text{CF}=\text{CF}(\eta^5\text{-C}_5\text{H}_4)\}(\eta^5\text{-C}_5\text{H}_5)_2\text{Fe}_2$. *Journal of the American Chemical Society*, 131(3), 1251–1268.
- Feng, Y., Wang, T., Wu, J., Feng, L., Xiang, J., Ma, Y., Zhang, Z., Jiang, L., Shu, C., & Wang, C. (2013). Structural and electronic studies of metal carbide clusterfullerene $\text{Sc}_2\text{C}_2@C_s\text{-C}_{72}$. *Nanoscale*, 5(15), 6704–6707.
- Fowler, P. W., & Manolopoulos, D. E. (1995). *An atlas of fullerenes*. Oxford: Clarendon Press.
- Fowler, P. W., & Zerbetto, F. (1995). Charging and equilibration of fullerene isomers. *Chemical Physics Letters*, 243(1–2), 36–41.
- Fowler, P. W., Heine, T., Mitchell, D., Orlandi, G., Schmidt, R., Seifert, G., & Zerbetto, F. (1996). Energetics of fullerenes with heptagonal rings. *Journal of the Chemical Society-Faraday Transactions*, 92(12), 2203–2210.
- Fowler, P. W., Rogers, K. M., Somers, K. R., & Troisi, A. (1999). Independent sets and the prediction of addition patterns for higher fullerenes. *Journal of the Chemical Society-Perkin Transactions*, 2(10), 2023–2027.
- Fritz, M. A., Kemnitz, E., & Troyanov, S. I. (2014). Capturing an unstable C_{100} fullerene as chloride, $\text{C}_{100}(1)\text{Cl}_{12}$, with a nanotubular carbon cage. *Chemical Communications*, 50(93), 14577–14580.
- Furche, F., & Ahlrichs, R. (2001). Fullerene C_{80} : Are there still more isomers? *Journal of Chemical Physics*, 114(23), 10362–10367.
- Gakh, A. A., & Tuinman, A. A. (2001). ‘Fluorine dance’ on the fullerene surface. *Tetrahedron Letters*, 42(41), 7137–7139.
- Gakh, A. A., Romanovich, A. Y., & Bax, A. (2003). Thermodynamic rearrangement synthesis and NMR structures of C_1 , C_3 , and T isomers of $\text{C}_{60}\text{H}_{36}$. *Journal of the American Chemical Society*, 125(26), 7902–7906.
- Gan, L. H., & Yuan, R. (2006). Influence of cluster size on the structures and stability of trimetallic nitride fullerenes $\text{M}_3\text{N}@C_{80}$. *ChemPhysChem*, 7(6), 1306–1310.
- García-Borràs, M., Osuna, S., Swart, M., Luis, J. M., & Solà, M. (2013). Maximum aromaticity as a guiding principle for the most suitable hosting cages in endohedral metallofullerenes. *Angewandte Chemie International Edition*, 52(35), 9275–9278.
- Goryunkov, A. A., Kuvychko, I. V., Ioffe, I. N., Dick, D. L., Sidorov, L. N., Strauss, S. H., & Boltalina, O. V. (2003). Isolation of $\text{C}_{60}(\text{CF}_3)_n$ ($n = 2, 4, 6, 8, 10$) with high compositional purity. *Journal of Fluorine Chemistry*, 124(1), 61–64.
- Goryunkov, A. A., Ioffe, I. N., Kuvychko, I. V., Yankova, T. S., Markov, V. Y., Streletskii, A. A., Dick, D. L., Sidorov, L. N., Boltalina, O. V., & Strauss, S. H. (2004a). Trifluoromethylated [60]fullerenes: Synthesis and characterization. *Fullerenes, Nanotubes, and Carbon Nanostructures*, 12(1–2), 181–185.
- Goryunkov, A. A., Markov, V. Y., Ioffe, I. N., Bolskar, R. D., Diener, M. D., Kuvychko, I. V., Strauss, S. H., & Boltalina, O. V. (2004b). $\text{C}_{74}\text{F}_{38}$: An exohedral derivative of a small-bandgap fullerene with D_3 symmetry. *Angewandte Chemie-International Edition*, 43(8), 997–1000.
- Goryunkov, A. A., Dorozhkin, E. I., Ignat’eva, D. V., Sidorov, L. N., Kemnitz, E., Sheldrick, G., & Troyanov, S. I. (2005). Crystal and molecular structures of $\text{C}_{70}(\text{CF}_3)_8\text{-PhMe}$. *Mendeleev Communications*, 15(6), 225–227.
- Goryunkov, A. A., Ignat’eva, D. V., Tamm, N. B., Moiseeva, N. N., Loffe, I. N., Avdoshenko, S. M., Markov, V. Y., Sidorov, L. N., Kemnitz, E., & Troyanov, S. I. (2006a). Preparation, crystallographic characterization, and theoretical study of $\text{C}_{70}(\text{CF}_3)_{14}$. *European Journal of Organic Chemistry*, 2006, 2508–2512.
- Goryunkov, A. A., Kareev, I. E., Ioffe, I. N., Popov, A. A., Kuvychko, I. V., Markov, V. Y., Goldt, I. V., Pimenova, A. S., Serov, M. G., Avdoshenko, S. M., Khavrel, P. A., Sidorov, L. N., Lebedkin, S. F., Mazej, Z., Zemva, B., Strauss, S. H., & Boltalina, O. V. (2006b). Reaction of C_{60} with KMnF_4 – Isolation and characterization of a new isomer of C_{60}F_8 and re-evaluation of the structures of $\text{C}_{60}\text{F}_7(\text{CF}_3)$ and the known isomer of C_{60}F_8 . *Journal of Fluorine Chemistry*, 127(10), 1423–1435.
- Goryunkov, A. A., Dorozhkin, E. I., Tamm, N. B., Ignat’eva, D. V., Avdoshenko, S. M., Sidorov, L. N., & Troyanov, S. I. (2007). Synthesis and molecular structure of 1,6,11,16,18,24,27,36- $\text{C}_{60}(\text{CF}_3)_8$. *Mendeleev Communications*, 17(2), 110–112.

- Haddon, R. C. (1993). Chemistry of the fullerenes – The manifestation of strain in a class of continuous aromatic-molecules. *Science*, 261(5128), 1545–1550.
- Hao, Y., Feng, L., Xu, W., Gu, Z., Hu, Z., Shi, Z., Slanina, Z., & Uhlík, F. (2015). Sm@C_{2v}(19,138)-C₇₆: A non-IPR cage stabilized by a divalent metal ion. *Inorganic Chemistry*, 54(9), 4243–4248.
- Haser, M., Almlof, J., & Scuseria, G. E. (1991). The equilibrium geometry of C₆₀ as predicted by 2nd-Order (MP2) perturbation-theory. *Chemical Physics Letters*, 181(6), 497–500.
- Hauffler, R. E., Conceicao, J., Chibante, L. P. F., Chai, Y., Byrne, N. E., Flanagan, S., Haley, M. M., O'Brien, S. C., Pan, C., Xiao, Z., Billups, W. E., Ciufolini, M. A., Hauge, R. H., Margrave, J. L., Wilson, L. J., Curl, R. F., & Smalley, R. E. (1990). Efficient production of C₆₀ (Buckminsterfullerene), C₆₀H₃₆, and the solvated buckide ion. *Journal of Physical Chemistry*, 94(24), 8634–8636.
- Heath, J. R., O'Brien, S. C., Zhang, Q., Liu, Y., Curl, R. F., Tittel, F. K., & Smalley, R. E. (1985). Lanthanum complexes of spheroidal carbon shells. *Journal of the American Chemical Society*, 107(25), 7779–7780.
- Henderson, C. C., & Cahill, P. A. (1993). C₆₀H₂ - synthesis of the simplest C₆₀ hydrocarbon derivative. *Science*, 259(5103), 1885–1887.
- Hennrich, F. H., Michel, R. H., Fischer, A., Richard-Schneider, S., Gilb, S., Kappes, M. M., Fuchs, D., Burk, M., Kobayashi, K., & Nagase, S. (1996). Isolation and characterization of C₈₀. *Angewandte Chemie-International Edition in English*, 35(15), 1732–1734.
- Hirsch, A., & Brettreich, M. (2005). *Fullerenes. Chemistry and reactions*. Weinheim: Wiley.
- Hirsch, A., & Vostrowsky, O. (2001). C₆₀ hexakisadducts with an octahedral addition pattern – A new structure motif in organic chemistry. *European Journal of Organic Chemistry*, 2001, 829–848.
- Hitchcock, P. B., & Taylor, R. (2002). Single crystal X-ray structure of tetrahedral C₆₀F₃₆: The most aromatic and distorted fullerene. *Chemical Communications*, 2002(18), 2078–2079.
- Hitchcock, P. B., Avent, A. G., Martsinovich, N., Troshin, P. A., & Taylor, R. (2005). C₂-C₇₀F₃₈ is aromatic, contains three planar hexagons, and has equatorial addends. *Chemical Communications*, 2005(1), 75–77.
- Ignat'eva, D. V., Goryunkov, A. A., Tamm, N. B., Ioffe, I. N., Avdoshenko, S. M., Sidorov, L. N., Dimitrov, A., Kemnitz, E., & Troyanov, S. I. (2006). Preparation, crystallographic characterization and theoretical study of two isomers of C₇₀(CF₃)₁₂. *Chemical Communications*, 2006(16), 1778–1780.
- Iiduka, Y., Wakahara, T., Nakahodo, T., Tsuchiya, T., Sakuraba, A., Maeda, Y., Akasaka, T., Yoza, K., Horn, E., Kato, T., Liu, M. T. H., Mizorogi, N., Kobayashi, K., & Nagase, S. (2005). Structural determination of metallofullerene Sc₃C₈₂ revisited: A surprising finding. *Journal of the American Chemical Society*, 127(36), 12500–12501.
- Iiduka, Y., Wakahara, T., Nakajima, K., Tsuchiya, T., Nakahodo, T., Maeda, Y., Akasaka, T., Mizorogi, N., & Nagase, S. (2006). ¹³C NMR spectroscopic study of scandium dimetallofullerene, Sc₂@C₈₄ vs. Sc₂C₂@C₈₂. *Chemical Communications*, 2006(19), 2057–2059.
- Inoue, T., Tomiyama, T., Sugai, T., Okazaki, T., Suematsu, T., Fujii, N., Utsumi, H., Nojima, K., & Shinohara, H. (2004). Trapping a C₂ radical in endohedral metallofullerenes: Synthesis and structures of (Y₂C₂)@C₈₂ (Isomers I, II, and III). *Journal of Physical Chemistry B*, 108(23), 7573–7579.
- Irle, S., Zheng, G. S., Elstner, M., & Morokuma, K. (2003). From C₂ molecules to self-assembled fullerenes in quantum chemical molecular dynamics. *Nano Letters*, 3(12), 1657–1664.
- Irle, S., Zheng, G. S., Wang, Z., & Morokuma, K. (2006). The C₆₀ formation puzzle “solved”: QM/MD simulations reveal the shrinking hot giant road of the dynamic fullerene self-assembly mechanism. *Journal of Physical Chemistry B*, 110(30), 14531–14545.
- Karataev, V. I. (1998). Anomalous properties of C₇₄ fullerene. *Molecular Crystals and Liquid Crystals Science and Technology Section C-Molecular Materials*, 11(1–2), 57–58.
- Kareev, I. E., Kuvychko, I. V., Popov, A. A., Lebedkin, S. F., Miller, S. M., Anderson, O. P., Strauss, S. H., & Boltalina, O. V. (2005). High-temperature synthesis of the surprisingly stable C₁-C₇₀(CF₃)₁₀ isomer with a *para*⁷-*meta*-*para* ribbon of nine C₆(CF₃)₂ edge-sharing hexagons. *Angewandte Chemie-International Edition*, 44(48), 7984–7987.

- Kareev, I. E., Shustova, N. B., Kuvychko, I. V., Lebedkin, S. F., Miller, S. M., Anderson, O. P., Popov, A. A., Strauss, S. H., & Boltalina, O. V. (2006). Thermally stable perfluoroalkylfullerenes with the skew-pentagonal-pyramid pattern: $C_{60}(C_2F_5)_4O$, $C_{60}(CF_3)_4O$, and $C_{60}(CF_3)_6$. *Journal of the American Chemical Society*, 128(37), 12268–12280.
- Kareev, I. E., Shustova, N. B., Peryshkov, D. V., Lebedkin, S. F., Miller, S. M., Anderson, O. P., Popov, A. A., Boltalina, O. V., & Strauss, S. H. (2007). X-ray structure and DFT study of C_1 - $C_{60}(CF_3)_{12}$. A high-energy, kinetically-stable isomer prepared at 500 degrees C. *Chemical Communications*, 2007(16), 1650–1652.
- Kareev, I. E., Kuvychko, I. V., Shustova, N. B., Lebedkin, S. F., Bubnov, V. P., Anderson, O. P., Popov, A. A., Boltalina, O. V., & Strauss, S. H. (2008a). C_1 -(C_{84} - $C_2(11)$)(CF_3)₁₂: Trifluoromethylation yields structural proof of a minor C_{84} cage and reveals a principle of higher fullerene reactivity. *Angewandte Chemie-International Edition in English*, 47, 6204–6207.
- Kareev, I. E., Popov, A. A., Kuvychko, I. V., Shustova, N. B., Lebedkin, S. F., Bubnov, V. P., Anderson, O. P., Seppelt, K., Strauss, S. H., & Boltalina, O. V. (2008b). Synthesis and X-ray or NMR/DFT structure elucidation of Twenty-One new trifluoromethyl derivatives of soluble cage isomers of C_{76} , C_{78} , C_{84} , and C_{90} . *Journal of the American Chemical Society*, 130, 13471–13489.
- Kato, H., Taninaka, A., Sugai, T., & Shinohara, H. (2003). Structure of a missing-caged metallofullerene: $La_2@C_{72}$. *Journal of the American Chemical Society*, 125(26), 7782–7783.
- Kemnitz, E., & Troyanov, S. I. (2009). Connectivity patterns of two C_{90} isomers provided by the structure elucidation of $C_{90}Cl_{32}$. *Angewandte Chemie International Edition*, 48(14), 2584–2587.
- Keperth, D. L., & Clare, B. W. (1996). Stereochemical patterns formed by addition to fullerene C_{60} . *Coordination Chemistry Reviews*, 155, 1–33.
- Kessler, B., Bringer, A., Cramm, S., Schlebusch, C., Eberhardt, W., Suzuki, S., Achiba, Y., Esch, F., Barnaba, M., & Cocco, D. (1997). Evidence for incomplete charge transfer and La-derived states in the valence bands of endohedrally doped $La@C_{82}$. *Physical Review Letters*, 79(12), 2289–2292.
- Kikuchi, K., Nakahara, N., Wakabayashi, T., Suzuki, S., Shiromaru, H., Miyake, Y., Saito, K., Ikemoto, I., Kainosho, M., & Achiba, Y. (1992). NMR characterization of isomers of C_{78} , C_{82} and C_{84} fullerenes. *Nature*, 357(6374), 142–145.
- Kobayashi, K., & Nagase, S. (1997). Structures of the $Ca@C_{82}$ isomers: A theoretical prediction. *Chemical Physics Letters*, 274(1–3), 226–230.
- Kobayashi, K., & Nagase, S. (1998). Structures and electronic states of $M@C_{82}$ ($M = Sc, Y, La$ and lanthanides). *Chemical Physics Letters*, 282(3–4), 325–329.
- Kobayashi, K., & Nagase, S. (1999). Bonding features in endohedral metallofullerenes. Topological analysis of the electron density distribution. *Chemical Physics Letters*, 302(3–4), 312–316.
- Kobayashi, K., & Nagase, S. (2002a). A stable unconventional structure of $Sc_2@C_{66}$ found by density functional calculations. *Chemical Physics Letters*, 362(5–6), 373–379.
- Kobayashi, K., & Nagase, S. (2002b). Structures and electronic properties of endohedral metallofullerenes; theory and experiment. In T. Akasaka & S. Nagase (Eds.), *Endofullerenes: A new family of carbon clusters* (pp. 99–119). Dordrecht: Kluwer.
- Kobayashi, K., Nagase, S., & Akasaka, T. (1995). A theoretical study of C_{80} and $La_2@C_{80}$. *Chemical Physics Letters*, 245(2–3), 230–236.
- Kobayashi, K., Nagase, S., Yoshida, M., & Osawa, E. (1997). Endohedral metallofullerenes. Are the isolated pentagon rule and fullerene structures always satisfied? *Journal of the American Chemical Society*, 119(51), 12693–12694.
- Kobayashi, K., Sano, Y., & Nagase, S. (2001). Theoretical study of endohedral metallofullerenes: $Sc_{3-n}La_nN@C_{80}$ ($n = 0–3$). *Journal of Computational Chemistry*, 22(13), 1353–1358.
- Komatsu, K., Murata, M., & Murata, Y. (2005). Encapsulation of molecular hydrogen in fullerene C_{60} by organic synthesis. *Science*, 307(5707), 238–240.
- Kovalenko, V. I., & Khamatgalimov, A. R. (2003). Open-shell fullerene C_{74} : Phenalenyl-radical substructures. *Chemical Physics Letters*, 377(3–4), 263–268.
- Krapp, A., & Frenking, G. (2007). Is this a chemical bond? A theoretical study of $Ng_2@C_{60}$ ($Ng = He, Ne, Ar, Kr, Xe$). *Chemistry - A European Journal*, 13(29), 8256–8270.

- Krätschmer, W., Lamb, L. D., Fostiropoulos, K., & Huffman, D. R. (1990). Solid C₆₀ – A new form of carbon. *Nature*, 347(6291), 354–358.
- Krause, M., & Dunsch, L. (2005). Gadolinium nitride Gd₃N in carbon cages: The influence of cluster size and bond strength. *Angewandte Chemie-International Edition*, 44(10), 1557–1560.
- Krause, M., Popov, A., & Dunsch, L. (2006). Vibrational structure of endohedral fullerene Sc₃N@C₇₈(D_{3h}⁺): Evidence for a strong coupling between the Sc₃N cluster and C₇₈ cage. *ChemPhysChem*, 7(8), 1734–1740.
- Krause, M., Ziegls, F., Popov, A. A., & Dunsch, L. (2007). Entrapped bonded hydrogen in a fullerene: The five-atom cluster Sc₃CH in C₈₀. *Chemphyschem*, 8(4), 537–540.
- Kroto, H. W. (1987). The Stability of the fullerenes C₂₄, C₂₈, C₃₂, C₃₆, C₅₀, C₆₀ and C₇₀. *Nature*, 329(6139), 529–531.
- Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F., & Smalley, R. E. (1985). C₆₀ – Buckminsterfullerene. *Nature*, 318(6042), 162–163.
- Lamparth, I., Maichlemosmer, C., & Hirsch, A. (1995). Reversible template-directed activation of equatorial double-bonds of the fullerene framework – Regioselective direct synthesis, crystal-structure, and aromatic properties of T_h-C₆₆(COOEt)₁₂. *Angewandte Chemie-International Edition in English*, 34(15), 1607–1609.
- Liu, D., Hagelberg, F., & Park, S. S. (2006). Charge transfer and electron backdonation in metallofullerenes encapsulating NSc₃. *Chemical Physics*, 330(3), 380–386.
- Lu, J., Zhang, X. W., Zhao, X. G., Nagase, S., & Kobayashi, K. (2000). Strong metal-cage hybridization in endohedral La@C₈₂, Y@C₈₂ and Sc@C₈₂. *Chemical Physics Letters*, 332(3–4), 219–224.
- Lu, X., Nikawa, H., Nakahodo, T., Tsuchiya, T., Ishitsuka, M. O., Maeda, Y., Akasaka, T., Toki, M., Sawa, H., Slanina, Z., Mizorogi, N., & Nagase, S. (2008). Chemical understanding of a Non-IPR metallofullerene: Stabilization of encaged metals on fused-pentagon bonds in La₂@C₇₂. *Journal of the American Chemical Society*, 130, 9129–9136.
- Macchi, P., & Sironi, A. (2003). Chemical bonding in transition metal carbonyl clusters: Complementary analysis of theoretical and experimental electron densities. *Coordination Chemistry Reviews*, 238–239, 383–412.
- Manolopoulos, D. E., May, J. C., & Down, S. E. (1991). Theoretical-studies of the fullerenes – C₃₄ to C₇₀. *Chemical Physics Letters*, 181(2–3), 105–111.
- Matsuzawa, N., Dixon, D. A., & Fukunaga, T. (1992). Semiempirical calculations of dihydrogenated buckminsterfullerenes, C₆₀H₂. *Journal of Physical Chemistry*, 96(19), 7594–7604.
- Matta, C. F., & Boyd, R. J. (Eds.). (2007). *The quantum theory of atoms in molecules. From solid state to DNA and drug design*. Weinheim: Wiley.
- Mercado, B. Q., Beavers, C. M., Olmstead, M. M., Chaur, M. N., Walker, K., Holloway, B. C., Echegoyen, L., & Balch, A. L. (2008). Is the isolated pentagon rule merely a suggestion for endohedral fullerenes? The structure of a second egg-shaped endohedral fullerene—Gd₃N@C_s(39663)-C₈₂. *Journal of the American Chemical Society*, 130(25), 7854–7855.
- Miyake, Y., Minami, T., Kikuchi, K., Kainosho, M., & Achiba, Y. (2000). Trends in structure and growth of higher fullerenes isomer structure of C₈₆ and C₈₈. *Molecular Crystals and Liquid Crystals*, 340, 553–558.
- Muthukumar, K., & Larsson, J. A. (2008). A density functional study of Ce@C₈₂: Explanation of the Ce preferential bonding site. *Journal of Physical Chemistry A*, 112, 1071–1075.
- Mutig, T., Ioffe, I. N., Kemnitz, E., & Troyanov, S. I. (2008). Crystal and molecular structures of C₂-C₇₀(CF₃)₈·1.5PhMe. *Mendeleev Communications*, 18, 73–75.
- Nagase, S., Kobayashi, K., & Akasaka, T. (1997). Recent progress in endohedral dimetallofullerenes. *Theochem-Journal of Molecular Structure*, 398, 221–227.
- Nagase, S., Kobayashi, K., & Akasaka, T. (1999). Unconventional cage structures of endohedral metallofullerenes. *Journal of Molecular Structure-Theochem*, 462, 97–104.
- Neretin, I. S., Lyssenko, K. A., Antipin, M. Y., Slovokhotov, Y. L., Boltalina, O. V., Troshin, P. A., Lukonin, A. Y., Sidorov, L. N., & Taylor, R. (2000). C₆₀F₁₈, a flattened fullerene: Alias a hexa-substituted benzene. *Angewandte Chemie-International Edition*, 39(18), 3273–3276.

- Nishibori, E., Takata, M., Sakata, M., Taninaka, A., & Shinohara, H. (2001). Pentagonal-dodecahedral La_2 charge density in $[\text{80-}I_h]\text{fullerene}$: $\text{La}_2@C_{80}$. *Angewandte Chemie-International Edition*, 40(16), 2998–2999.
- Nishibori, E., Ishihara, M., Takata, M., Sakata, M., Ito, Y., Inoue, T., & Shinohara, H. (2006a). Bent $(\text{metal})_2C_2$ clusters encapsulated in $(\text{Sc}_2C_2)@C_{82}(\text{III})$ and $(Y_2C_2)@C_{82}(\text{III})$ metallofullerenes. *Chemical Physics Letters*, 433(1–3), 120–124.
- Nishibori, E., Narioka, S., Takata, M., Sakata, M., Inoue, T., & Shinohara, H. (2006b). A C_2 molecule entrapped in the pentagonal-dodecahedral Y_2 cage in $Y_2C_2@C_{82}(\text{III})$. *ChemPhysChem*, 7(2), 345–348.
- Olmstead, M. H., de Bettencourt-Dias, A., Duchamp, J. C., Stevenson, S., Marciu, D., Dorn, H. C., & Balch, A. L. (2001). Isolation and structural characterization of the endohedral fullerene $\text{Sc}_3N@C_{78}$. *Angewandte Chemie-International Edition*, 40(7), 1223–1225.
- Olmstead, M. M., Lee, H. M., Duchamp, J. C., Stevenson, S., Marciu, D., Dorn, H. C., & Balch, A. L. (2003). $\text{Sc}_3N@C_{68}$: Folded pentalene coordination in an endohedral fullerene that does not obey the isolated pentagon rule. *Angewandte Chemie-International Edition*, 42(8), 900–903.
- Omelyanyuk, N. A., Goryunkov, A. A., Tamm, N. B., Avdoshenko, S. M., Ioffe, I. N., Sidorov, L. N., Kemnitz, E., & Troyanov, S. I. (2007). New trifluoromethylated derivatives of $[\text{60}]\text{fullerene}$, $C_{60}(\text{CF}_3)_n$ with $n = 12$ and 14 . *Chemical Communications*, 2007(45), 4794–4796.
- Popov, A. A. (2009). Metal-cage bonding, molecular structures and vibrational spectra of endohedral fullerenes: Bridging experiment and theory. *Journal of Computational and Theoretical Nanoscience*, 6(2), 292–317.
- Popov, A. A., & Dunsch, L. (2007). Structure, stability, and cluster-cage interactions in nitride clusterfullerenes $M_3N@C_{2n}$ ($M = \text{Sc}, Y$; $2n = 68–98$): A density functional theory study. *Journal of the American Chemical Society*, 129(38), 11835–11849.
- Popov, A. A., & Dunsch, L. (2008). Hindered cluster rotation and ^{45}Sc hyperfine splitting constant in distonoid anion radical $\text{Sc}_3N@C_{80}$, and spatial spin charge separation as a general principle for anions of endohedral fullerenes with metal-localized lowest unoccupied molecular orbitals. *Journal of the American Chemical Society*, 130(52), 17726–17742.
- Popov, A. A., & Dunsch, L. (2009). The bonding situation in endohedral metallofullerenes as studied by quantum theory of atoms in molecules (QTAIM). *Chemistry - A European Journal*, 15(38), 9707–9729.
- Popov, A. A., Senyavin, V. M., Granovsky, A. A., & Lobach, A. S. (2004). Vibrational spectra and molecular structure of the hydrofullerenes $C_{60}H_{18}$, $C_{60}D_{18}$ and $C_{60}H_{36}$ as studied by IR and Raman spectroscopy and first-principle calculations. In T. N. Veziroglu (Ed.), *Hydrogen materials science and chemistry of carbon nanomaterials* (NATO science series II: Mathematics, physics and chemistry, Vol. 172, pp. 347–356). Dordrecht: Kluwer.
- Popov, A. A., Senyavin, V. M., Boltalina, O. V., Seppelt, K., Spandl, J., Feigler, C. S., & Compton, R. N. (2006). Infrared, Raman, and DFT vibrational spectroscopic studies of $C_{60}F_{36}$ and $C_{60}F_{48}$. *Journal of Physical Chemistry A*, 110(28), 8645–8652.
- Popov, A. A., Kareev, I. E., Shustova, N. B., Lebedkin, S. F., Strauss, S. H., Boltalina, O. V., & Dunsch, L. (2007a). Synthesis, spectroscopic and electrochemical characterization, and DFT study of seventeen $C_{70}(\text{CF}_3)_n$ derivatives ($n = 2, 4, 6, 8, 10, 12$). *Chemistry - A European Journal*, 14(1), 107–121.
- Popov, A. A., Kareev, I. E., Shustova, N. B., Stukalin, E. B., Lebedkin, S. F., Seppelt, K., Strauss, S. H., Boltalina, O. V., & Dunsch, L. (2007b). Electrochemical, spectroscopic, and DFT study of $C_{60}(\text{CF}_3)_n$ frontier orbitals ($n = 2–18$): The link between double bonds in pentagons and reduction potentials. *Journal of the American Chemical Society*, 129(37), 11551–11568.
- Popov, A. A., Krause, M., Yang, S. F., Wong, J., & Dunsch, L. (2007c). C_{78} cage isomerism defined by trimetallic nitride cluster size: A computational and vibrational spectroscopic study. *Journal of Physical Chemistry B*, 111(13), 3363–3369.
- Popov, A. A., Yang, S., & Dunsch, L. (2013). Endohedral Fullerenes. *Chemical Reviews*, 113(8), 5989–6113.

- Raghavachari, K. (1992). Ground state of C_{84} : Two almost isoenergetic isomers. *Chemical Physics Letters*, 190(5), 397–400.
- Reimers, J. R., Cai, Z.-L., Bilic, A., & Hush, N. S. (2003). The appropriateness of density-functional theory for the calculation of molecular electronics properties. *Annals of the New York Academy of Sciences*, 1006, 235–251.
- Rodriguez-Forste, A., Alegret, N., Balch, A. L., & Poblet, J. M. (2010). The maximum pentagon separation rule provides guideline for the structures of endohedral metallofullerenes. *Nature Chemistry*, 2(11), 955–961.
- Rogers, K. M., & Fowler, P. W. (1999). A model for pathways of radical addition to fullerenes. *Chemical Communications*, 1999(23), 2357–2358.
- Ruchardt, C., Gerst, M., Ebenhoch, J., Beckhaus, H. D., Campbell, E. E. B., Tellgmann, R., Schwarz, H., Weiske, T., & Pitter, S. (1993). Bimolecular radical formation through H-transfer 3. Transfer hydrogenation and deuteration of buckminsterfullerene C_{60} by 9,10-dihydroanthracene and 9,9', 10,10'-[D₄]dihydroanthracene. *Angewandte Chemie-International Edition in English*, 32(4), 584–586.
- Samokhvalova, N. A., Khavrel, P. A., Markov, V. Y., Samokhvalov, P. S., Goryunkov, A. A., Kemnitz, E., Sidorov, L. N., & Troyanov, S. I. (2009). Isolation and structural characterization of the most stable, highly symmetric isomer of $C_{60}(CF_3)_{18}$. *European Journal of Organic Chemistry*, 2009(18), 2935–2938.
- Sandall, J. P. B., & Fowler, P. W. (2003). The energies of some isomers of $C_{60}F_8$: The use of experimental and theoretical considerations to limit candidate structures. *Organic & Biomolecular Chemistry*, 1(6), 1061–1066.
- Sato, Y., Yumura, T., Suenaga, K., Moribe, H., Nishide, D., Ishida, M., Shinohara, H., & Iijima, S. (2006). Direct imaging of intracage structure in titanium-carbide endohedral metallofullerene. *Physical Review B*, 73(19), 193401.
- Saunders, M., Jimenezvazquez, H. A., Cross, R. J., & Poreda, R. J. (1993). Stable compounds of Helium and Neon – He@ C_{60} and Ne@ C_{60} . *Science*, 259(5100), 1428–1430.
- Schmalz, T. G., Seitz, W. A., Klein, D. J., & Hite, G. E. (1988). Elemental carbon cages. *Journal of the American Chemical Society*, 110(4), 1113–1127.
- Shao, N., Gao, Y., Yoo, S., An, W., & Zeng, X. C. (2006). Search for lowest-energy fullerenes: C_{98} to C_{110} . *Journal of Physical Chemistry A*, 110(24), 7672–7676.
- Shao, N., Gao, Y., & Zeng, X. C. (2007). Search for lowest-energy fullerenes 2: C_{38} to C_{80} and C_{112} to C_{120} . *Journal of Physical Chemistry C*, 111(48), 17671–17677.
- Shi, Z. Q., Wu, X., Wang, C. R., Lu, X., & Shinohara, H. (2006). Isolation and characterization of $Sc_2C_2@C_{68}$: A metal-carbide endofullerene with a non-IPR carbon cage. *Angewandte Chemie-International Edition*, 45(13), 2107–2111.
- Shimotani, H., Ito, T., Iwasa, Y., Taninaka, A., Shinohara, H., Nishibori, E., Takata, M., & Sakata, M. (2004). Quantum chemical study on the configurations of encapsulated metal ions and the molecular vibration modes in endohedral dimetallofullerene $La_2@C_{80}$. *Journal of the American Chemical Society*, 126(1), 364–369.
- Shinohara, H. (2000). Endohedral metallofullerenes. *Reports on Progress in Physics*, 63(6), 843–892.
- Shustova, N. B., Kuvychko, I. V., Bolskar, R. D., Seppelt, K., Strauss, S. H., Popov, A. A., & Boltalina, O. V. (2006). Trifluoromethyl derivatives of insoluble small-HOMO–LUMO-gap hollow higher fullerenes. NMR and DFT structure elucidation of $C_2-(C_{74}-D_{3h})(CF_3)_{12}$, $C_5-(C_{76}-T_d(2))(CF_3)_{12}$, $C_2-(C_{78}-D_{3h}(5))(CF_3)_{12}$, $C_5-(C_{80}-C_{2v}(5))(CF_3)_{12}$, and $C_2-(C_{82}-C_2(5))(CF_3)_{12}$. *Journal of the American Chemical Society*, 128(49), 15793–15798.
- Shustova, N. B., Newell, B. S., Miller, S. M., Anderson, O. P., Bolskar, R. D., Seppelt, K., Popov, A. A., Boltalina, O. V., & Strauss, S. H. (2007). Discovering and verifying elusive fullerene cage isomers: Structures of $C_2-p^{11}-(C_{74}-D_{3h})(CF_3)_{12}$ and $C_2-p^{11}-(C_{78}-D_{3h}(5))(CF_3)_{12}$. *Angewandte Chemie-International Edition*, 46(22), 4111–4114.
- Shustova, N. B., Chen, Y.-S., Mackey, M. A., Coumbe, C. E., Phillips, J. P., Stevenson, S., Popov, A. A., Strauss, S. H., & Boltalina, O. V. (2009). $Sc_3N@(C_{80}-I_h(7))(CF_3)_{14}$ and $Sc_3N@(C_{80}-I_h(7))(CF_3)_{16}$. Endohedral metallofullerene derivatives with exohedral addends on four and

- eight triple-hexagon junctions. Does the Sc₃N cluster control the addition pattern or vice versa? *Journal of the American Chemical Society*, 131(48), 17630–17637.
- Shustova, N. B., Mazej, Z., Chen, Y.-S., Popov, A. A., Strauss, S. H., & Boltalina, O. V. (2010). Saturnene revealed: X-ray crystal structure of D_{5d}-C₆₀F₂₀ formed in reactions of C₆₀ with A_xMF_y fluorinating agents (A = Alkali Metal; M = 3d Metal). *Angewandte Chemie International Edition*, 49, 812–815.
- Shustova, N. B., Peryshkov, D. V., Kuvychko, I. V., Chen, Y.-S., Mackey, M. A., Coumbe, C. E., Heaps, D. T., Confait, B. S., Heine, T., Phillips, J. P., Stevenson, S., Dunsch, L., Popov, A. A., Strauss, S. H., & Boltalina, O. V. (2011). Poly(perfluoroalkylation) of metallic nitride fullerenes reveals addition-pattern guidelines: Synthesis and characterization of a family of Sc₃N@C₈₀(CF₃)_n (n = 2–16) and their radical anions. *Journal of the American Chemical Society*, 133(8), 2672–2690.
- Simeonov, K. S., Amsharov, K. Y., & Jansen, M. (2007). Connectivity of the chiral D₂-symmetric isomer of C₇₆ through a crystal-structure determination C₇₆Cl₁₈·TiCl₄. *Angewandte Chemie-International Edition*, 46(44), 8419–8421.
- Simeonov, K. S., Amsharov, K. Y., Krokos, E., & Jansen, M. (2008). An Epilogue on the C₇₈-fullerene family: The discovery and characterization of an elusive isomer. *Angewandte Chemie International Edition*, 47, 6283–6285.
- Slanina, Z., Zhao, X., Deota, P., & Osawa, E. (2000). Relative stabilities of C₉₂ IPR fullerenes. *Journal of Molecular Modeling*, 6(2), 312–317.
- Slanina, Z., Ishimura, K., Kobayashi, K., & Nagase, S. (2004a). C₇₂ isomers: The IPR-satisfying cage is disfavored by both energy and entropy. *Chemical Physics Letters*, 384(1–3), 114–118.
- Slanina, Z., Zhao, X., Uhlik, F., Lee, S. L., & Adamowicz, L. (2004b). Computing enthalpy-entropy interplay for isomeric fullerenes. *International Journal of Quantum Chemistry*, 99(5), 640–653.
- Slanina, Z., Chen, Z. F., Schleyer, P. V., Uhlik, F., Lu, X., & Nagase, S. (2006). La₂@C₇₂ and Sc₂@C₇₂: Computational characterizations. *Journal of Physical Chemistry A*, 110(6), 2231–2234.
- Slanina, Z., Uhlik, F., Sheu, J. H., Lee, S. L., Adamowicz, L., & Nagase, S. (2008). Stabilities of fullerenes: Illustration on C₈₀. *Match-Communications in Mathematical and in Computer Chemistry*, 59, 225–238.
- Spielmann, H. P., Weedon, B. R., & Meier, M. S. (2000). Preparation and NMR characterization of C₇₀H₁₀: Cutting a fullerene pi-system in half. *Journal of Organic Chemistry*, 65(9), 2755–2758.
- Stevenson, S., Rice, G., Glass, T., Harich, K., Cromer, F., Jordan, M. R., Craft, J., Hadju, E., Bible, R., Olmstead, M. M., Maitra, K., Fisher, A. J., Balch, A. L., & Dorn, H. C. (1999). Small-bandgap endohedral metallofullerenes in high yield and purity. *Nature*, 401(6748), 55–57.
- Stevenson, S., Fowler, P. W., Heine, T., Duchamp, J. C., Rice, G., Glass, T., Harich, K., Hajdu, E., Bible, R., & Dorn, H. C. (2000). Materials science – A stable non-classical metallofullerene family. *Nature*, 408(6811), 427–428.
- Stevenson, S., Mackey, M. A., Stuart, M. A., Phillips, J. P., Easterling, M. L., Chancellor, C. J., Olmstead, M. M., & Balch, A. L. (2008). A distorted tetrahedral metal oxide cluster inside an icosahedral carbon cage. Synthesis, isolation, and structural characterization of Sc₄(μ³-O)₂@I_h-C₈₀. *Journal of the American Chemical Society*, 130(36), 11844–11845.
- Sun, G. Y. (2003a). Assigning the major isomers of fullerene C₈₈ by theoretical ¹³C NMR spectra. *Chemical Physics Letters*, 367(1–2), 26–33.
- Sun, G. Y. (2003b). Theoretical ¹³C NMR chemical shifts of the stable isomers of fullerene C₉₀. *Chemical Physics*, 289(2–3), 371–380.
- Sun, G. Y., & Kertesz, M. (2000). Theoretical ¹³C NMR spectra of IPR isomers of fullerenes C₆₀, C₇₀, C₇₂, C₇₄, C₇₆, and C₇₈ studied by density functional theory. *Journal of Physical Chemistry A*, 104(31), 7398–7403.
- Sun, G. Y., & Kertesz, M. (2001a). Identification for IPR isomers of fullerene C₈₂ by theoretical ¹³C NMR spectra calculated by density functional theory. *Journal of Physical Chemistry A*, 105(22), 5468–5472.
- Sun, G. Y., & Kertesz, M. (2001b). Isomer identification for fullerene C₈₄ by ¹³C NMR spectrum: A density-functional theory study. *Journal of Physical Chemistry A*, 105(21), 5212–5220.

- Sun, G. Y., & Kertesz, M. (2002). ^{13}C NMR spectra for IPR isomers of fullerene C_{86} . *Chemical Physics*, 276(2), 107–114.
- Suzuki, M., Mizorogi, N., Yang, T., Uhlík, F., Slanina, Z., Zhao, X., Yamada, M., Maeda, Y., Hasegawa, T., Nagase, S., Lu, X., & Akasaka, T. (2013). $\text{La}_2@C_s(17,490)\text{-C}_{76}$: A new non-IPR dimetallic metallofullerene featuring unexpectedly weak metal–pentalene interactions. *Chemistry - A European Journal*, 19(50), 17125–17130.
- Tagmatarchis, N., Avent, A. G., Prassides, K., Dennis, T. J. S., & Shinohara, H. (1999). Separation, isolation and characterisation of two minor isomers of the [84]fullerene C_{84} . *Chemical Communications*, 1999(11), 1023–1024.
- Tagmatarchis, N., Arcon, D., Prato, M., & Shinohara, H. (2002). Production, isolation and structural characterization of [92]fullerene isomers. *Chemical Communications*, 2002(24), 2992–2993.
- Takata, M., Nishibori, E., Sakata, M., & Shinohara, H. (2003). Synchrotron radiation for structural chemistry – Endohedral natures of metallofullerenes found by synchrotron radiation powder method. *Structural Chemistry*, 14(1), 23–38.
- Tamm, N. B., & Troyanov, S. I. (2014). Synthesis, isolation, and X-ray structural characterization of trifluoromethylated C_{90} fullerenes: $\text{C}_{90}(30)(\text{CF}_3)_{18}$ and $\text{C}_{90}(35)(\text{CF}_3)_{14}$. *Nanosystems: Physics, Chemistry, Mathematics*, 5(1), 39–45.
- Tamm, N. B., & Troyanov, S. I. (2015a). Capturing C_{90} isomers as CF_3 derivatives: $\text{C}_{90}(30)(\text{CF}_3)_{14}$, $\text{C}_{90}(35)(\text{CF}_3)_{16/18}$, and $\text{C}_{90}(45)(\text{CF}_3)_{16/18}$. *Chemistry – An Asian Journal*, 10(8), 1622–1625.
- Tamm, N. B., & Troyanov, S. I. (2015b). New isolated-pentagon-rule isomer of C_{92} isolated as trifluoromethyl and chlorido derivatives: $\text{C}_{92}(38)(\text{CF}_3)_{14/16}$ and $\text{C}_{92}(38)\text{Cl}_{20/22}$. *Inorganic Chemistry*, 54(22), 10527–10529.
- Tamm, N. B., Sidorov, L. N., Kemnitz, E., & Troyanov, S. I. (2009a). Crystal structures of $\text{C}_{94}(\text{CF}_3)_{20}$ and $\text{C}_{96}(\text{C}_2\text{F}_5)_{12}$ reveal the cage connectivities in $\text{C}_{94}(61)$ and $\text{C}_{96}(145)$ fullerenes. *Angewandte Chemie International Edition*, 48, 9102–9104.
- Tamm, N. B., Sidorov, L. N., Kemnitz, E., & Troyanov, S. I. (2009b). Isolation and structural X-ray investigation of perfluoroalkyl derivatives of six cage isomers of C_{84} . *Chemistry - A European Journal*, 15(40), 10486–10492.
- Tamm, N. B., Yang, S., Wei, T., & Troyanov, S. I. (2015). Five isolated pentagon rule isomers of higher fullerene C_{94} captured as chlorides and CF_3 derivatives: $\text{C}_{94}(34)\text{Cl}_{14}$, $\text{C}_{94}(61)\text{Cl}_{20}$, $\text{C}_{94}(133)\text{Cl}_{22}$, $\text{C}_{94}(42)(\text{CF}_3)_{16}$, and $\text{C}_{94}(43)(\text{CF}_3)_{18}$. *Inorganic Chemistry*, 54(6), 2494–2496.
- Tan, K., & Lu, X. (2005). T_2C_{80} is more likely a titanium carbide endohedral metallofullerene (Ti_2C_2)@ C_{78} . *Chemical Communications*, 2005(35), 4444–4446.
- Tan, Y.-Z., Xie, S.-Y., Huanh, R.-B., & Zheng, I.-S. (2009). The stabilization of fused-pentagon fullerene molecules. *Nature Chemistry*, 1, 450–460.
- Tebbe, F. N., Harlow, R. L., Chase, D. B., Thorn, D. L., Campbell, G. C., Calabrese, J. C., Herron, N., Young, R. J., & Wasserman, E. (1992). Synthesis and single-crystal X-ray structure of a highly symmetrical C_{60} derivative, $\text{C}_{60}\text{Br}_{24}$. *Science*, 256(5058), 822–825.
- Troyanov, S. I., & Kemnitz, E. (2005). Synthesis and structures of fullerene bromides and chlorides. *European Journal of Organic Chemistry*, 2005(23), 4951–4962.
- Troyanov, S. I., & Kemnitz, E. (2012). Synthesis and structure of halogenated fullerenes. *Current Organic Chemistry*, 16(9), 1060–1078.
- Troyanov, S. I., & Tamm, N. B. (2009). Cage connectivities of $\text{C}_{88}(33)$ and $\text{C}_{92}(82)$ fullerenes captured as trifluoromethyl derivatives, $\text{C}_{88}(\text{CF}_3)_{18}$ and $\text{C}_{92}(\text{CF}_3)_{16}$. *Chemical Communications*, 2009(40), 6035–6037.
- Troyanov, S. I., Popov, A. A., Denisenko, N. I., Boltalina, O. V., Sidorov, L. N., & Kemnitz, E. (2003). The first X-ray crystal structures of halogenated [70]fullerene: $\text{C}_{70}\text{Br}_{10}$ and $\text{C}_{70}\text{Br}_{10}\cdot 3\text{Br}_2$. *Angewandte Chemie-International Edition*, 42(21), 2395–2398.
- Troyanov, S. I., Shustova, N. B., Ioffe, I. N., Turnbull, A. P., & Kemnitz, E. (2005). Synthesis and structural characterization of highly chlorinated C_{70} , $\text{C}_{70}\text{Cl}_{28}$. *Chemical Communications*, 1, 72–74.
- Troyanov, S. I., Dimitrov, A., & Kemnitz, E. (2006). Selective synthesis of a trifluoromethylated fullerene and the crystal structure of $\text{C}_{60}(\text{CF}_3)_{12}$. *Angewandte Chemie-International Edition*, 45(12), 1971–1974.

- Troyanov, S. I., Goryunkov, A. A., Dorozhkin, E. I., Ignat'eva, D. V., Tamm, N. B., Avdoshenko, S. M., Ioffe, I. N., Markov, V. Y., Sidorov, L. N., Scheurel, K., & Kemnitz, E. (2007). Higher trifluoromethylated derivatives of C₆₀, C₆₀(CF₃)₁₆ and C₆₀(CF₃)₁₈ – Synthesis, structure, and theoretical study. *Journal of Fluorine Chemistry*, 128(5), 545–551.
- Troyanov, S. I., Yang, S., Chen, C., & Kemnitz, E. (2011). Six IPR isomers of C₉₀ fullerene captured as chlorides: Carbon cage connectivities and chlorination patterns. *Chemistry - A European Journal*, 17(38), 10662–10669.
- Tsuchiya, T., Wakahara, T., Maeda, Y., Akasaka, T., Waelchli, M., Kato, T., Okubo, H., Mizorogi, N., Kobayashi, K., & Nagase, S. (2005). 2D NMR characterization of the La@C₈₂ anion. *Angewandte Chemie-International Edition*, 44(21), 3282–3285.
- Valencia, R., Rodríguez-Fortea, A., & Poblet, J. M. (2007). Large fullerenes stabilized by encapsulation of metallic clusters. *Chemical Communications*, 40, 4161–4163.
- Valencia, R., Rodríguez-Fortea, A., & Poblet, J. M. (2008). Understanding the stabilization of metal carbide endohedral fullerenes M₂C₂@C₈₂ and related systems. *Journal of Physical Chemistry A*, 112(20), 4550–4555.
- Wakabayashi, T., Kikuchi, K., Suzuki, S., Shiromaru, H., & Achiba, Y. (1994). Pressure-controlled selective isomer formation of fullerene-C₇₈. *Journal of Physical Chemistry*, 98(12), 3090–3091.
- Wakahara, T., Kobayashi, J., Yamada, M., Maeda, Y., Tsuchiya, T., Okamura, M., Akasaka, T., Waelchli, M., Kobayashi, K., Nagase, S., Kato, T., Kako, M., Yamamoto, K., & Kadish, K. M. (2004). Characterization of Ce@C₈₂ and its anion. *Journal of the American Chemical Society*, 126(15), 4883–4887.
- Wakahara, T., Nikawa, H., Kikuchi, T., Nakahodo, T., Rahman, G. M. A., Tsuchiya, T., Maeda, Y., Akasaka, T., Yoza, K., Horn, E., Yamamoto, K., Mizorogi, N., Slanina, Z., & Nagase, S. (2006). La@C₇₂ having a non-IPR carbon cage. *Journal of the American Chemical Society*, 128(44), 14228–14229.
- Wang, C. R., Kai, T., Tomiyama, T., Yoshida, T., Kobayashi, Y., Nishibori, E., Takata, M., Sakata, M., & Shinohara, H. (2000a). C₆₆ fullerene encaging a scandium dimer. *Nature*, 408, 426–427.
- Wang, C. R., Sugai, T., Kai, T., Tomiyama, T., & Shinohara, H. (2000b). Production and isolation of an ellipsoidal C₈₀ fullerene. *Chemical Communications*, 7, 557–558.
- Wang, C. R., Kai, T., Tomiyama, T., Yoshida, T., Kobayashi, Y., Nishibori, E., Takata, M., Sakata, M., & Shinohara, H. (2001). A scandium carbide endohedral metallofullerene: (Sc₂C₂)@C₈₄. *Angewandte Chemie-International Edition*, 40(2), 397–399.
- Wang, T.-S., Chen, N., Xiang, J.-F., Li, B., Wu, J.-Y., Xu, W., Jiang, L., Tan, K., Shu, C.-Y., Lu, X., & Wang, C.-R. (2009). Russian-doll-type metal carbide endofullerene: Synthesis, isolation, and characterization of Sc₄C₂@C₈₀. *Journal of the American Chemical Society*, 131(46), 16646–16647.
- Wang, Z., Yang, H., Jiang, A., Liu, Z., Olmstead, M. M., & Balch, A. L. (2010). Structural similarities in C₃(16)-C₈₆ and C₂(17)-C₈₆. *Chemical Communications*, 46(29), 5262–5264.
- Wang, S., Yang, S., Kemnitz, E., & Troyanov, S. I. (2016). Unusual chlorination patterns of three IPR isomers of C₈₈ fullerene in C₈₈(7)Cl_{12/24}, C₈₈(17)Cl₂₂, and C₈₈(33)Cl_{12/14}. *Chemistry – An Asian Journal*. doi:10.1002/asia.201501152.
- Whitaker, J. B., Kuvychko, I. V., Shustova, N. B., Chen, Y.-S., Strauss, S. H., & Boltalina, O. V. (2013). An elusive fulvene 1,7,11,24-C₆₀(CF₃)₄ and its unusual reactivity. *Chemical Communications*, 50, 1205–1208.
- Wu, J., & Hagelberg, F. (2008). Computational study on C₈₀ enclosing mixed trimetallic nitride clusters of the form Gd_xM_{3-x}N (M = Sc, Sm, Lu). *Journal of Physical Chemistry C*, 112(15), 5770–5777.
- Xiao, Z., Wang, F. D., Huang, S. H., Gan, L. B., Zhou, J., Yuan, G., Lu, M. J., & Pan, J. Q. (2005). Regiochemistry of [70]fullerene: Preparation of C₇₀((OOBu)⁻tBu)_n, (n = 2, 4, 6, 8, 10) through both equatorial and cyclopentadienyl addition modes. *Journal of Organic Chemistry*, 70(6), 2060–2066.
- Xu, L., Cai, W. S., & Shao, X. G. (2006). Prediction of low-energy isomers of large fullerenes from C₁₃₂ to C₁₆₀. *Journal of Physical Chemistry A*, 110(29), 9247–9253.

- Xu, L., Cai, W. S., & Shao, X. G. (2007). Performance of the semiempirical AM1, PM3, MNDO, and tight-binding methods in comparison with DFT method for the large fullerenes C_{116} – C_{120} . *Journal of Molecular Structure-Theochem*, 817(1–3), 35–41.
- Xu, L., Cai, W., & Shao, X. (2008). Systematic search for energetically favored isomers of large fullerenes C_{122} – C_{130} and C_{162} – C_{180} . *Computational Materials Science*, 41, 522–528.
- Yamada, M., Wakahara, T., Tsuchiya, T., Maeda, Y., Kako, M., Akasaka, T., Yoza, K., Horn, E., Mizorogi, N., & Nagase, S. (2008). Location of the metal atoms in $Ce_2@C_{78}$ and its bis-silylated derivative. *Chemical Communications*, 5, 558–560.
- Yamada, M., Kurihara, H., Suzuki, M., Guo, J. D., Waelchli, M., Olmstead, M. M., Balch, A. L., Nagase, S., Maeda, Y., Hasegawa, T., Lu, X., & Akasaka, T. (2014). $Sc_2@C_{66}$ revisited: An endohedral fullerene with scandium ions nestled within two unsaturated linear triquinanes. *Journal of the American Chemical Society*, 136(21), 7611–7614.
- Yang, S., Popov, A. A., & Dunsch, L. (2007a). The role of an asymmetric nitride cluster on a fullerene cage: The non-IPR endohedral $DySc_2N@C_{76}$. *Journal of Physical Chemistry B*, 111(49), 13659–13663.
- Yang, S. F., Popov, A. A., & Dunsch, L. (2007b). Violating the isolated pentagon rule (IPR): The endohedral non-IPR cage of $Sc_3N@C_{70}$. *Angewandte Chemie-International Edition*, 46(8), 1256–1259.
- Yang, S., Yoon, M., Hicke, C., Zhang, Z., & Wang, E. (2008). Electron transfer and localization in endohedral metallofullerenes: Ab initio density functional theory calculations. *Physical Review B*, 78, 115435.
- Yang, H., Beavers, C. M., Wang, Z., Jiang, A., Liu, Z., Jin, H., Mercado, B. Q., Olmstead, M. M., & Balch, A. L. (2010). Isolation of a small carbon nanotube: The surprising appearance of $D_{5h}(1)$ - C_{90} . *Angewandte Chemie International Edition*, 49(5), 886–890.
- Yang, H., Mercado, B. Q., Jin, H., Wang, Z., Jiang, A., Liu, Z., Beavers, C. M., Olmstead, M. M., & Balch, A. L. (2011). Fullerenes without symmetry: Crystallographic characterization of $C_1(30)$ - C_{90} and $C_1(32)$ - C_{90} . *Chemical Communications*, 47(7), 2068–2070.
- Yang, S., Wei, T., Kemnitz, E., & Troyanov, S. I. (2012a). The most stable IPR isomer of C_{88} fullerene, C_s - $C_{88}(17)$, revealed by X-ray structures of $C_{88}Cl_{16}$ and $C_{88}Cl_{22}$. *Chemistry – An Asian Journal*, 7(2), 290–293.
- Yang, S., Wei, T., Kemnitz, E., & Troyanov, S. I. (2012b). Four isomers of C_{96} fullerene structurally proven as $C_{96}Cl_{22}$ and $C_{96}Cl_{24}$. *Angewandte Chemie International Edition*, 51(33), 8239–8242.
- Yang, S., Wei, T., & Troyanov, S. I. (2013). A new isomer of pristine higher fullerene C_s - $C_{82}(4)$ captured by chlorination as $C_{82}Cl_{20}$. *Chemistry – An Asian Journal*, 8(2), 351–353.
- Yang, S., Wei, T., Kemnitz, E., & Troyanov, S. I. (2014). First isomers of pristine C_{104} fullerene structurally confirmed as chlorides, $C_{104}(258)Cl_{16}$ and $C_{104}(812)Cl_{24}$. *Chemistry – An Asian Journal*, 9(1), 79–82.
- Yanov, I., Kholod, Y., Simeon, T., Kaczmarek, A., & Leszczynski, J. (2006). Local minima conformations of the $Sc_3N@C_{80}$ endohedral complex: Ab initio quantum chemical study and suggestions for experimental verification. *International Journal of Quantum Chemistry*, 106(14), 2975–2980.
- Yumura, T., Sato, Y., Suenaga, K., & Iijima, S. (2005). Which do endohedral Ti_2C_{80} metallofullerenes prefer energetically: $Ti_2@C_{80}$ or $Ti_2C_2@C_{78}$? A theoretical study. *Journal of Physical Chemistry B*, 109(43), 20251–20255.
- Zalibera, M., Rapta, P., & Dunsch, L. (2007). In situ ESR–UV/VIS/NIR spectroelectrochemistry of an empty fullerene anion and cation: The $C_{82:3}$ isomer. *Electrochemistry Communications*, 9(12), 2843–2847.
- Zhang, J., Hao, C., Li, S. M., Mi, W. H., & Jin, P. (2007). Which configuration is more stable for $La_2@C_{80}$, D_{3d} or D_{2h} ? Recomputation with ZORA methods within ADF. *Journal of Physical Chemistry C*, 111(22), 7862–7867.
- Zhang, J., Bowles, F. L., Bearden, D. W., Ray, W. K., Fuhrer, T., Ye, Y., Dixon, C., Harich, K., Helm, R. F., Olmstead, M. M., Balch, A. L., & Dorn, H. C. (2013). A missing link in the transformation from asymmetric to symmetric metallofullerene cages implies a top-down fullerene formation mechanism. *Nature Chemistry*, 5(10), 880–885.

- Zhang, Y., Ghiassi, K. B., Deng, Q., Samoylova, N. A., Olmstead, M. M., Balch, A. L., & Popov, A. A. (2015). Synthesis and structure of $\text{LaSc}_2\text{N}@C_5(\text{hept})\text{-C}_{80}$ with one heptagon and thirteen pentagons. *Angewandte Chemie International Edition*, 52(2), 495–499.
- Zhao, X., Slanina, Z., Goto, H., & Osawa, E. (2003). Theoretical investigations on relative stabilities of fullerene C_{94} . *Journal of Chemical Physics*, 118(23), 10534–10540.
- Zheng, G. S., Irle, S., & Morokuma, K. (2005). Performance of the DFTB method in comparison to DFT and semiempirical methods for geometries and energies of $\text{C}_{20}\text{--C}_{86}$ fullerene isomers. *Chemical Physics Letters*, 412(1–3), 210–216.
- Zuo, T. M., Beavers, C. M., Duchamp, J. C., Campbell, A., Dorn, H. C., Olmstead, M. M., & Balch, A. L. (2007). Isolation and structural characterization of a family of endohedral fullerenes including the large, chiral cage fullerenes $\text{Tb}_3\text{N}@C_{88}$ and $\text{Tb}_3\text{N}@C_{86}$ as well as the I_h and D_{5h} isomers of $\text{Tb}_3\text{N}@C_{80}$. *Journal of the American Chemical Society*, 129(7), 2035–2043.
- Zuo, T., Walker, K., Olmstead, M. M., Melin, F., Holloway, B. C., Echegoyen, L., Dorn, H. C., Chaur, M. N., Chancellor, C. J., Beavers, C. M., Balch, A. L., & Athans, A. J. (2008). New egg-shaped fullerenes: non-isolated pentagon structures of $\text{Tm}_3\text{N}@C_5(51365)\text{-C}_{84}$ and $\text{Gd}_3\text{N}@C_5(51365)\text{-C}_{84}$. *Chemical Communications*, 2008(9), 1067–1069.
- Zywietz, T. K., Jiao, H., Schleyer, R., & de Meijere, A. (1998). Aromaticity and antiaromaticity in oligocyclic annelated five-membered ring systems. *Journal of Organic Chemistry*, 63(10), 3417–3422.

Panagiotis Karamanis

Contents

Introduction	1098
Structural Properties of Semiconductor Clusters	1099
The Ground-State Structure	1099
Structural Determination	1101
Silicon Clusters	1103
III–V and II–VI Semiconductor Clusters	1108
Electric Polarizabilities and Hyperpolarizabilities of Clusters	1114
Definitions and Theory	1114
Cluster (Hyper)Polarizabilities: Computational Approach	1117
Semiconductor Cluster (Hyper)polarizabilities: General Trends and Selected Studies ..	1123
Concluding Remarks	1128
Bibliography	1129

Abstract

Materials that exhibit an electrical resistivity between that of conductor and insulator are called semiconductors. Devices based on semiconductor materials, such as transistors, solar cells, light-emitting diodes, digital integrated circuits, solar photovoltaics, and much more, are the base of modern electronics. Silicon is used in most of the semiconductor devices while other materials such as germanium, gallium arsenide, and silicon carbide are used for specialized applications. The obvious theoretical and technological importance of semiconductor materials has led to phenomenal success in making semiconductors with near-atomic precision such as quantum wells, wires, and dots. As a result, there is a lot of undergoing research in semiconductor clusters of small and medium

P. Karamanis (✉)

Groupe de Chimie Théorique et Réactivité, ECP, IPREM UMR 5254, Université de Pau et de Pays de l'Adour, PAU Cedex, France

e-mail: panos@chemistry.upatras.gr; panagiotis.karamanis@univ-pau.fr

sizes both experimentally and by means of computational chemistry since the miniaturization of devices still continues. In the next pages, we are going to learn which the most studied semiconductor clusters are, we will explore their basic structural features and visit some of the most representative *ab initio* studies that are considered as works of reference in this research realm. Also, we are going to be introduced to the theory of the electric properties applied in the case of clusters by visiting some of the most illustrative studies into this research area. It is one of the purposes of this presentation to underscore the strong connection between the electric properties of clusters and their structure.

Introduction

About two decades ago Robert Pool stated: “Clusters are strange morsels of matter: when metals or semiconductors are shrunk down to clumps only 10 or 100 atoms in size, they become a “totally new class of materials” with potentially valuable applications” (Pool 1990). This distinctive behavior of clusters has triggered an explosion of scientific work in this area and it is in close relation with the dramatic development of a new field in science. This field is now widely known as “nanoscience.” The term nanoscience comes from the terms “nano” and “science.” The word “nano” is used to define a specific length of one billionth of a meter. This size scale resides between bulk materials and typical molecular dimensions. Thus, nanoscience is associated with the study of structures, materials, and devices, which are larger than the typical molecules but extremely small, usually in the range from 0.1 to 100 nm. The increasing importance of this area and its rapid expansion is owed mainly to the requirements of modern technologies for a vast diversity of new materials with more than one function. The field of nanoscience is highly multidisciplinary and joins different disciplines such as chemistry, applied physics, materials science, colloidal science, device physics, supramolecular chemistry, and even more, mechanical and electrical engineering. As a matter of fact, nanoscience can be viewed as extension of most of existing sciences into the nanoscale or the upgrade of the existing molecular sciences to the nanoscale dimensions.

Among nano-objects clusters occupy a large part of the big picture. The so-called nanoclusters apart from a bridge between molecules and solids are systems with properties of their own which are considerably different from those of the bulk materials and from those of their constituent parts (atoms and molecules). Under specific conditions those species can be considered as the building modules of nano-objects which are the foundations of nanoscience. Clusters are aggregates of atoms or molecules in the size of few nanometers and their constitution ranges between 10 and 10^6 atoms. They can be built from only one type of atoms such as the homogeneous silicon clusters, or from more than one element such as the heterogeneous gallium arsenide binary clusters. Atomic clusters can be met in neutral or charged forms while their atoms bind together by almost all of the main bond types such as covalent, ionic, van der Waals, and metallic.

Amongst the different types of clusters that have been the subject of numerous studies in the past, clusters which originate from semiconductor materials have attracted considerable attention. In spite of their origin, these entities are not necessarily semiconductors. Their basic bonding features resemble those of the metals but they keep as well some of the covalent bonding characteristics which dominate in semiconductor nanocrystals.

The purpose of this presentation is to bring together the most representative *ab initio* studies of the non-bulk-like structural and electric properties of semiconductor clusters. Both of these two different property classes are closely related to each other and have occupied the interest of the researcher in numerous studies.

Structural Properties of Semiconductor Clusters

Excluding clusters built from carbon (fullerenes) then the most studied semiconductor clusters are those made from silicon (Si) and gallium arsenide (GaAs). GaAs clusters belong to the so-called III–V group of semiconductors which are formed between the elements of groups 13 and 15. Also, there is a significant amount of work focusing on the II–VI semiconductor clusters which are formed between the elements of the 12 and 15 groups. Especially for the latter clusters there is significant amount of reported experimental and theoretical work due to their unique optical properties. The most studied II–VI semiconductor clusters are those built from cadmium sulfide which have served as prototypes for studies of the effect of quantum confinement on electronic and optical properties of semiconductor nano-objects. For instance, CdS quantum dots which are water soluble and biocompatible (Bruchez et al. 1998), are potentially useful as fluorescent biological labels since they are characterized by large optical gaps. Also, cadmium selenide and telluride clusters have attracted considerable attention due to their narrow bandgaps and large Bohr exciton radii. The last characteristic implies that nanostructures built from those systems are expected to exhibit strong quantum confinement and this feature makes both CdSe and CdTe nano-objects very interesting to explore their properties. Other kinds of semiconductor clusters that have been the subject of a notable number of computational studies are clusters built from GaN, GaP, InP, InAs, ZnO, ZnS, AlP, HgTe, Ge, SnTe, BeTe, BiTe, BN, BAs, and BP.

The Ground-State Structure

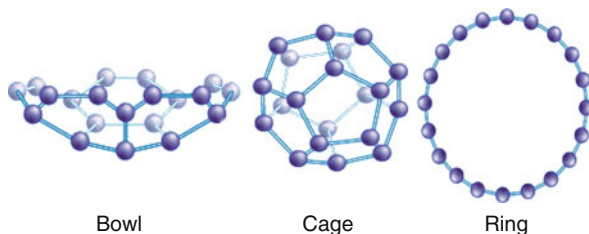
In principle, the most favorable atom arrangement of a cluster should be the one that is characterized by the smallest total energy. This structure is generally called the ground-state structure and at zero or low enough temperatures should be the mathematical global minimum of the potential energy of a cluster as a function of the coordinates of the atomic centers (the potential energy surface, PES). The arrangement of the atoms in the ground-state structure is the one that is considered in the majority of cluster property studies, assuming that most of

those properties are defined by the properties of the lowest energy configuration. As a result, a respectable number of various computational schemes have been developed and proposed in order to treat the ground-state structure problem in an accurate and efficient manner. However, the solution to this problem is neither obvious nor simple. In the contrary, it corresponds to a task of extremely high computational complexity. The problem becomes even harder given that the available experimental methods are not yet in the position to provide the necessary microscopic structural details for the complete understanding of clusters structures in order for the developed theoretical methods to be verified or improved. Also, the developed computational methods face certain difficulties that are mainly delivered from “practical” limitations of the algorithms used or restrictions of the quantum chemical methods in the precise calculation of the potential energy surface of a given cluster. Accordingly, in several cases the global minimum of the potential energy cluster surface exhibits a strong dependence on the quantum chemical method one uses in order to explore it (Karamanis et al. 2007b). A classical example of the above statement is the case of C_{20} . This cluster is the smallest possible fullerene, and experiments suggest that it should be a planar ring (Castro et al. 2002; Dugourd et al. 1998; Prinzbach et al. 2000). However, detailed theoretical studies (Grossman et al. 1995) by means of quantum Monte Carlo, all-electron fixed-node quantum Monte Carlo (Sokolova et al. 2000), single- and multireference MP2 methods (Grimme and Mück-Lichtenfeld 2002) and high-level ab initio methods (An et al. 2005) pointed out the bowl and the cage structures as more stable. The answer to this puzzle is that both experiments and theory are correct since if one introduces finite-temperature corrections to the calculations then ring structure becomes more stable (see Fig. 1).

Furthermore, some specific clusters show a peculiar behavior well known as “fluxional behavior.” This behavior is connected with the fact that at a given temperature the clusters are expected to be in some sort of thermal equilibrium rapidly interconverting from one stable isomer to the other due to fast and sometimes drastic atomic rearrangements. During a structural rearrangement of this kind, the resulting “isomerization” leads to structures characterized by different bonding patterns. Thus, it would not be irrational for one to claim that there is no guarantee that the theoretically predicted ground states are favored in all situations and condition.

In order to overcome such problems many studies extend their property investigations to stationary points which correspond to cluster conurations energetically

Fig. 1 The three different structures of C_{20} competing for the ground state



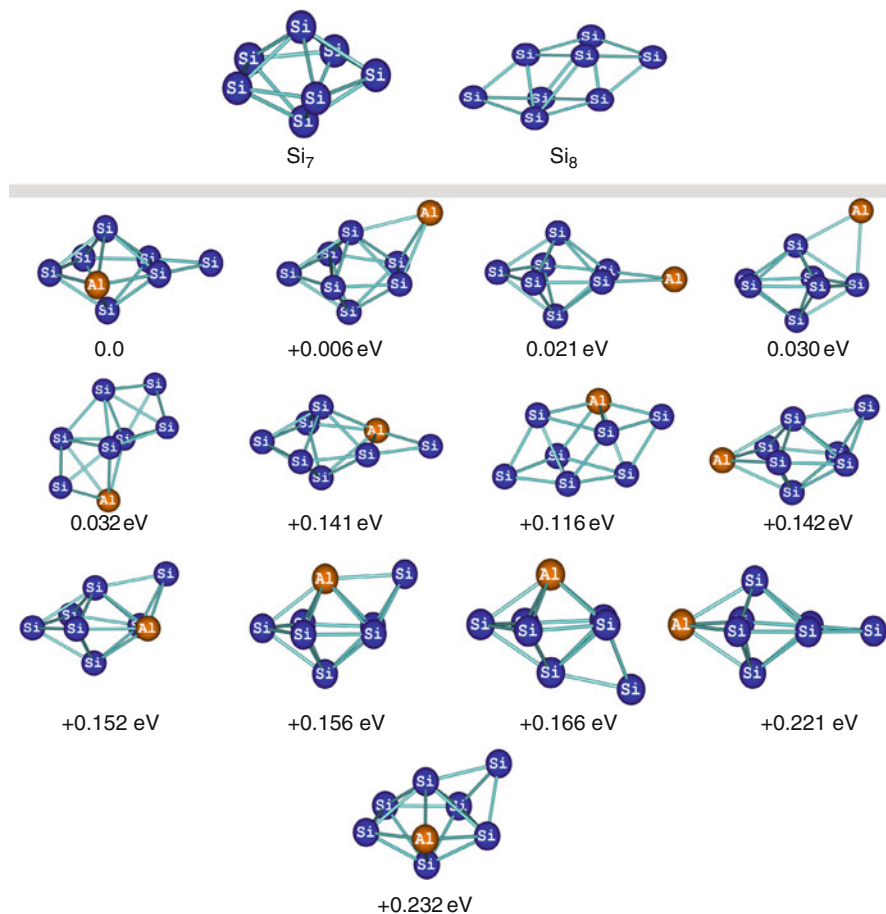


Fig. 2 Ground state structures and low lying isomers of AlSi_7 in comparison with the lowest energy structures of Si_7 and Si_8

close to the ground state. These clusters are meta-stable species and their study is expected to complement the investigations which focus on the narrow area of the global minima. A demonstration of the numerous stable local minima that exist even for clusters of small size is given in Fig. 2 which depicts the ground-state structures of aluminum-doped clusters along with several low-lying isomers which are extremely close in energy (Karamanis et al. 2010).

Structural Determination

There are two major classes of global computational methods: the unbiased and the seeded methods. The first class works independently of initial cluster configurations

and is considered as the most significant and reliable but the more computationally demanding. The second class utilizes a set of initial structures as seeds and they are usually faster than the unbiased; however, there is always the possibility of disregarding configurations of low energy that are not expected. Another optional way of finding the most favorable structures of clusters is based on guiding principles and structural rules stability principles, such those that have been originally developed and tested over the years for boranes and carboranes and bisboranes. Although this kind of approach is more of an art than of an automatic algorithmic computational approach it has been utilized with remarkable success in various systems such as pure or doped silicon cluster, or clusters made from gallium arsenide, cadmium sulfide, selenide, or telluride (see for instance, Al-Laham and Raghavachari 1991, 1993; Lipscomb 1966; Zdetsis 2008, 2009; Williams 1992; Gutsev et al. 2008a, b; Raghavachari and Logovinsky 1985, Raghavachari and Rohlfing 1988; Matxain et al. 2000, 2001, 2003, 2004). In any case though, whatever the method chosen, one has to deal with the existence of a vast number of local minima even in the case of clusters composed by few atoms.

Let us now see in brief some of the developed algorithms that have been used in cases that involve semiconductor clusters. The most commonly used computational schemes in finding the ground-state structures of semiconductor clusters are those based on genetic algorithms that simulate the evolution in nature through natural selection. This approach utilizes concepts such as chromosome mixing, mutations, and the selection of the fittest. Another family of algorithms is the so-called basin hopping which use canonical Monte Carlo simulations at a certain temperature. The basic idea of the basin hopping algorithms is to lower the energetic barriers that separate the various local minima of a cluster by leaving the number and the kinds of the local minima unchanged. This has been proven very efficient in scanning different areas of the potential energy surface and has been successfully applied in cases of clusters up to 100 atoms. A different approach which has been applied in a great variety of systems is the thermal simulated annealing. This method lets the system to evolve initially at a high temperature and then a cooling down procedure starts by reducing gradually the temperature of the system to zero temperature. At each step some neighbor structures are considered based on probability rules and then it is decided whether this new structures are accepted or not following certain energy criteria (e.g. by moving to states of lower energy). As a result, the search space which contains appropriate solutions to the problem considerably narrows down. However, if the free energy global minimum changes at low temperatures where dynamical relaxation is slow, the algorithms may become stuck in the structure corresponding to the high-temperature free energy global minimum. Thus, although this algorithm is easy to follow and understand because of its physical concepts, it is less efficient than methods based on genetic and basin hopping algorithms. Finally, another method that has been developed to treat the problem of rich potential energy surfaces close to the minimum is the so-called Global Search Algorithm of Minima (GSAM) (Marchal et al. 2009, 2010, 2011). In brief, the GSAM includes three major parts: the first part is devoted to the generation of an initial guess set of structures for a large number of cluster structures

randomly generated by several techniques. The second part performs an automatic selection of the most appropriate structures through a special pattern recognition technique (Karamanis et al. 2009), which are expected to lead to different local minima. Finally, the third part comprises full geometry optimizations of the selected configurations. This last part yields the most stable cluster structures and the structure which corresponds to the ground state among them.

Further information into the subject of the methods that have been used to treat cluster structural problems in a global manner can be found in the Refs. Bazterra et al. (2004), Biswas and Hamann (1986), Blaisten-Barojas and Levesque (1986), Hamad et al. (2005), Ho et al. (1998), Hossain et al. (2007), Jelski et al. (1991), Menon and Subbaswamy (1995), Nair et al. (2004), Pedroza and Da Silva (2007), Sokolova et al. (2000), Tekin and Hartke (2004), Yoo et al. (2004, 2008), Yoo and Zeng (2005, 2004).

Silicon Clusters

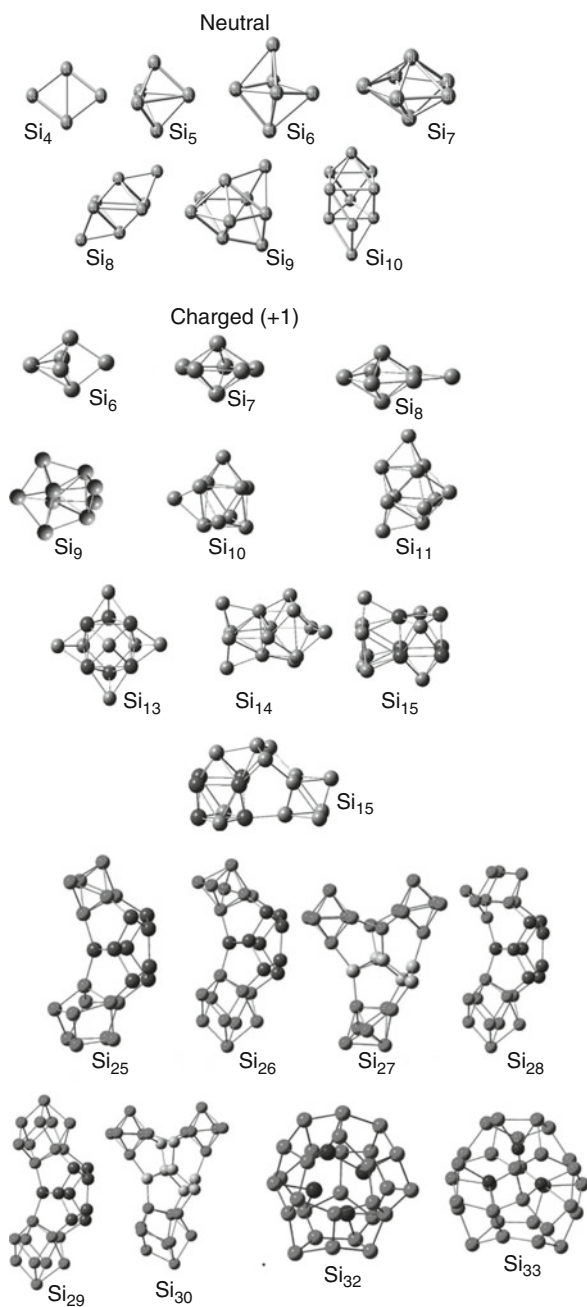
Selected Structural Studies

As we have already mentioned, among the various semiconductor clusters studied during the past two decades, those made from silicon have attracted most of the attention. A significant number of algorithms and computational strategies have been employed to identify the ground states and the low-lying isomers of silicon clusters of small and medium size with remarkable success. As a result, it is now a common sense that the silicon atoms of a cluster are “different” from those of the bulk material.

It is evident that most of the studies on silicon clusters concern their neutral forms. However, it is impossible to neglect investigations on charged clusters since they have led to very significant results and conclusions (Lyon et al. 2009; Xiao et al. 2002; Nigam et al. 2004; Wei et al. 1997; Li et al. 2003, Li 2005; Zhou and Pan 2008). One of the most impressive investigations on this subject is the recent study reported by Lyon et al. who studied for the first time the experimental infrared spectra of positively charged silicon bare clusters up to 21 atoms in the gas phase. In this study the experimental spectra were compared to theoretical spectral predictions and the authors managed to make unambiguous structural assignments for these species (Fig. 3). This work is one of the most recent proofs confirming that the silicon clusters which have been predicted one way or another by theory also exist in the gas phase.

In the case of the neutral clusters up to 13 atoms, the most stable structures are not planar except Si_3 and Si_4 . Mainly they prefer polyhedral structures of rather high symmetry. Raghavachari and Logovinsky (1985), Fournier et al. (1992), Li (2005), and Yu et al. (2002) have established the basic structural patterns of those small species which nowadays are considered as systems of reference. For those species, Zhu and Zeng (2003) provided accurate geometries computed at high levels of theory. A recent study performed by Zdetsis (2007a, b) have showed that silicon clusters of 5 up to 13 atoms, and their dianions follow common structural

Fig. 3 Ground-state structures of neutral and charged silicon clusters. Stable structures of neutral silicon clusters



motifs of the corresponding closo-boranes. Among the small clusters with 2 up to 13 atoms there are two species of remarkable interest. These species are the Si_6 and Si_{10} which are considered as magic clusters because of their high stability. The six-atomic silicon cluster has attracted the attention of many researchers due the existence of three almost isoenergetic structures that compete for the ground state. Namely, the distorted (compressed) octahedron of D_{4h} symmetry and two trigonal bipyramidal shapes of lower C_{2v} symmetry, the edge-capped and face-capped trigonal bipyramids. On the other hand Si_{10} is also known as an extremely stable entity and it has been used as a building module of larger structures (Raghavachari and Rohlffing 1988; Mitas et al. 2000; Rohlffing and Raghavachari 1990; Grossman et al. 1995).

From $n = 14$ the shapes of the most stable structures become prolate (Zhu et al. 2004; Yoo and Zeng 2005). An interesting structural feature of the stacked prolate structures is that they are built from blocks that in fact are smaller stable clusters such as the Si_9 tricapped trigonal prism or the Si_6 puckered-hexagonal-ring. For instance Si_{11} and Si_{15} are built upon the tricapped-trigonal-prism of Si_9 while Si_6 puckered-hexagonal-ring motifs can be identified for silicon clusters with 16–20 atoms. This is the most reliable explanation of their stability since it is related to the stability of their building blocks. This finding is of crucial importance since it may open new directions in the nanomaterial and single nanoparticle design.

After $n = 19$, hollow near-spherical shapes start to show significant stability. From $n = 24$, the near-spherical structure with interior atoms becomes nearly degenerated with the prolate configurations (Jarrold and Bower 1992; Jarrold and Constant 1991; Kaxiras and Jackson 1993). However, up to $n = 25$, the stacked prolate structures are still the most stable. Finally, although clusters with more than 25 atoms are not very well studied and their ground-state structures are either unknown or under debate (Bai et al. 2006), there are recent studies that yielded very interesting results. For instance, Yoo and coworkers (Avramov et al. 2007; Yoo et al. 2008, 2004; Yoo and Zeng 2006, 2005) searched for generic structural features as well as patterns of structural evolution for the low-lying silicon clusters up to the size of 80 atoms by studying the structures and relative stability of different cluster shapes families. Their results indicate that for some cluster sizes such as Si_{39} , Si_{40} , and Si_{50} , the fullerene cage motifs consistently are more stable than other kind of cage structures.

The Si_6 and Si_{36} Cases

To demonstrate the difficulty in studying the structures and properties of clusters even of small size even when are built from only one element, and the complexity in interpreting computational and experimental results of diverse sources we shall revisit two distinctive cases: One concerns the small Si_6 cluster and the other the larger Si_{36} . As mentioned, Si_6 is a particularly stable cluster and along with Si_{10} coincides with clusters found abundantly in the experiment. Figure 4 shows three optimized structures of Si_6 at the CCSD(T) level and the natural atomic charges

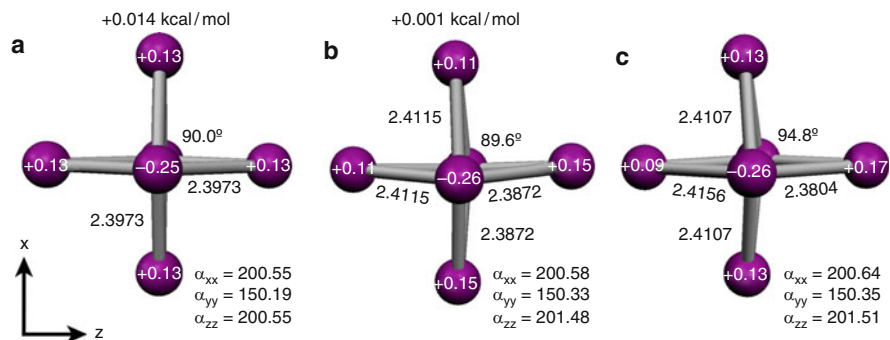


Fig. 4 CCSD(T)/6-311G(2d) optimized structures of the three isomers of Si₆ atomic charges after a natural bond orbital analysis at the CCD level and MP2/6-311G(2d) polarizabilities (au). All drawn bondlengths are in Å and dihedral angles in degrees. Structure (a) is of D_{4h} symmetry, structures (b) (face-capped) and (c) (edge-capped) of C_{2v} symmetry.

computed with the CCD density. It is obvious that although the two C_{2v} isomers keep the characteristics of their symmetry group as they were defined in the starting point geometries, they are almost identical with the distorted octahedron and their energy separations are extremely smaller than expected. The C_{2v} face-capped isomer and the distorted octahedron are lying only 0.001 and 0.014 kcal/mol higher in energy than the edge-capped isomer, respectively. Additionally, MP2 polarizability tensor values (au) at the CCSD(T) optimized geometries (also included in Fig. 2) show that the two C_{2v} structures are almost as polarizable as the D_{4h} structure in all three directions. Consequently, the two trigonal bipyramid structures could be viewed as slightly distorted tetragonal bipyramidal shapes belonging to the C_{2v} point group.

A careful literature search shows that these three structures compete for the ground state. More specifically, all electron and core potential geometry optimizations at Hartree–Fock (HF) level by Raghavachari and Rohlfing prefer an edge-capped trigonal bipyramidal shape of C_{2v} symmetry while MP2 geometry optimizations favor a distorted octahedron of D_{4h} symmetry. On the other hand, higher order correlated method based on Möller–Plesset perturbation theory with the standard 6-31G(d) basis set by Zdzetsis (2001) showed that the MP2 predicted structure of D_{4h} symmetry is unstable. Additionally, density functional theory (DFT) (Karamanis et al. 2007b) shows an fluctuating preference between the C_{2v} face-capped, C_{2v} edge-capped trigonal bipyramids. It is evident that there is an obvious disagreement between different theoretical approaches.

Let us now explore the most important experimental results concerning this problem. Raman experiments by Honea et al. (1993) of matrix isolated clusters, supported indirectly the MP2/6-31G(d) structure of D_{4h} symmetry. Their conclusions were based on two arguments: The first was the number of visible bands in both spectra which characterize species of D_{4h} symmetry, and second, the agreement between the experimental and scaled theoretical harmonic frequencies computed at the MP2/6-31G(d) level with the small 6-31G(d) standard basis

set. The strongest arguments used to reject the other two candidates were two: first, geometry optimizations at MP2 level starting from either face-capped or edge-capped trigonal bipyramidal shapes collapsed to the distorted octahedron of D_{4h} symmetry; second, the other two structures were predicted to have a larger number of active vibrational bands than observed in both experimental Raman and IR spectra. Obviously such a puzzling problem rarely has solely one answer. Indeed, the resolution to this puzzle is twofold. On one hand the source of the disagreement among the various methods used hides in the particular capabilities of each method in the treatment of the pseudo Jahn–Teller effect (PJTE) (Karamanis et al. 2007b). According to general Jahn–Teller theory the only source of instability in any molecular systems is the proper Jahn–Teller effect, for degenerated states (Renner Teller effect in linear systems), and pseudo-Jahn–Teller for non-degenerated ones. PJTE is a part of modern JT interaction theory (Bersuker 2001) and its importance in the formation of the equilibrium structures of clusters has been demonstrated by Garcia-Fernandez et al. (2006) and by Pushpa et al. (2004). In the Si_6 case, geometry optimizations with methods based on Møller–Plesset perturbation theory, which have been demonstrated not to treat the PJTE correctly, predict that the distorted octahedron of D_{4h} symmetry is the ground-state structure, while, methods which do provide PJTE treatment (DFT, HF) suggest that the distorted octahedron is unstable and undergoes PJTE distortions. On the other hand, the flat potential energy surface is connected with a fluxional behavior of Si_6 (Zdetsis 2007b) caused by the couplings between the electronic motion and soft nuclear vibrations due to the second-order Jahn–Teller effect under the influence of “electron deficiency” and the resulting charge smoothing process. The soft nuclear motion that is implied by the calculation is shown in Fig. 5. Of course, this kind of discrepancies in cluster structures can hardly be identified by experimental means since in such a flexible (or fluxional) system with low barriers, the consideration of separate “conformers” is probably not appropriate at all, and experimental spectroscopic properties are just an average of all structures. This has been highlighted in the recent work of Fielicke et al. (2009) who used tunable far-infrared-vacuum-ultraviolet two-color ionization to obtain vibrational spectra of neutral silicon clusters in the gas phase. Their results show clearly that it is almost impossible to distinguish the three different structures by means of vibrational spectroscopy.

Another example that demonstrates the complexity in determining the ground-state structures of larger clusters, and the importance of exploring the complete space when searching for atomic cluster is the case of Si_{36} (Bazterra et al. 2002; Sun et al. 2003). This cluster is one of the first large species that have been studied using first principles within the density functional framework (Sun et al. 2003). The structure that was initially proposed as the most stable was a stuffed fullerene-like configuration of near-spherical shape (Fig. 6 structure a). This structure seemed to confirm the up to that time experimental findings according to which a shape transition from prolate shapes to compact ones is occurring as the size of silicon clusters grows. Nonetheless, a subsequent study (Bazterra et al. 2002) based on a global approach using genetic algorithms, showed that a totally different structure lies lower in energy than the one previously adopted (Fig. 6 structure b). Once more,

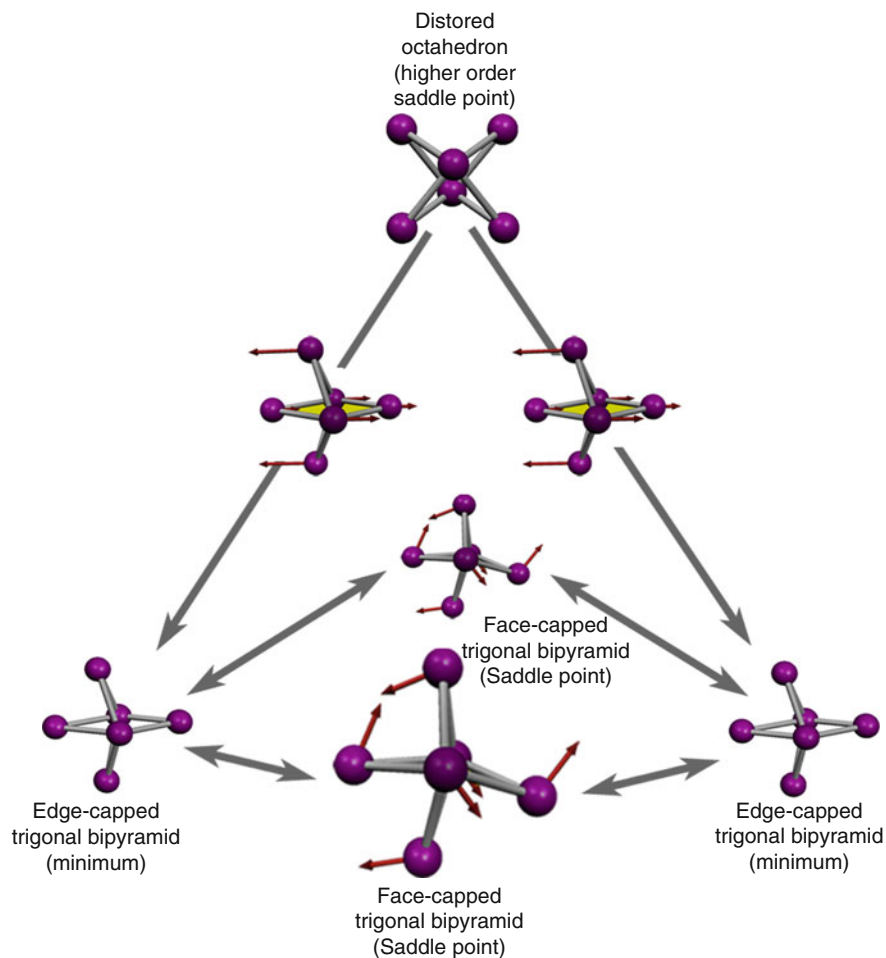


Fig. 5 Schematic representation of the three competing structures of Si_6 based on the normal modes of the imaginary frequencies at CCSD(T) level of theory with the split valence Pople-like Basis set 6-311G(2d)

the structure of Si_{36} proposed by Bazterra et al. can be envisioned as an assembly of smaller, stable clusters.

III–V and II–VI Semiconductor Clusters

General Features

The III–V and II–VI semiconductor clusters are isoelectronic with the elements of the group 13 and 15. For instance, gallium arsenide clusters of the type Ga_nAs_n have the same number of electrons with Ge_{2n} clusters while the analogue Al_nP_n clusters are isoelectronic with the corresponding Si_{2n} cluster species. For clusters

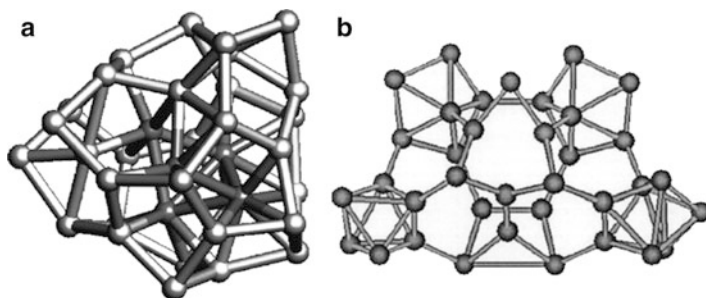


Fig. 6 Two competing structures of Si₃₆

of these two semiconductor families with even number of electrons, the preferred spin multiplicity is the singlet, where all electrons are in pairs. Triplets and larger spin multiplicities are always higher in energy (Gutsev et al. 2010). In the case of nonstoichiometric GaAs clusters which are species with even number of electrons, thus, open shell systems the preferred spin multiplicity is the doublet (Karamanis et al. 2009).

The basic bonding pattern in the compound III–V and II–VI semiconductor clusters is the alternating arrangement of their atoms (Figs. 7 and 8). The alternating A–B type of bonding is a result of the polarity of these entities which is owed to the electronegativity difference between their atoms. As a result, the bonding in clusters formed by atoms with large electronegativity differences is expected to be more ionic than in clusters built from atoms characterized by smaller differences. For instance, the spectroscopic electronegativity differences $\Delta\chi_{spec}$ converted to Pauling scaling between the As and the atoms Ga and In are 0.455 and 0.555, respectively. Accordingly, the bonding in III–V compound clusters built from As and Ga is expected to be of less ionic (or of more covalent character) than clusters built from In and As. However, although the alternating A–B type of bonding is generally observed for this class of semiconductor clusters due to the polarity of their bonding, there are some cases in which homo-atomic bonds are detected between the electronegative atoms. For example the ground-state structures of Ga₄As₄ (Fig. 9) and Al₅P₅ (Fig. 8) are characterized by As–As and P–P bonds. On the other hand, this bonding feature is not favorable for II–VI semiconductors which also do not favor endohedral structures.

Gallium Arsenide Clusters

Combined experimental and computational studies have led to very interesting and important conclusions concerning microscopic feature of those systems such as their bonding and electronic structures. For instance, the revealed even/odd alternation in the photoionization cross section obtained with an ArF excimer laser (6.4 eV) indicated that clusters of any composition composed by even numbers of atoms (e.g., Ga₁As₅, Ga₂As₄, Ga₃As₃, Ga₄As₂, Ga₅As₁) have singlet ground states, whereas odd numbered clusters of any composition (e.g., Ga₁As₅, Ga₁As₄, Ga₂,

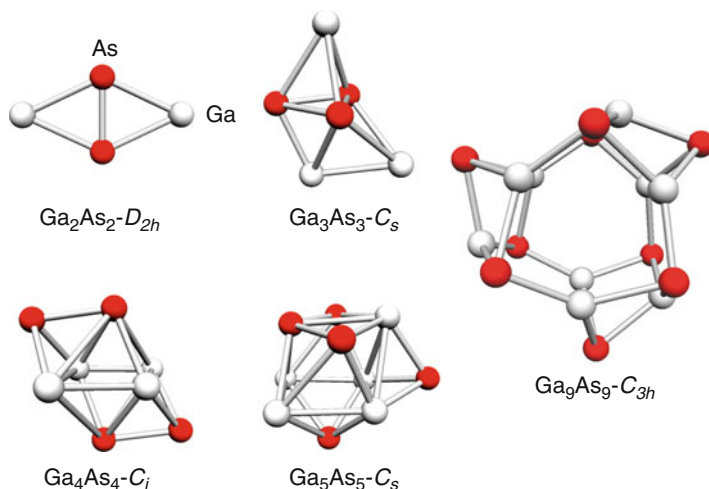


Fig. 7 Established ground-state structures of Ga_nAs_n clusters with $n = 2-5$ and 9. Such structure as for $n = 9$ is the most stable also in the cases of Al₉P₉ and Cd₉(S, Se, Te)₉

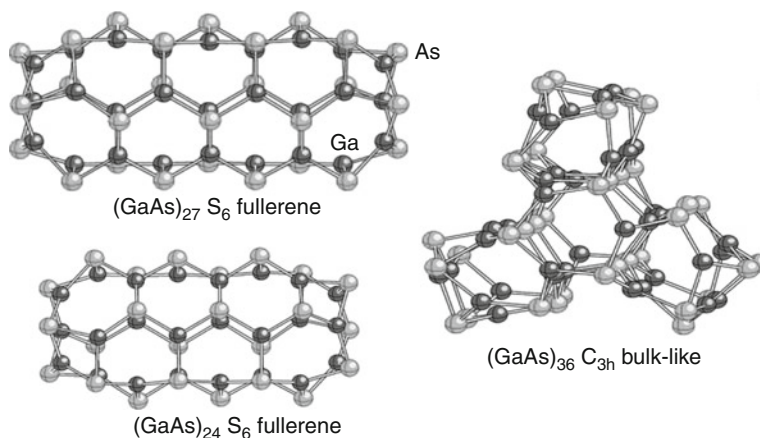


Fig. 8 Fullerene and bulk-like structures of large GaAs clusters (Gutsev et al. 2010)

As₃, Ga₃, As₂, Ga₄, As₁) have doublet ground states with an unpaired electron in a weakly bound HOMO. This sort of electronic structure clearly indicates that the structural and bonding features of the produced small GaAs clusters should be dramatically different from the tetrahedral sp³ hybridization of the bulk. Both of those indirect experimental conclusions—assumptions have been verified by all the later computational studies on those species which clearly showed that the singlet closed shell forms of the even-numbered GaAs clusters are the most stable than any other electronic configuration and their structures are completely different from the bulk (see: O'Brien et al. 1985; Lou et al. 1992; Karamanis et al. 2007a; Gutsev et al. 2008a, b).

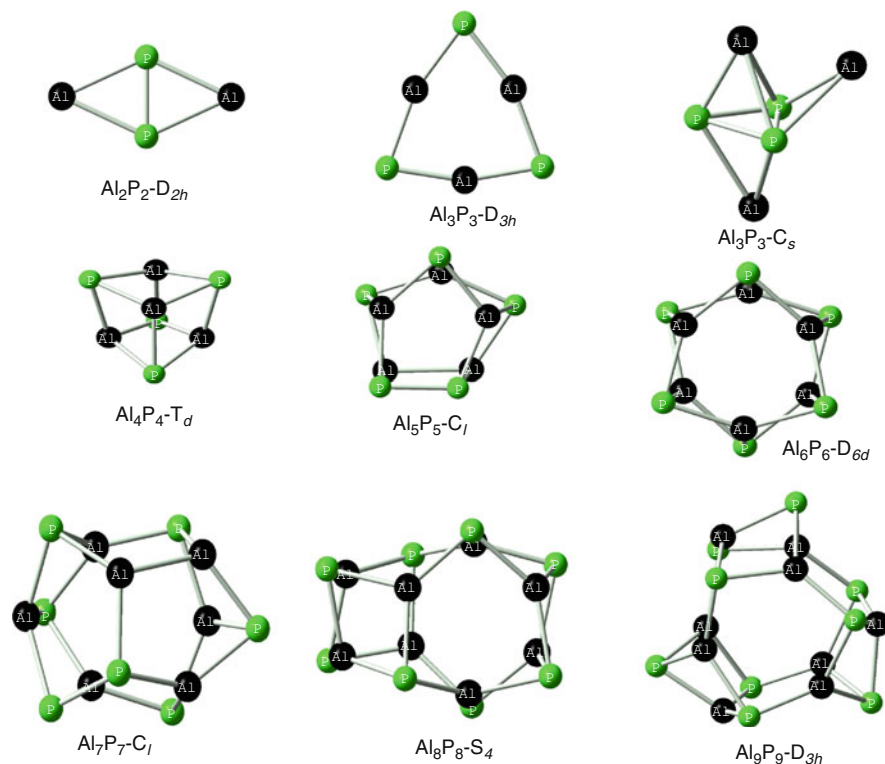


Fig. 9 Structures of aluminum phosphide clusters

Let us now make a brief review of the most representative studies of those important species starting from GaAs clusters which are the most studied species from those belonging to the III–V semiconductor family. Current and future applications of those species, in their nano or bulk, forms are very promising and interestingly diverse, extending from fuel cells Schaller and Klimov (2006) and optoelectronic sensors (Chen et al. 2004) to spintronics (spin-based electronics), medical diagnostics, and quantum computing (see Calarco et al. 2003; Michalet et al. 2005; Wolf et al. 2001).

Pioneering studies performed by O'Brien et al. on supersonic cluster beams of GaAs clusters generated by laser vaporization of discs made from pure GaAs brought theory closer to the experimental reality by providing cluster species of sizes suitable for very precise theoretical computations. As a result, the structural and bonding features of small GaAs clusters have occupied the interest of the researchers in the early cluster computational studies. For instance, Graves and Scuseria (1991) studied the first members of these systems using the Hartree–Fock (HF) method. Liao and Balasubramanian (1992) by means of the multireference configuration interaction method with singles and doubles (MRCISD) made a reference estimation of the bond lengths and angles of $(\text{GaAs})_2$; Lou et al. (1992)

studied the structures and the bonding of small GaAs clusters up to 10 atoms while Andreoni within the Car–Parinello molecular dynamics approach studied the stoichiometric systems up to the ten-atomic Ga_5As_5 . Later, Song et al. (1994) applied a fourth-order many body perturbation theory to study species from 2 to 4 atoms (Ga_nAs_n with $n = 2\text{--}4$). The interest in studying those systems remains intense even the recent years and a respectable variety of different methods have been applied. For instance, Costales et al. (2002) applied a DFT-GGA method for the systems up to the trimer, Zhao et al. (2006), Zhao and Cao (2001), Zhao et al. (2000) used another level of DFT-GGA approach for $(\text{GaAs})_n$ ($n = 6\text{--}9$) while Karamanis et al. (2007b) revisited the previous reported results using second-order many body perturbation theory (MP2) up to clusters with 18 atoms. More recently, Gutsev et al. (2008a, b) extended the earlier studies and reported structures up to 32 atoms with the DFT framework. In addition, there are some reported attempts in studying nonstoichiometric species and the influence of the composition on their properties. One of the most representative works in this field is the investigation of structural and electronic properties of Ga_mAs_n clusters reported by Feng et al. (2007); Gutsev et al. (2010) studied the electronic and geometrical structures of neutral $(\text{GaAs})_n$ clusters using density functional theory with generalized gradient approximation and relativistic effective core potentials for $n = 17\text{--}40$, 48, and 54. Also, they examined the energetic preference of tubular and non-tubular cages with increasing n and found that tubular cages are generally lower in total energy than the other cage isomers. Along with the cage structures, they probed structures cut off from a zinc blende lattice, tailoring the surface atoms in such a way as to have no dangling bonds. They found that the lowest energy state of $(\text{GaAs})_{36}$ has a bulk-like structure with 12 tetrahedrally coordinated inner atoms (Fig. 7). Finally, Gutsev and coworkers (Karamanis et al. 2011) studied the structure and properties of prolate $(\text{GaAs})_n$ clusters up to 120 atoms corresponding to the (2, 2) and (3, 3) armchair and (6, 0) zigzag capped single-wall tubes within the generalized gradient approximation (DFT-GGA). It was found that the bandgap in all three series does not converge to the GaAs bulk value when the cluster length increases.

II–VI Semiconductor Clusters

The synthesis of ZnO, ZnS, ZnSe, CdS, CdSe, and CdTe nanostructures which shape an important category of optically active materials with strong size dependence of their exciton energy (see: Alivisatos 1996; Swaminathan et al. 2006; Murray et al. 2000; Peng et al. 1998, 2000) have inspired most of the studies on II–VI semiconductor clusters. The respective nanomaterials can be synthesized in a great variety of sizes and shapes (spheres, rods, tetrapods or branched and core-shell structures) and they have been used for a wide assortment of applications, including quantum light emitting diodes pigments, biological tagging, and solar cells (Gur et al. 2005). Particularly, for CdSe clusters, correlations between their structure and properties have been highlighted by Jose et al. (2006).

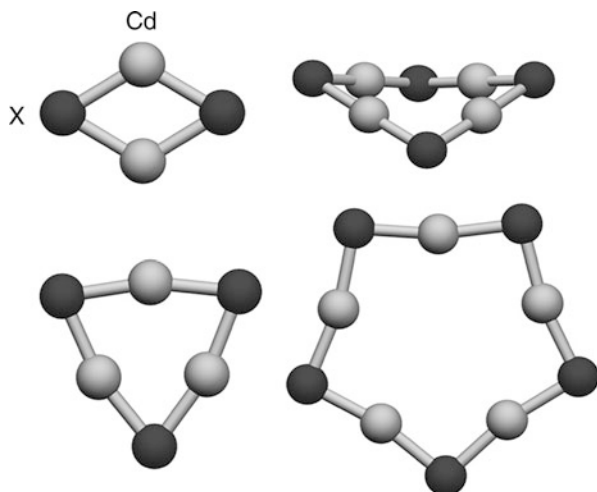
One of the most fascinating features of this class of semiconductor clusters is the size dependence of the HOMO–LUMO bandgap which is distinguishable in their luminescence properties. For instance, it has been shown that by adjusting the

particle size of colloidal CdSe–CdS core–shell nanoparticles the fluorescence can be tuned between blue and red. Motivated by those finding, in nanoparticles, Sanville et al. (2006) used direct laser ablation to produce ZnS, CdS, and CdSe clusters and found that clusters composed of 6 and 13 monomer units were ultrastable in all cases. Kasuya et al. (2004) by means of mass spectral analysis of CdSe nanoparticles found that $(\text{CdSe})_{33}$ and $(\text{CdSe})_{34}$ are so stable that they grow in preference to any other chemical compositions to produce macroscopic quantities with a simple solution method. First-principles calculations predict that these clusters are puckered $(\text{CdSe})_{28}$ -cages, with four- and six-membered rings based on the highly symmetric octahedral analogues of fullerenes, accommodating inside their framework either $(\text{CdSe})_5$ or $(\text{CdSe})_6$ to form a three-dimensional network with essentially heteropolar sp^3 -bonding.

Let us now see in brief the most cited computational studies for this very interesting class of cluster species. Behrman et al. (1994) have shown that ZnO clusters can form stable, fullerene-like structures while Hamad et al. (2005) provided evidences that ZnS clusters can form onion-like or alternatively “double bubble” structures. Matxain et al. (2000) studied systematically the structures of small ZnO clusters with up to nine atoms and demonstrated that three-dimensional structures may be envisioned as being built from Zn_2O_2 and Zn_3O_3 rings. Also, the same group (Matxain et al. 2001, 2003) have found that clusters built from ZnS, ZnSe, and CdO clusters can form stable one-dimensional ring structures (Fig. 10).

For species made from Cd and the chalcogenides S, Se Gurin(1998) studied small Cd_xS_y and Zn_xS_y ($x, y \leq 6$) clusters, Troparevsky and Chelikowsky(2001), Troparevsky et al. (2002) reported structures of $\text{Cdn}(\text{S, Se})_n$ with up to 8 atoms, Deglmann et al. (2002) used density functional theory potential surface investigations to study small CdS, and CdSe clusters, up to the heptamer while CdSe large clusters have been investigated further using large crystal structure sections

Fig. 10 Ring structure-types of MnX_n ($M = \text{Zn, Cd}$ and $X = \text{S, Se, Te}$) small clusters



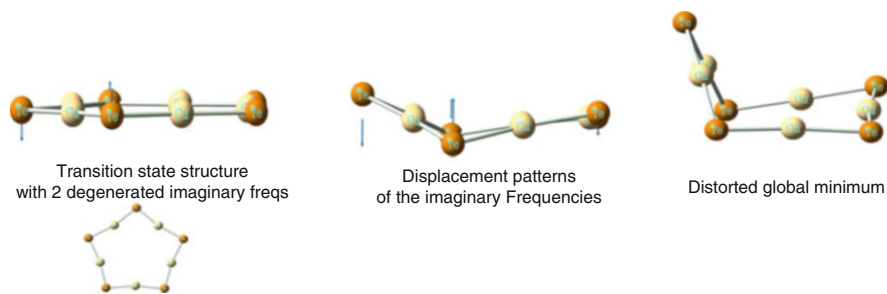


Fig. 11 Displacement patterns and ground-state structure of Cd_5Te_5

with up to 198 atoms. Also, Matxain et al. (2004) and Jha et al. (2008) reported ring and spherical structures with up to 16 atoms while Sanville et al. (2006) obtained geometries and energies of the neutral and positively charged M_nX_n clusters up to $n = 16$ ($\text{M} = \text{Zn}, \text{Cd}$ and $\text{X} = \text{S}, \text{Se}$). Their results showed that one-dimensional structures are favored for some cluster sizes while three-dimensional structures may be envisioned as being built from $\text{Cd}_2(\text{S}, \text{Se})_2$ and $\text{Cd}_3(\text{S}, \text{Se})_3$ rings. At this point it is important to stress that contrary to what is reported in several studies about CdX clusters, the five-atom rings are not totally planar. Instead, they maintain distorted envelope structures of cm^{-1} symmetry due to the weakening of the steric repulsion between the metal atoms as the size of the ring increases. As, a result the planar structure is a transition point in each cluster's potential energy surface characterized by soft imaginary frequencies frequencies of a few cm^{-1} which indicate a very floppy system that is expected to exhibit fluxional behavior (Fig. 11).

Lastly, small Cd_nTe_n and Hg_nTe_n clusters have attracted the attention of several groups (see for instance Wang et al. 2004, 2008, 2009). Especially HgTe clusters are very interesting since this material is considered as a semimetal with a small negative bandgap around $=0.14$ eV at 300 K and a semimetal-to-semiconductor can be produced by size quantization making HgTe nano-objects promising candidates as optical amplifiers (Fig. 12).

Electric Polarizabilities and Hyperpolarizabilities of Clusters

Definitions and Theory

The electric dipole (hyper)polarizabilities describe the response of a molecular system to an external electric field. Briefly, the polarizabilities (or linear polarizabilities) are used to express the ability of the electronic density of a molecular or cluster entity to distort under the influence of weak external fields such as the field generated by a charge in close distance. The hyperpolarizabilities (or nonlinear polarizabilities) describe the same response to laser beams. Large polarizabilities imply easily polarized molecules, whereas, small polarizabilities correspond to molecules in which the electronic density is considerably insensitive

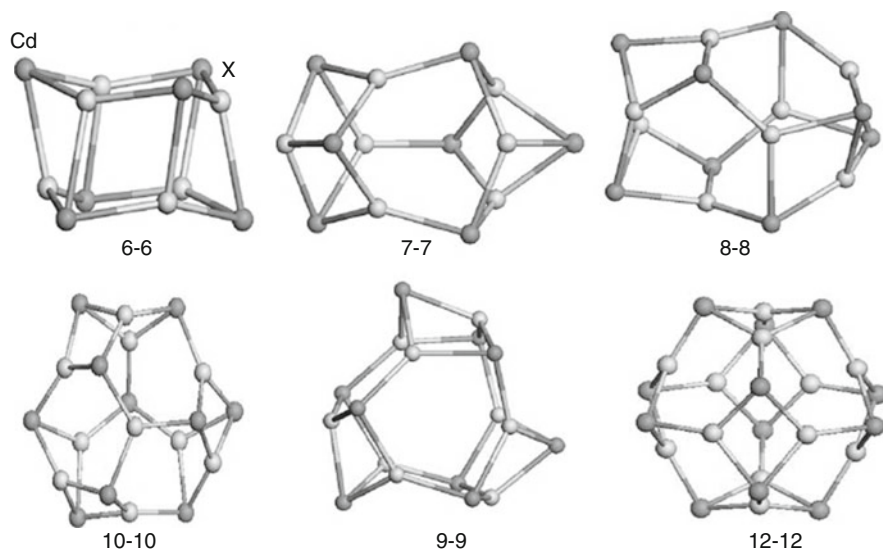


Fig. 12 Typical cage structures of Cd_nX_n ($X = S, Se, Te$) clusters

to electric field perturbations. Large hyperpolarizabilities involve molecules which can interact with intense electric fields in a “nonlinear” manner and the corresponding materials are known as nonlinear optical materials. The polarizabilities are associated with fundamental chemical concepts such as global softness (Vela and Gázquez 1990), basicity–acidity (Headley 1987), electronegativity (Nagle 1990), and stability (Hohm 2000), (Parr and Chattaraj 1991). On the other hand, the hyperpolarizabilities are closely related to the nonlinear optical behavior of materials in such processes as the second harmonic generation (SHG) (or frequency doubling), the third harmonic generation (THG), the sum frequency generation (SFG), the optical Kerr effect, and others which are expected to find application in a wide range of future technologies (see Bloembergen 1996 and references therein).

For clusters, polarizability is one of the few microscopic quantities that are available from experiment (Backer 1997) and is linked both to macroscopic and microscopic features of the matter such as the bulk dielectric constants and chemical bonding, respectively. Hence, this property can provide valuable information about crucial microscopic cluster features and the behavior of low dimensional systems that may be used in nanotechnology applications where nano-objects are subjected in small electric perturbations. On the other hand, given that the macroscopic nonlinear optical properties of materials are governed by the microscopic molecular hyperpolarizabilities, the study of those properties on clusters (Leitsmann et al. 2005) may offer a new ground in the search of new nonlinear optical materials with potential applications in future nanostructure technologies. The nonlinear behavior could be employed for amplification, modulation, and changing the frequency of

optical signals, similar to how nonlinearities of valves and transistors are used to govern the processes in traditional electric chips and circuits.

The microscopic (hyper)polarizabilities are studied by means of the so-called theory of the response functions which is of importance for all molecular and cluster entities (Roman et al. 2006). The most commonly used approach in studying the linear and nonlinear optical properties of clusters is the so-called semiclassical one. According to this approach a classical treatment is used to describe the response of the cluster to an external field (radiation) while the system itself is treated using the laws and techniques of quantum mechanics. This is done by using a Hamiltonian which combines both of the above treatments:

$$H = H^0 - \hat{\mu}_a F_a - \frac{1}{3} \hat{\Theta}_{\alpha\beta} F_{a\beta} - \frac{1}{15} \hat{\Omega}_{\alpha\beta\gamma} F_{\alpha\beta\gamma} - \frac{1}{105} \hat{\Theta}_{\alpha\beta\gamma\delta} F_{\alpha\beta\gamma\delta} + \dots \quad (1)$$

H^0 is the Hamiltonian which represents the free of the field system and the coefficients denoted with the Greek capital letters stand for the tensors of the dipole, quadrupole, octupole, and hexadecapole moments which are used to implement the field perturbation (F) into the new Hamiltonian. The expectation value after the implementation of this modified operator on the wavefunction is the perturbed energy of the system due to the influence of the field which in the case of a static field F can be written using a Taylor series expansion as

$$\begin{aligned} E^p = E^0 - \mu_a F_a - \frac{1}{3} \Theta_{\alpha\beta} F_{a\beta} - \frac{1}{15} \Omega_{\alpha\beta\gamma} F_{\alpha\beta\gamma} - \frac{1}{105} \Phi_{\alpha\beta\gamma\delta} F_{\alpha\beta\gamma\delta} + \dots & \frac{1}{2} a_{a\beta} F_a F_\beta \\ - \frac{1}{3} A_{\alpha,\beta\gamma} F_{a,\beta\gamma} - \frac{1}{6} C_{\alpha\beta,\gamma\delta} F_{a\beta,\gamma\delta} - \frac{1}{15} E_{\alpha,\beta\gamma\delta} F_{a,\beta\gamma\delta} + \dots & - \frac{1}{6} \beta_{\alpha\beta\gamma} F_a F_\beta F_\gamma \\ - \frac{1}{6} B_{\alpha\beta,\gamma\delta} F_a F_\beta F_{\gamma\delta} - \frac{1}{24} \gamma_{\alpha\beta\gamma\delta} F_a F_\beta F_\gamma F_\delta + \dots & \end{aligned} \quad (2)$$

E^p is the energy of the atomic or molecular system in the presence of the static electric field (F), E^0 is its energy in the absence of the field, μ_a corresponds to the permanent dipole moment of the system, $\alpha_{\alpha\beta}$ to the static dipole polarizability tensor and $\beta_{\alpha\beta\gamma}$, $\gamma_{\alpha\beta\gamma\delta}$ to the first and second dipole hyperpolarizabilities, respectively. Greek subscripts denote tensor components and can be equal to x, y, and z and each repeated subscript implies summation over x, y, and z as follows:

$$\frac{1}{2} a_{a\beta} F_a F_\beta = \frac{1}{2} \sum_{\alpha,\beta=x,y,z} a_{\alpha,\beta} F_\alpha F_\beta \quad (3)$$

In the case of a weak uniform (homogenous) field, the above expression becomes more simplified:

$$E^p = E^0 - \mu_\alpha F_\alpha - \frac{1}{2}\alpha_{\alpha\beta} F_\alpha F_\beta - \frac{1}{6}\beta_{\alpha\beta\gamma} F_\alpha F_\beta F_\gamma - \frac{1}{24}\gamma_{\alpha\beta\gamma\delta} F_\alpha F_\beta F_\gamma F_\delta + \dots \quad (4)$$

The properties that are routinely computed and discussed are the mean (or average) static dipole polarizability ($\bar{\alpha}$), the anisotropy ($\Delta\alpha$) of the polarizability tensor, the vector component of the first hyperpolarizability tensor in the direction of the ground state permanent dipole moment ($\beta_{||}$), the total first-order hyperpolarizability β_{tot} , and the scalar component of the second hyperpolarizability tensor $\bar{\gamma}$. Those quantities are related to the experiment and in terms of the Cartesian components are defined as

$$\bar{\alpha} = \frac{1}{3} (\alpha_{xx} + \alpha_{yy} + \alpha_{zz}) \quad (5)$$

$$\Delta\alpha = \left(\frac{1}{2}\right)^{1/2} \left[(\alpha_{xx} - \alpha_{yy})^2 + (\alpha_{xx} - \alpha_{zz})^2 + (\alpha_{zz} - \alpha_{yy})^2 + 6(\alpha_{xy}^2 + \alpha_{xz}^2 + \alpha_{zy}^2) \right]^{1/2} \quad (6)$$

$$\beta_{||} = \frac{3/5 (\beta_x \mu_x + \beta_y \mu_y + \beta_z \mu_z)}{|\mu|} \quad (7)$$

where $\beta_x = \beta_{xxx} + \beta_{xyy} + \beta_{xzz}$
 $\beta_y = \beta_{yxx} + \beta_{yyy} + \beta_{yzz}$
 $\beta_z = \beta_{zxx} + \beta_{zyy} + \beta_{zzz}$

$$\beta_{tot} = \left(\beta_x^2 + \beta_y^2 + \beta_z^2\right)^{1/2} \quad (8)$$

$$\bar{\gamma} = 1/5 (\gamma_{xxxx} + \gamma_{yyyy} + \gamma_{zzzz} + 2\gamma_{xxyy} + 2\gamma_{yyzz} + 2\gamma_{xxzz}) \quad (9)$$

Cluster (Hyper)Polarizabilities: Computational Approach

The majority of the undergoing theoretical (hyper)polarizability studies on semiconductor clusters rely on conventional ab-initio and density functional methods (DFT) As it is broadly accepted the ab initio approach leads to accurate (hyper)polarizability predictions and most of the reported attempts rely on a hierarchy of ab initio methods of increasing predictive capability such as the Hartree–Fock approximation (HF), the Møller–Plesset (MP) many body perturbation theory, and

also on the coupled cluster theory such as singles and doubles coupled-cluster (CCSD), and singles and doubles coupled-cluster with an estimate of connected triple excitations via a perturbational treatment (CCSD(T)) (further reading for those quantum chemical methods can be found in Szabo and Ostlun(1989), and Helgaker et al. (2000). Although this computational strategy is not always straightforward and in most of the cases computationally demanding, leads to reliable predictions of cluster (hyper)polarizabilities giving also crucial and detailed information about the electron correlation effects on the properties of interest. On the other hand the density functional approach (Koch and Holthausen 2000) is by far less computationally costly and allows the treatment of large systems; however, the reliability of DFT results rely heavily on the ability of the functional applied each time to describe the response of the system to an external electric field (see Karamanis et al. 2011 and references therein)

Looking carefully at Eq. 4 it becomes obvious that the computation of polarizabilities and hyperpolarizabilities in the case of static fields is in fact a mathematical derivation issue of the cluster energy with respect to the applied field. Consequently, the best accuracy can be achieved by finding and using the analytic expressions of the following relations:

$$-\left(\frac{\partial^2 E}{\partial F^2}\right) = \alpha_{\text{static}, F=0} \quad (10)$$

$$-\left(\frac{\partial^3 E}{\partial F^3}\right) = \beta_{\text{static}, F=0} \quad (11)$$

$$-\left(\frac{\partial^4 E}{\partial F^4}\right) = \gamma_{\text{static}, F=0} \quad (12)$$

In the case of the linear polarizability, the majority of the available quantum chemical software is able to evaluate analytically the first field derivative of the energy yielding in a straightforward manner all components of the polarizability tensor. In most of the cases the type of the calculation has to be indicated in the command line of the input file and the polarizabilities of the system are returned after the calculation ends. Unfortunately no analytical evaluation of the high-order field derivatives of the energy is available in all commercial packages for all methods. What is more, for methods of high accuracy such as the CC methods, this option is not available even for the linear polarizabilities. To overcome this problem several numerical procedures have been developed. The most applied numerical solution to this mathematical problem is the one provided by the so called finite field approach. According to this approach, first a field perturbation of different magnitudes and along different directions is added to the Hamiltonian. Then the wavefunction is “solved” for each perturbation and the energy of the perturbed system is obtained. Lastly the differentiation of energy of the system with respect to the applied field is performed numerically and the coefficients of the Eq. 4 are obtained (Figs. 13 and 14).

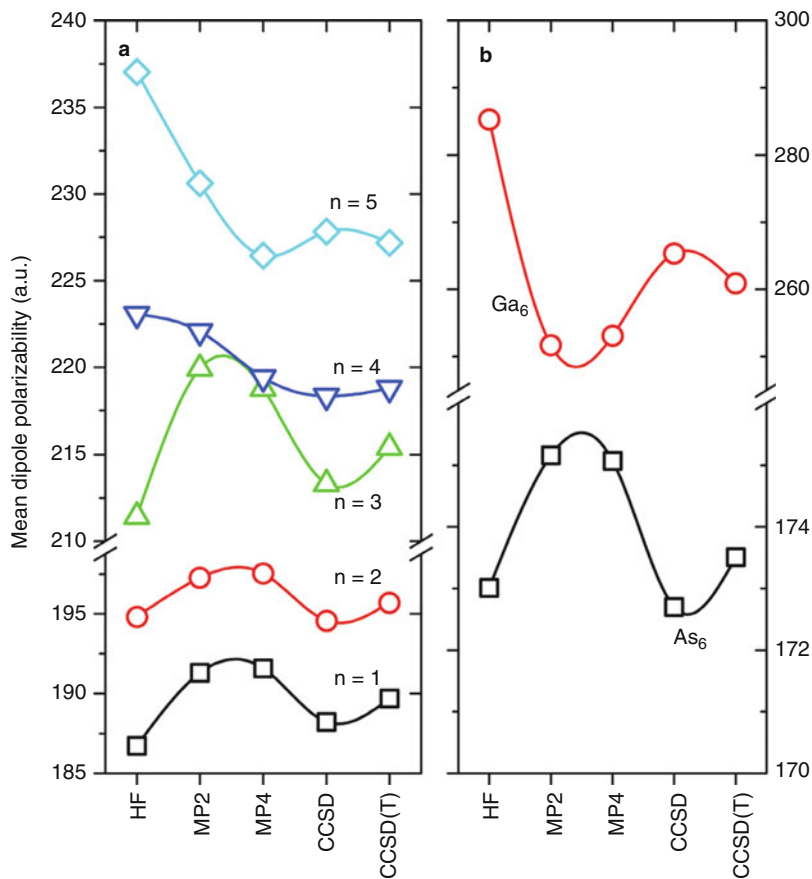


Fig. 13 Method performance on the computation of the mean polarizabilities of Ga_nAs_m clusters with $m + n = 6$ in comparison with the performance on the computation of the mean polarizabilities of Ga_6 and As_6 clusters

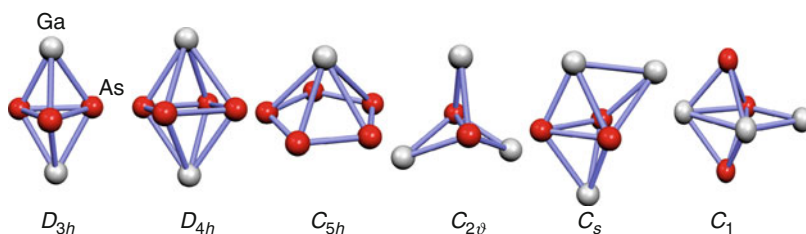


Fig. 14 Gallium arsenic clusters of various symmetries

Let us now demonstrate a simple example of the above process which will show us how the Eq. 4 can be utilized in order to calculate all the components of the polarizability tensor for a cluster without symmetry (C_1 point group). For

this symmetry group there are three independent components of the dipole moment (μ_x , μ_y and μ_z) and the six for the dipole polarizability: three principal (axial) components α_{xx} , α_{yy} , and α_{zz} and three transversal ones $\alpha_{xy} = \alpha_{yx}$, $\alpha_{zy} = \alpha_{yz}$, and $\alpha_{xz} = \alpha_{zx}$. The three principal ones can be computed by applying a weak electric field (of 0.0005–0.005 au strength) along each Cartesian direction, as follows:

$$\mu_a \approx \frac{E(-F_a) - E(F_a)}{2F_a} \quad (13)$$

$$\alpha_{aa} \approx \frac{2E(0) - E(-F_a) - E(F_a)}{F_a^2} \quad (14)$$

In the same manner, the transversal components can be obtained by applying fields in each $\alpha\beta$ plane along each $\pi/4$ diagonal between x , y , and z axis (this is done by applying fields of equal strengths ($F_\alpha = F_\beta$) at each xy , yz and xz planes). In this case one obtains the following equation:

$$\alpha_{a\beta} \approx \frac{E(F_a) + E(F_\beta) - E(F_a, F_\beta) - E(0)}{F_a F_\beta} \quad (15)$$

In the case of centrosymmetric structures, the computation of the dipole polarizability becomes more straightforward since the permanent dipole moment and the transversal components of dipole polarizability of those systems are vanishing. In this case we have

$$\alpha_{aa} \approx \frac{E(0) - E(F_a)}{F_a^2} \quad (16)$$

Provided that a suitable field is chosen, the above computational approach leads to quite reliable polarizability values. The advantage of this approach is that it gives the flexibility to the user to compute the polarizabilities of a given system using a variety of post Hartree–Fock methods only by computing the energies of the free clusters with and without the field.

The described approach can be applied to other response properties apart from the perturbed energies provided that the method used for their computations satisfy the Hellmann–Feynman theorem (Feynman 1939; Hellmann 1937) according to which the derivative of the total energy with respect to the field is equivalent with the expectation value of the derivative of the Hamiltonian with respect to the field. In this case, one can safely use the following relationships:

$$-\left(\frac{\partial^4 E}{\partial F^4}\right) = \left(\frac{\partial^3 \mu}{\partial F^3}\right) = \left(\frac{\partial^2 \alpha}{\partial F^2}\right) = \left(\frac{\partial \beta}{\partial F}\right) = \gamma_{static, F=0} \quad (17)$$

$$-\left(\frac{\partial^3 E}{\partial F^3}\right) = \left(\frac{\partial^2 \mu}{\partial F^2}\right) = \left(\frac{\partial \alpha}{\partial F}\right) = \beta_{\text{static}, F=0} \quad (18)$$

$$-\left(\frac{\partial^2 E}{\partial F^2}\right) = \left(\frac{\partial \mu}{\partial F}\right) = a_{\text{static}, F=0} \quad (19)$$

These relations can be used in all cases where the approximation to the true wavefunction is variationally optimized with respect to the Hamiltonian. This holds in the case of the SCF approach and all the DFT methods in which the energy of the systems is obtained through an iterative process. Contrary, in the case of the finite-order Møller–Pleset perturbation theory, (MP2, MP3, MP4), which is not variational, the relations Eqs. 17–19 do not hold and the calculations of the properties of interest should be extracted using only the energy derivatives.

Another key point in (hyper)polarizability calculations of clusters, as in all molecular systems, is the basis set choice. This has been clearly demonstrated in previous basis set studies on semiconductor clusters (see for instance Karamanis et al. 2006a, b, 2007a, 2008a; Maroulis et al. 2003, 2007; Papadopoulos et al. 2005, 2006). In general the polarizabilities are less sensitive to the basis set used than the first and second hyperpolarizabilities (see Fig. 15). Systematic basis set studies on cluster polarizabilities using basis sets of increasing size emphasize the importance of the diffuse s p Gaussian type functions (GTFs) and the effect of GTFs of polarization functions and functions of higher angular momentum such as ffunctions. Also it has been illustrated that as the cluster size increases the overall basis set effect on the mean polarizability gradually decreases (Karamanis et al. 2011).

Besides the methods and techniques described above there are also some theoretical methods that have been developed to treat the polarizabilities of clusters in a different manner from that described above. One of those schemes has been used by Jackson et al. (2007) and relies on partitioning the volume of a system into atomic volumes and using the charge distributions within the atomic volumes to define quantities such as atomic dipole moments and polarizabilities. In a second step each of the latter quantities are further decomposed into two components: the dipole (or local) and the charge-transfer component. The summation of each component over all of the atoms gives the local and charge-transfer parts of the total system dipole moment and polarizability. This treatment allows one to perform a site-specific analysis of dielectric properties of finite systems and, also, to obtain an insight about the contribution of each atom or of any desired building block of the cluster. The specific method has been applied to silicon clusters with up to 28 atoms which were subjected to a small external electric field. The obtained results showed clearly a strong dependence of the individual atomic polarizabilities, as well as their dipole and charge-transfer components, on the site location of the atom within the cluster. For atoms at peripheral sites those properties are substantially larger than for

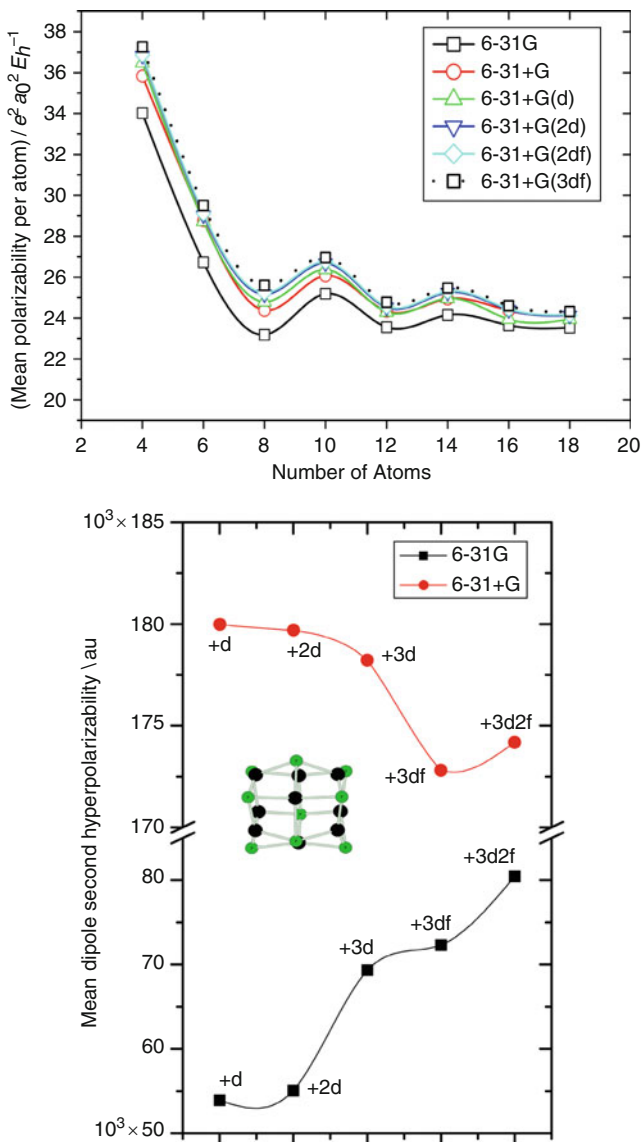


Fig. 15 Basis set effect, at the Hartree–Fock level of theory, on the mean dipole polarizabilities per atom of the ground state structures of aluminum phosphide cluster of the type Al_nP_n with $n = 2–9$, and basis set effect on the mean second hyperpolarizability of Al_9P_9

atoms at the interior sites, and the more peripheral atoms are characterized by larger polarizabilities. The same conclusion was drawn later by Krishtal et al. (2010).

An alternative methodology which uses electrostatic interaction schemes to predict the evolution of cluster polarizabilities has been developed and tested on

silicon clusters by Guillaume et al. (2009). This approach accepts the electrostatic nature of the dominant intra- and intermolecular interactions of a bulk material which is under the influence of external electric fields. This bottom-up approach allows the treatment very large clusters, enabling to reach the convergence of the response of the bulk material up to a point where adding new cluster units in the system would not change significantly the response of the material.

Further reading about the computational aspects of (hyper)polarizabilities on clusters and other molecules can be found in the following: Buckingham(1967), McLean and Yoshimine(1967), Marks and Ratner(1995), Kanis et al. (1994), Brédas et al. (1994), Bishop et al. (1997), Xenides(2006), Xenides and Maroulis(2000, 2006), Champagne et al. (2002), Maroulis(2008, 2004, 2003), Maroulis et al. (2007, 2003), Pouchan et al. (2004), Maroulis and Pouchan (2003), Papadopoulos and coworkers (Avramopoulos et al. 2004; Raptis et al. 1999; Reis et al. 2000); Karamanis and coworkers (Hohm et al. 2000; Kurtz et al. 1990; Luis et al. 2000).

Semiconductor Cluster (Hyper)polarizabilities: General Trends and Selected Studies

General Trends

Semiconductor cluster polarizabilities have been the subject of some very important experimental studies via beam-deflection techniques (Backer 1997; Schlecht et al. 1995; Schnell et al. 2003; Schäfer et al. 1996; Kim et al. 2005) while they have been extensively studied using quantum chemical and density functional theory. In this research realm, one of the areas intensively discussed is the evolution of the cluster's polarizabilities per atom (PPA) with the cluster size. The PPA is obtained by dividing the mean polarizability of a given system by the number of its atoms. Such property offers a straightforward tool to compare the microscopic polarizability of a given cluster with the polarizability of the bulk (see Fig. 16) as the latter is obtained by the "hard sphere" model with the bulk dielectric constant via the Clausius–Mossotti relation :

$$\bar{\alpha} = \frac{3}{4\pi} \left(\frac{\varepsilon - 1}{\varepsilon + 2} \right) v_{at} \quad (20)$$

where v_{at} is the volume per atom in the unit cell of the bulk and ε is the static dielectric constant of the bulk.

Most of the obtained knowledge on the polarizabilities of semiconductor clusters has emerged as the result of extensive studies of such species as silicon and gallium arsenide which have served well as models for the better understanding of the (hyper)polarizability evolution of semiconductor clusters. Also other kinds of clusters built from different elements such as aluminum phosphide, cadmium sulphide, selenide, and telluride have attracted the attention of several researchers and the corresponding studies have led to interesting conclusions. Those studies suggest that cluster polarizabilities are very different from the known polarizabilities

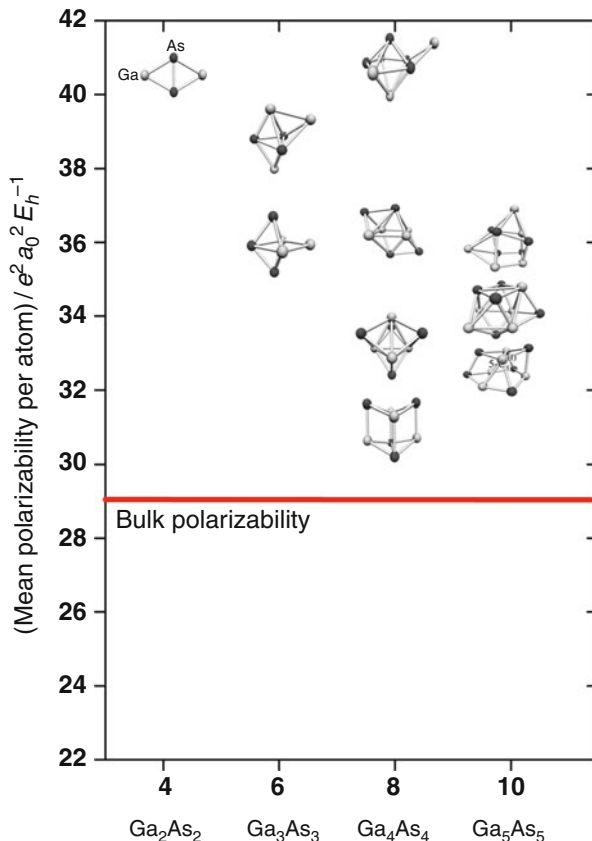


Fig. 16 Comparison between computed polarizabilities of small GaAs clusters and the polarizability of the bulk. The polarizabilities have been computed at the MP2 level of theory

of bulk materials. In general the PPAs of small semiconductor cluster are larger than the polarizability of the bulk (Fig. 16). For instance, in all studied cases the theoretical predicted polarizabilities of small semiconductor clusters (e.g., Si, AlP, CdSe, CdS, GaAs) are found higher than the bulk polarizability and converge to the bulk limit from above. Also, the polarizability dependence of the cluster size is strong and as a rule cluster species exhibit strong and nonmonotonic size variations. In addition there is a strong dependence of the polarizability on the shape of the cluster. It has been shown that for species which maintain closed cage-like structures (i.e., close to the spherical shape), the polarizability per atom (PPA) decreases with clusters size (Fig. 15) (see for instance (Jackson et al. 2004, 2005; Karamanis et al. 2008b, c; Zhang et al. 2004). On the other hand the per atom polarizabilities of clusters which maintain prolate structures (e.g., tubular clusters) increase as a function of the cluster size (Jackson et al. 2005; Karamanis et al. 2008b, 2011). Another important issue is the dependence of the electric polarizabilities

on the cluster bonding. It has been shown that semiconductor clusters which favor structures characterized by covalent bonds are expected to exhibit larger polarizabilities than clusters which tend to form ionic structures. This dependence is significantly more intense in the case of the second hyperpolarizability. The general rule that holds in this case is that semiconductor clusters which are characterized by significant electron transfer from the electropositive atoms (ionic bonding) to the electronegative ones are less hyperpolarizable than species in which the electron transfer is smaller (covalent bonding) (Karamanis et al. 2008d; Krishtal et al. 2010). This variation explains the nonmonotonic size variations of the mean PPAs of small binary semiconductor clusters. Also, both polarizabilities and hyperpolarizabilities of clusters have been proven very sensitive to the cluster elemental constitution. For instance, InAs clusters are more (hyper)polarizable than GaAs, AlP, AlAs, InP, and GaP clusters while CdTe clusters, are far more (hyper)polarizable than CdS and CdSe clusters (Karamanis and Pouchan 2009). What is more, the strong correlation between the polarizabilities of semiconductor clusters and their composition is also significant. In the case of GaAs clusters it has been demonstrated both experimentally and theoretically that Ga-rich gallium arsenide clusters are more polarizable than As-rich clusters due to the larger polarizability of Ga (Fig. 17) (Karamanis et al. 2009). Lastly, it was shown that the cluster shape dominates the magnitude of the second hyperpolarizabilities of large clusters (Karamanis and Pouchan 2011).

Selected Studies

Polarizabilities

Let us now briefly explore some of the most representative studies in the subject of the polarizabilities of Si and GaAs clusters that have attracted considerable attention. One of the first studies has been reported by Vasiliev et al. (1997). In

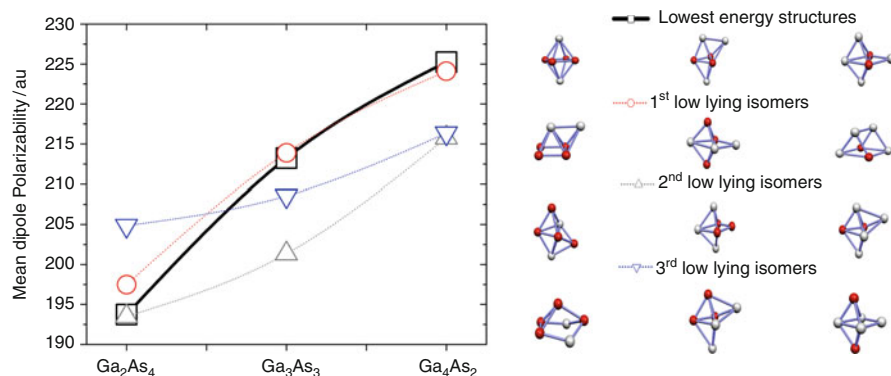


Fig. 17 Composition-dependent polarizabilities of small Ga_nAs_m with clusters $n + m = 6$ computed at the B2PLYP/aug-cc-pVDZ-PP level

that work the authors computed the polarizabilities of small Si_n , Ge_n ($n \leq 10$) and Ga_nAs_m ($n + m \leq 8$) clusters and they found that the polarizabilities of all systems are higher than the values estimated from the “hard sphere” model using Eq. 20. Also they evidenced that the computed per atom polarizabilities tend to decrease with increasing the cluster size and they related this trend to the one observed in the case of metallic clusters, for which the polarizability bulk limit is approached from above. This observation was considered as an evidence for the “metallic-like” nature of small semiconductor clusters. Two similar studies about silicon clusters up to 20 atoms by Jackson et al. (1999) and up to 28 atoms by Deng et al. (2000) confirmed the reported trends regarding to the PPA evolution with size and the metallic nature of silicon semiconductor clusters. The same conclusion has been reported also in later studies by Bazterra et al. (2002) and Maroulis et al. (2003). In the case of GaAs clusters the early results reported by Vasiliev et al. (1997) have been confirmed by Torrens(2002) with an interacting-induced-dipoles polarization model and have been extended up to 18 atoms by Karamanis et al. (2007a, 2008). Furthermore, the polarizability ordering in the case of the III–V dimers has been established: $\text{In}_2\text{As}_2 > \text{In}_2\text{P}_2 > \text{Al}_2\text{As}_2 > \text{Ga}_2\text{As}_2 > \text{Al}_2\text{P}_2 > \text{Ga}_2\text{P}_2$ (Karamanis et al. 2008a).

Hyperpolarizabilities

In the domain of the nonlinear polarizabilities of semiconductor clusters (i.e., the second dipole hyperpolarizability and third polarizability) the reported studies are fewer than in the case of the linear polarizabilities. However, this field is very promising due to the properties that the compound semiconductors possess. With respect to their interaction with intense electric fields compound semiconductors are characterized by high sensitivity, high speed, and they may operate at a wide variety of wavelengths in the near-infrared range. In particular the compound III–V and II–VI semiconductors such as GaAs, InP, and CdS possess the two properties needed for the existence of the photorefractive effect, the photoconductivity, and the electro-optic effect. These properties are very crucial in electro-optical technology; and alloptical information technology, thus the study of clusters and nano-objects of those materials are essential, first for better understanding the properties of the bulk material and second for the research of new nano-materials with improved nonlinear optical properties (Bechstedt et al. 2007; Adolph and Bechstedt 1998; Hughes and Sipe 1996; Bergfeld and Daum 2003; Bloembergen 1996; Butcher and Cotter 1990; Powell et al. 2002; Vijayalakshmi et al. 2002).

At the microscopic level, one of the first studies into this matter was reported by Korambath and Karna(2000) who studied the first dipole hyperpolarizabilities of GaN, GaP, and GaAs clusters within the ab initio time-dependent Hartree–Fock method using an even-tempered Gaussian basis set and demonstrated that the magnitudes of the calculated (hyper)polarizabilities depend on the size of the cluster. Lan et al. (2003) studied the hyperpolarizabilities of Ga_3As_3 , Ga_3Sb_3 , In_3P_3 , and In_3Sb_3 clusters using a similar approach based on the time-dependent Hartree–Fock (TDHF) formalism combined with sum-over-states (SOS) method and found that the charge transferred from the π - bonding to π - antibonding orbitals between III and V group atoms under the influence of an electric field make a significant

contribution to the second-order polarizability. The first ab initio attempt in treating the hyperpolarizabilities of semiconductor using methods of high sophistication such as CC and MP has been reported by Maroulis and Pouchan who explored the second hyperpolarizabilities of cadmium sulfide clusters. Their results showed that the mean second hyperpolarizability per atom of those species decreases rapidly with the cluster size. A second attempt on systems of increasing size has been reported by Lan et al. (2006). These authors studied several small GaAs clusters up to 10 atoms within the time-dependent density functional theory and reported that the dynamic behavior of β and γ show that the small Ga_nAs_m clusters are expected as good candidates for nonlinear optical properties due to the avoidance of linear resonance photoabsorption. An application of the time-dependent DFT (TD-DFT) theory on the hyperpolarizabilities of clusters can be found in the work of Sen and Chakrabarti(2006) who studied the second hyperpolarizabilities of CdSe clusters. In their conclusions, it is highlighted that these clusters have a high nonlinear optical activity. Karamanis et al. (2007a) reported ab initio values at the HF, MP2 level of theory for stoichiometric GaAs clusters up to 18 atoms and showed that as the size of the cluster increases the per atom mean second hyperpolarizability decreases fast and converges to a specific value. Finally, the significance of the bonding on the nonlinear polarizabilities of small clusters has been demonstrated by Karamanis and Leszczynski(2008) in a hyperpolarizability study of stoichiometric aluminum phosphide clusters up to 18 atoms. The conclusions of investigations on aluminum phosphide clusters have been extended later to several types of III–V semiconductor in a comparative hyperpolarizability study by Karamanis et al. (2008a). In that study the relative mean static dipole hyperpolarizabilities of the X_2Y_2 type ($\text{X} = \text{Al}, \text{Ga}, \text{In}$ and $\text{Y} = \text{P}, \text{As}$) clusters have been established in several levels of theory, $\bar{\gamma} : \text{In}_2\text{As}_2 > \text{In}_2\text{P}_2 \cong \text{Al}_2\text{As}_2 > \text{Al}_2\text{P}_2 \cong \text{Ga}_2\text{As}_2 > \text{Ga}_2\text{P}_2$. This classification follows the size ordering of the electropositive atoms. Ga is slightly smaller than Al, while In is far more larger than both Ga and Al. The atomic radius of $\text{Ga}(3d^{10}4s^{24}p^1)$ is slightly smaller than the $\text{Al}(3s^23p^1)$ atomic radius, despite its larger number of electrons. This is caused by the insufficient shielding that the $3d^{10}$ electrons provide to the $4s^24p^1$ valence electrons of Ga. As a result, the valence electrons of Ga experience a larger attraction from the nuclear charge than the electrons of Al. Considering that the (hyper)polarizabilities are valence-related properties, the described peculiarity of the Ga atoms can explain qualitatively the smaller hyperpolarizabilities of both Ga_2P_2 and Ga_2As_2 compared to Al_2P_2 and Al_2As_2 clusters.

The dynamic-second order hyperpolarizabilities of Si_3 and Si_4 have been explored by Lan et al. (2008) using the coupled cluster cubic response theory. These authors reported that Si_3 and Si_4 clusters exhibit wide non-resonant optical region. The same approach has been used by Lan and Feng (2009) in the study of the nonlinear response of another class of semiconductor clusters, namely, the SiC_n and Si_nC . In this work it was demonstrated that the size dependence of the first-order hyperpolarizabilities of the SiC_n clusters, which have approximate Si-terminated linear chain geometry, is similar to that observed in π -conjugated organic molecules. Lastly concerning the (hyper)polarizabilities of other II–VI clusters apart

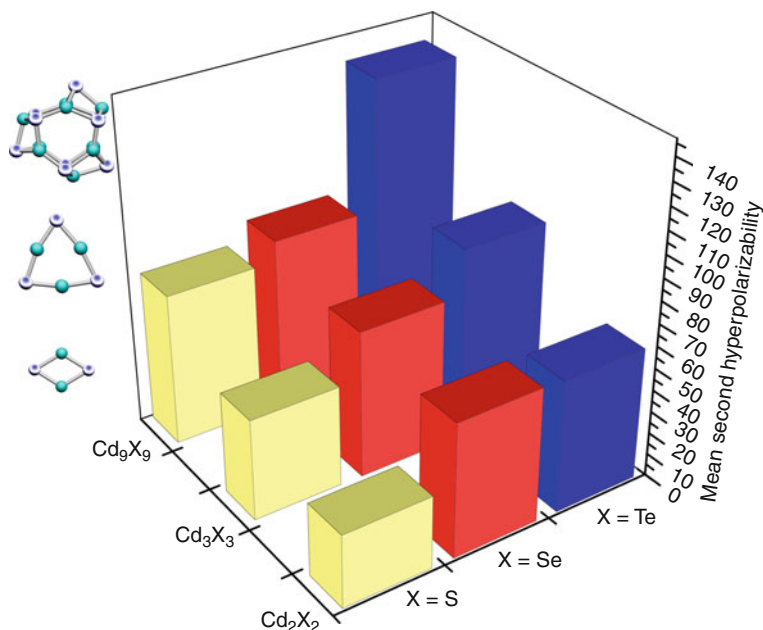


Fig. 18 Comparisons between the second-order mean hyperpolarizabilities (divided by 10^3) of CdS, CdSe, and CdTe clusters computed at the HF/aug-cc-pVDZ-PP level of theory

from CdS there two reported studies on their hyperpolarizabilities. The first (Li et al. 2008) deals with the ground states and low-lying isomers of stoichiometric $(\text{ZnO})_n$ clusters with $n = 2-12$ using finite field approach within density functional theory framework and the second (Karamanis et al. 2009) involves a comparative study of CdS, CdSe, CdTe with respect to their second hyperpolarizabilities (Fig. 18). Both of those studies demonstrate the structural and size dependence of those properties while the second one reveal that CdTe clusters are far more (hyper)polarizable than CdS and CdSe, verifying previous experimental work reported by Wu et al. (2003) which provided evidences that that CdTe nanoparticles exhibit stronger nonlinear optical properties than nanoparticles built from CdS and CdSe.

Concluding Remarks

In this brief review, we made an attempt to present some of the most illustrative computational investigations of the structural properties of semiconductor clusters. We have chosen to present studies that had as subject cluster species which have attracted most of the attention of the researchers the last two decades and have served as models in the better understanding of cluster properties. Therefore it must be emphasized that although this chapter does not cover the vast number of studies

that have been reported during the previous years, it can be utilized as starting material for those who are interested in contributing in this field, or simply, as a guide for those who wish to get some general information about the structural and bonding features of semiconductor clusters. Finally, in this chapter we have revisited most of the recent studies that focus on the electric properties of those species which also demonstrate in a comprehensive manner that clusters and electric properties such as the dipole polarizabilities and especially hyperpolarizabilities are strongly related.

Bibliography

- Adolph, B., & Bechstedt, F. (1998). Ab initio second-harmonic susceptibilities of semiconductors: generalized tetrahedron method and quasiparticle effects. *Physical Review B - Condensed Matter and Materials Physics*, 57(11), 6519–6526.
- Alivisatos, A. P. (1996). Semiconductor clusters, nanocrystals, and quantum dots. *Science*, 271(5251), 933–937.
- Al-Laham, M. A., & Raghavachari, K. (1991). Theoretical study of small gallium arsenide clusters. *Chemical Physics Letters*, 187(1–2), 13–20.
- Al-Laham, M. A., & Raghavachari, K. (1993). Theoretical study of Ga₄As₄, Al₄P₄, and Mg₄S₄ clusters. *Journal of Chemical Physics*, 98(11), 8770–8776.
- An, W., Gao, Y., Bulusu, S., & Zeng, X. (2005). Ab initio calculation of bowl, cage, and ring isomers of C₂₀ and C₂₀. *Journal of Chemical Physics*, 122(20), 4109/1–204109/8.
- Avramopoulos, A., Reis, H., Li, J., & Papadopoulos, M. G. (2004). The dipole moment, polarizabilities, and first hyperpolarizabilities of HARF. A computational and comparative study. *Journal of the American Chemical Society*, 126(19), 6179–6184.
- Avramov, P. V., Fedorov, D. G., Sorokin, P. B., Chernozatonskii, L. A., & Gordon, M. S. (2007). Atomic and electronic structure of new hollow-based symmetric families of silicon nanoclusters. *Journal of Physical Chemistry C*, 111(51), 18824–18830.
- Backer, J. A. (1997). Molecular beam studies on semiconductor clusters: polarizabilities and chemical bonding. *Angewandte Chemie International Edition in English*, 36(13–14), 1390–1404.
- Bai, J., Cui, L.-F., Wang, J., Yoo, S., Li, X., Jellinek, J., et al. (2006). Structural evolution of anionic silicon clusters Si_N (20 ≤ N ≤ 45). *Journal of Physical Chemistry A*, 110(3), 908–912.
- Bazterra, V. E., Caputo, M. C., Ferraro, M. B., & Fuentealba, P. (2002). On the theoretical determination of the static dipole polarizability of intermediate size silicon clusters. *Journal of Chemical Physics*, 117(24), 11158–11165.
- Bazterra, V. E., Oña, O., Caputo, M. C., Ferraro, M. B., Fuentealba, P., & Facelli, J. C. (2004). Modified genetic algorithms to model cluster structures in medium-size silicon clusters. *Physical Review A - Atomic, Molecular, and Optical Physics*, 69(5B), 053202/1–053202/7.
- Bechstedt, F., Adolph, B., & Schmidt, W. G. (2007). Ab initio calculation of linear and nonlinear optical properties of semiconductor structures. *Brazilian Journal of Physics*, 29(4), 643–651.
- Behrman, E. C., Foehrweiser, R. K., Myers, J. R., French, B. R., & Zandler, M. E. (1994). Possibility of stable spheroid molecules of ZnO. *Physical Review A*, 49(3), R1543–R1549.
- Bergfeld, S., & Daum, W. (2003). Second-harmonic generation in GaAs: experiment versus theoretical predictions of $\chi(2)_{xyz}$. *Physical Review Letters*, 90(3), 036801/1–036801/4.
- Bersuker, I. B. (2001). Modern aspects of the Jahn-Teller effect theory and applications to molecular problems. *Chemical Reviews*, 101(4), 1067–1114.
- Bishop, D. M., Kirtman, B., & Champagne, B. (1997). Differences between the exact sum-over-states and the canonical approximation for the calculation of static and dynamic hyperpolarizabilities. *Journal of Chemical Physics*, 107(15), 5780–5787.

- Biswas, R., & Hamann, D. R. (1986). Simulated annealing of silicon atom clusters in langevin molecular dynamics. *Physical Review B*, *34*(2), 895–901.
- Blaisten-Barojas, E., & Levesque, D. (1986). Molecular-dynamics simulation of silicon clusters. *Physical Review B*, *34*(6), 3910–3916.
- Bloembergen, N. (1996). *In Nonlinear optics* (4th ed.). Singapore: World Scientific.
- Brédas, J. L., Adant, C., Tackx, P., Persoons, A., & Pierce, B. M. (1994). Third-order nonlinear optical response in organic materials: theoretical and experimental aspects. *Chemical Reviews*, *94*(1), 243–278.
- Bruchez, M., Jr., Moronne, M., Gin, P., Weiss, S., & Alivisatos, A. P. (1998). Semiconductor nanocrystals as fluorescent biological labels. *Science*, *281*(5385), 2013–2016.
- Buckingham, A. D. (1967). Permanent and induced molecular moments and long-range intermolecular forces. *Advances in Chemical Physics*, *12*, 107–142.
- Butcher, P. N., & Cotter, D. (1990). *The elements Of nonlinear optics*. Cambridge: Cambridge University Press.
- Calarco, T., Datta, A., Fedichey, P., Pazy, E., & Zoller, P. (2003). Spin-based all-optical quantum computation with quantum dots: understanding and suppressing decoherence. *Physical Review A - Atomic, Molecular, and Optical Physics*, *68*(1), 012310/1–012310/21.
- Castro, A., Marques, M. A. L., Alonso, J. A., Bertsch, G. F., Yabana, K., & Rubio, A. (2002). Can optical spectroscopy directly elucidate the ground state of C₂₀? *Journal of Chemical Physics*, *116*(5), 1930–1933.
- Champagne, B., Spassova, M., Jadin, J.-B., & Kirtman, B. (2002). Ab initio investigation of doping-enhanced electronic and vibrational second hyperpolarizability of polyacetylene chains. *Journal of Chemical Physics*, *116*(9), 3935–3946.
- Chen, W., Zhang, J. Z., & Joly, A. G. (2004). Optical properties and potential applications of doped semiconductor nanoparticles. *Journal of Nanoscience and Nanotechnology*, *4*(8), 919–947.
- Costales, A., Kandalam, A. K., Franco, R., & Pandey, R. (2002). Theoretical study of structural and vibrational properties of (AlAs)_n, (GaP)_n, (GaAs)_n, (InP)_n, and (InAs)_n clusters with n = 1, 2, 3. *Journal of Physical Chemistry B*, *106*(8), 1940–1944.
- Deglmann, P., Ahlrichs, R., & Tsereteli, K. (2002). Theoretical studies of ligand-free cadmium selenide and related semiconductor clusters. *Journal of Chemical Physics*, *116*(4), 1585–1597.
- Deng, K., Yang, J., & Chan, C. T. (2000). Calculated polarizabilities of small S clusters. *Physical Review A - Atomic, Molecular, and Optical Physics*, *61*(2), 252011–252014.
- Dugourd, P., Hudgins, R. R., Tenenbaum, J. M., & Jarrold, M. F. (1998). Observation of new ring isomers for carbon cluster anions. *Physical Review Letters*, *80*(19), 4197–4200.
- Feng, Y. P., Boo, T. B., Kwong, H. H., Ong, C. K., Kumar, V., & Kawazoe, Y. (2007). Composition dependence of structural and electronic properties of Ga_mAs_n clusters from first principles. *Physical Review B - Condensed Matter and Materials Physics*, *76*(4), 045336/1–045336/8.
- Feynman, R. P. (1939). Forces in molecules. *Forces in Molecules. Physical Reviews*, *56*(4), 340.
- Fielicke, A., Lyon, J. T., Haertelt, M., Meijer, G., Claes, P., De Haeck, J., et al. (2009). Vibrational spectroscopy of neutral silicon clusters via far-IR-VUV two color ionization. *Journal of Chemical Physics*, *131*(17), 171105/1–171105/6.
- Fournier, R., Sinnott, S. B., & DePristo, A. E. (1992). Density functional study of the bonding in small silicon clusters. *Journal of Chemical Physics*, *97*(6), 4149–4161.
- Garcia-Fernandez, P., Bersuker, I. B., & Boggs, J. E. (2006). The origin of molecular distortions: a proposed experimental test. *Journal of Chemical Physics*, *124*(6), 044321/1–044321/7.
- Graves, R. M., & Scuseria, G. E. (1991). Ab initio theoretical study of small GaAs clusters. *Journal of Chemical Physics*, *95*(9), 6602–6606.
- Grimme, S., & Mück-Lichtenfeld, C. (2002). Structural isomers of C₂₀ revisited: the cage and bowl are almost isoenergetic. *ChemPhysChem*, *3*(2), 207–209.
- Grossman, J. C., Mitas, L., & Raghavachari, K. (1995). Structure and stability of molecular carbon: importance of electron correlation. *Physical Review Letters*, *75*(21), 3870–3873.
- Guillaume, M., Champagne, B., Bégú, D., & Pouchan, C. (2009). Electrostatic interaction schemes for evaluating the polarizability of silicon clusters. *Journal of Chemical Physics*, *130*(13), 134715.

- Gur, I., Fromer, N. A., Geier, M. L., & Alivisatos, A. P. (2005). Materials science: air-stable all-inorganic nanocrystal solar cells processed from solution. *Science*, *310*(5747), 462–465.
- Gurin, V. S. (1998). Ab-initio calculations of small Cd_xS_y and Zn_xS_y ($x, y \leq 6$) clusters. *Solid State Communications*, *108*(6), 389–392.
- Gutsev, G. L., O'Neal, R. H., Jr., Saha, B. C., Mochena, M. D., Johnson, E., & Bauschlicher, C. W., Jr. (2008a). Optical properties of $(GaAs)_n$ clusters ($n = 2-16$). *Journal of Physical Chemistry A*, *112*(43), 10728–10735.
- Gutsev, G. L., Johnson, E., Mochena, M. D., & Bauschlicher, C. W., Jr. (2008b). The structure and energetics of $(GaAs)_n$, $(GaAs)^-_n$, and $(GaAs)^+_n$ ($n = 2-15$). *Journal of Chemical Physics*, *128*(14), 144707/1–144707/9.
- Gutsev, G. L., Mochena, M. D., Saha, B. C., & Derosa, P. A. (2010). Structures and properties of $(GaAs)_n$ clusters. *Journal of Computational and Theoretical Nanoscience*, *7*, 1–10.
- Hamad, S., Richard, C., Catlow, A., Spanó, E., Matxain, J. M., & Ugalde, J. M. (2005). Structure and properties of ZnS nanoclusters. *Journal of Physical Chemistry B*, *109*(7), 2703–2709.
- Headley, A. D. (1987). Substituent effects on the basicity of dimethylamines. *Journal of the American Chemical Society*, *109*(8), 2347–2348.
- Helgaker, T., Jørgensen, P., & Olsen, J. (2000). *Molecular Electronic-Structure Theory*. Chichester: Wiley.
- Hellmann, H. (1937). *Einführung in die Quantenchemie* (p. 285). Leipzig: Franz Deuticke.
- Ho, K.-M., Shvartsburg, A. A., Pan, B., Lu, Z.-Y., Wang, C.-Z., Wacker, J. G., Fye, J. L., & Jarrold, M. F. (1998). Structures of medium-sized silicon clusters. *Nature*, *392*, 582–585.
- Hohm, U. (2000). Is there a minimum polarizability principle in chemical reactions? *Journal of Physical Chemistry A*, *104*(36), 8418–8423.
- Hohm, U., Loose, A., Maroulis, G., & Xenides, D. (2000). Combined experimental and theoretical treatment of the dipole polarizability of P_4 clusters. *Physical Review A - Atomic, Molecular, and Optical Physics*, *61*(5), 532021–532026.
- Honea, E. C., Ogura, A., Murray, C. A., Raghavachari, K., Sprenger, W. O., Jarrold, M. F., & Brown, W. L. (1993). Raman spectra of size-selected silicon clusters and comparison with calculated structures. *Nature*, *366*(6450), 42–44.
- Hossain, D., Hagelberg, F., Pittman, C. U., Jr., & Saebo, S. (2007). Structures and stabilities of clusters of Si_{12} , Si_{18} and Si_{20} containing endohedral charged and neutral atomic species. *Journal of Physical Chemistry C*, *111*(37), 13864–13871.
- Jackson, K., Pederson, M., Wang, C.-Z., & Ho, K.-M. (1999). Calculated polarizabilities of intermediate-size Si clusters. *Physical Review A - Atomic, Molecular, and Optical Physics*, *59*(5), 3685–3689.
- Jackson, K. A., Horoi, M., Chaudhuri, I., Frauenheim, T., & Shvartsburg, A. A. (2004). Unraveling the shape transformation in silicon clusters. *Physical Review Letters*, *93*(1), 013401/1–013401/4.
- Jackson, K. A., Yang, M., Chaudhuri, I., & Frauenheim, T. (2005). Shape, polarizability, and metallicity in silicon clusters. *Physical Review A - Atomic, Molecular, and Optical Physics*, *71*(3), 1–6.
- Jackson, K., Yang, M., & Jellinek, J. (2007). Site-specific analysis of dielectric properties of finite systems. *Journal of Physical Chemistry C*, *111*(48), 17952–17960.
- Jarrold, M. F., & Bower, J. E. (1992). Mobilities of silicon cluster ions: the reactivity of silicon sausages and spheres. *The Journal of Chemical Physics*, *96*(12), 9180–9190.
- Jarrold, M. F., & Constant, V. A. (1991). Silicon cluster ions: evidence for a structural transition. *Physical Review Letters*, *67*(21), 2994–2997.
- Jelski, D. A., Swift, B. L., Rantala, T. T., Xia, X., & George, T. F. (1991). Structure of the Si_{45} cluster. *Journal of Chemical Physics*, *95*(11), 8552–8560.
- Jha, P. C., Seal, P., Sen, S., Ågren, H., & Chakrabarti, S. (2008). Static and dynamic polarizabilities of $(CdSe)_n$ ($n = 1-16$) clusters. *Computational Materials Science*, *44*(2), 728–732.
- Jose, R., Zhanpeisov, N. U., Fukumura, H., Baba, Y., & Ishikawa, I. (2006). Structure-property correlation of CdSe clusters using experimental results and first-principles DFT calculations. *Journal of the American Chemical Society*, *128*(2), 629–636.

- Kanis, D. R., Ratner, M. A., & Marks, T. J. (1994). Design and construction of molecular assemblies with large second-order optical nonlinearities. quantum chemical aspects. *Chemical Reviews*, 94(1), 195–242.
- Karamanis, P., & Leszczynski, J. (2008). Correlations between bonding, size, and second hyperpolarizability (γ) of small semiconductor clusters: ab initio study on Al_nP_n clusters with $n = 2, 3, 4, 6$, and 9 . *Journal of Chemical Physics*, 128(15), 154323/1–154323/10.
- Karamanis, P., & Pouchan, C. (2009). How large are the microscopic electronic dipole (hyper)polarizabilities of Cd_nTe_n bare clusters compared to those of Cd_nS_n and Cd_nSe_n ? A systematic ab initio study. *Chemical Physics Letters*, 474(1–3), 162–167.
- Karamanis, P., & Pouchan, C. (2011). On the shape dependence of cluster (hyper)polarizabilities. A combined ab initio and DFT study on large fullerene-like gallium arsenide semiconductor clusters. *International Journal of Quantum Chemistry*, 111(4), 788–796.
- Karamanis, P., Maroulis, G., & Pouchan, C. (2006a). Basis set and electron correlation effects in all-electron ab initio calculations of the static dipole polarizability of small cadmium selenide clusters, $(CdSe)_n$ $n = 1, 2, 3, 4$. *Chemical Physics*, 331(1), 19–25.
- Karamanis, P., Maroulis, G., & Pouchan, C. (2006b). Molecular geometry and polarizability of small cadmium selenide clusters from all-electron Ab initio and density functional theory calculations. *Journal of Chemical Physics*, 124(7), 071101/1–071101/2.
- Karamanis, P., Begue, D., & Pouchan, C. (2007a). Ab initio finite field (hyper)polarizability computations on stoichiometric gallium arsenide clusters Ga_nAs_n ($n = 2–9$). *Journal of Chemical Physics*, 127(9), 094706/1–094706/10.
- Karamanis, P., Zhang-Negrerie, D., & Pouchan, C. (2007b). A critical analysis of the performance of conventional ab initio and DFT methods in the computation of Si_5 ground state. *Chemical Physics*, 331(2–3), 417–426.
- Karamanis, P., Pouchan, C., & Leszczynski, J. (2008a). Electric dipole (hyper)polarizabilities of selected X_2Y_2 and X_3Y_3 ($X = Al, Ga, In$ and $Y = P, As$): III-V semiconductor clusters. An ab initio comparative study. *Journal of Physical Chemistry A*, 112(51), 13662–13671.
- Karamanis, P., Xenides, D., & Leszczynski, J. (2008b). Polarizability evolution on natural and artificial low dimensional binary semiconductor systems: a case study of stoichiometric aluminum phosphide semiconductor clusters. *Journal of Chemical Physics*, 129(9), 094708/1–094708/12.
- Karamanis, P., Xenides, D., & Leszczynski, J. (2008c). The polarizabilities of small stoichiometric aluminum phosphide clusters Al_nP_n ($n = 2–9$). Ab initio and density functional investigation. *Chemical Physics Letters*, 457(1–3), 137–142.
- Karamanis, P., Pouchan, C., & Maroulis, G. (2008d). Structure, stability, dipole polarizability and differential polarizability in small gallium arsenide clusters from all-electron ab initio and density-functional-theory calculations. *Physical Review A - Atomic, Molecular, and Optical Physics*, 77(1), 013201/1–013201/6.
- Karamanis, P., Carbonnière, P., & Pouchan, C. (2009). Structures and composition-dependent polarizabilities of open- and closed-shell gan asm semiconductor clusters. *Physical Review A - Atomic, Molecular, and Optical Physics*, 80(5), 053201/1–053201/11.
- Karamanis, P., Marchal, R., Carbonnière, P., & Pouchan, C. (2010). Doping effects on the electric response properties of Silicon clusters. A global structure-property investigation of $AlSi_{n-1}$ clusters ($n = 3–10$). *Chemical Physics Letters*, 474(1–3), 59–64.
- Karamanis, P., Pouchan, C., Weatherford, C. A., & Gutsev, G. L. (2011). Evolution of properties in prolate $(GaAs)_n$ clusters. *Journal of Physical Chemistry C*, 115(1), 97–107.
- Kasuya, A., Sivamohan, R., Barnakov, Y. A., Dmitruk, I. M., Nirasawa, T., Romanyuk, V. R., et al. (2004). Ultra-stable nanoparticles of CdSe revealed from mass spectrometry. *Nature Materials*, 3(2), 99–102.
- Kaxiras, E., & Jackson, K. (1993). Shape of small silicon clusters. *Physical Review Letters*, 71(5), 727–730.
- Kim, H.-Y., Sofo, J. O., Velegol, D., Cole, M. W., & Mukhopadhyay, G. (2005). Static polarizabilities of dielectric nanoclusters. *Physical Review A - Atomic, Molecular, and Optical Physics*, 72(5), 1–8.

- Koch, W., & Holthausen, M. C. (2000). *A Chemist's guide to density functional theory*. Chichester: Wiley.
- Korambath, P. P., & Karna, S. P. (2000). (Hyper)polarizabilities of GaN, GaP, and GaAs clusters: an ab initio time-dependent Hartree-Fock study. *Journal of Physical Chemistry A*, *104*(20), 4801–4804.
- Krishtal, A., Senet, P., & Van Alsenoy, C. (2010). Origin of the size-dependence of the polarizability per atom in heterogeneous clusters: the case of AlP clusters. *Journal of Chemical Physics*, *133*(15), 154310/1–154310/11.
- Kurtz, H. A., Stewart, J. J. P., & Dieter, K. M. (1990). Calculation of the nonlinear optical properties of molecules. *Journal of Computational Chemistry*, *11*(1), 82–87.
- Lan, Y.-Z., & Feng, Y.-L. (2009). Study of absorption spectra and (hyper)polarizabilities of SiC_n and Si_nC (n = 2–6) clusters using density functional response approach. *Journal of Chemical Physics*, *131*(5), 054509/1–054509/11.
- Lan, Y., Cheng, W., Wu, D., Li, X., Zhang, H., & Gong, Y. (2003). TDHF-SOS treatments on linear and nonlinear optical properties of III-V semiconductor clusters (Ga₃, As₃, Ga₃Sb₃, In₃P₃, In₃As₃, In₃Sb₃). *Chemical Physics Letters*, *372*(5–6), 645–649.
- Lan, Y.-Z., Cheng, W.-D., Wu, D.-S., Shen, J., Huang, S.-P., Zhang, H., Gong, Y.-J., & Li, F.-F. (2006). A theoretical investigation of hyperpolarizability for small Ga_nAs_m (n + m = 4–10) clusters. *Journal of Chemical Physics*, *124*(9), 094302/1–094302/8.
- Lan, Y.-Z., Feng, Y.-L., Wen, Y.-H., & Teng, B.-T. (2008). Dynamic second-order hyperpolarizabilities of Si₃ and Si₄ clusters using coupled cluster cubic response theory. *Chemical Physics Letters*, *461*(1–3), 118–121.
- Leitsmann, R., Schmidt, W. G., Hahn, P. H., & Bechstedt, F. (2005). Second-harmonic polarizability including electron-hole attraction from band-structure theory. *Physical Review B - Condensed Matter and Materials Physics*, *71*(19), 195209/1–195209/10.
- Li, B.-X. (2005). Stability of medium-sized neutral and charged silicon clusters. *Physical Review B - Condensed Matter and Materials Physics*, *71*(23), 1–7.
- Li, B.-X., Cao, P.-L., & Zhou, X.-Y. (2003). Electronic and geometric structures of Siⁿ⁻ and Siⁿ⁺ (n = 2–10) clusters and in comparison with Siⁿ Physica Status Solidi (B). *Basic Research*, *238*(1), 11–19.
- Li, L., Zhou, Z., Wang, X., Huang, W., He, Y., & Yang, M. (2008). First-principles study of static polarizability, first and second hyperpolarizabilities of small-sized ZnO clusters. *Physical Chemistry Chemical Physics*, *10*(45), 6829–6835.
- Liao, D. W., & Balasubramanian, K. (1992). Electronic structure of the III-V tetramer clusters and their positive ions. *Journal of Chemical Physics*, *96*(12), 8938–8947.
- Lipscomb, W. N. (1966). Framework rearrangement in boranes and carboranes. *Science*, *153*(3734), 373–378.
- Lou, L., Nordlander, P., & Smalley, R. E. (1992). Electronic structure of small GaAs clusters II. *Journal of Chemical Physics*, *97*(3), 1858–1864.
- Luis, J. M., Duran, M., Champagne, B., & Kirtman, B. (2000). Determination of vibrational polarizabilities and hyperpolarizabilities using field-induced coordinates. *Journal of Chemical Physics*, *113*(13), 5203–5213.
- Lyon, J. T., Gruene, P., Fielicke, A., Meijer, G., Janssens, E., Claes, P., et al. (2009). Structures of silicon cluster cations in the gas phase. *Journal of the American Chemical Society*, *131*(3), 1115–1121.
- Marchal, R., Carbonnière, P., & Pouchan, C. (2009). A global search algorithm of minima exploration for the investigation of low lying isomers of clusters from density functional theory-based potential energy surfaces: the example of Si_n (n = 3, 15) as a test case. *Journal of Chemical Physics*, *131*(11), 114105/1–114105/9.
- Marchal, R., Carbonnière, P., & Pouchan, C. (2010). A global search algorithm of minima exploration for the investigation of low lying isomers of clusters from DFT-based potential energy surface. A theoretical study of Si_n and Si_{n-1} Al clusters. *International Journal of Quantum Chemistry*, *110*(12), 2256–2259.

- Marchal, R., Carbonnière, P., & Pouchan, C. (2011). On the Structures of Non-Stoichiometric GaAs_n Clusters (5 n < + m < 8). *Journal of Computational and Theoretical Nanosciences*, 8(4), 568–578.
- Marks, T. J., & Ratner, M. A. (1995). Design, synthesis, and properties of molecule-based assemblies with large second-order optical nonlinearities. *Angewandte Chemie (International Edition in English)*, 34(2), 155–173.
- Maroulis, G. (2003). Accurate electric multipole moment, static polarizability and hyperpolarizability derivatives for N₂. *Journal of Chemical Physics*, 118(6), 2673–2687.
- Maroulis, G. (2004). Bonding and (hyper) polarizability in the sodium dimer. *Journal of Chemical Physics*, 121(21), 10519–10524.
- Maroulis, G. (2008). How large is the static electric (hyper)polarizability anisotropy in HXeI? *Journal of Chemical Physics*, 129(4), 044314/1–044314/6.
- Maroulis, G., & Pouchan, C. (2003). Size and electric dipole (hyper)polarizability in small cadmium sulfide clusters: an ab initio study on (Cd)_n, n = 1, 2, and 4. *Journal of Physical Chemistry B*, 107(39), 10683–10686.
- Maroulis, G., Begué, D., & Pouchan, C. (2003). Accurate dipole polarizabilities of small silicon clusters from ab initio and density functional theory calculations. *Journal of Chemical Physics*, 119(2), 794–797.
- Maroulis, G., Karamanis, P., & Pouchan, C. (2007). Hyperpolarizability of GaAs dimer is not negative. *Journal of Chemical Physics*, 126(15), 154316/1–154316/5.
- Matxain, J. M., Fowler, J. E., & Ugalde, J. M. (2000). Small clusters of II–VI materials: Zn_iO_i, i = 1–9. *Physical Review A - Atomic, Molecular, and Optical Physics*, 62(5), 053201/1–053201/10.
- Matxain, J. M., Mercero, J. M., Fowler, J. E., & Ugalde, J. M. (2001). Small clusters of group-(II–VI) materials: Zn_i, Xi, X = Se, Te, i = 1–9. *Physical Review A. Atomic, Molecular, and Optical Physics*, 64(5), 532011–532018.
- Matxain, J. M., Mercero, J. M., Fowler, J. E., & Ugalde, J. M. (2003). Clusters of group II–VI materials: Cd_iO_i (i ≤ 15). *Journal of Physical Chemistry A*, 107(46), 9918–9923.
- Matxain, J. M., Mercero, J. M., Fowler, J. E., & Ugalde, J. M. (2004). Clusters of II–VI materials: Cd_iX_i, X = S, Se, Te, i ≤ 16. *Journal of Physical Chemistry A*, 108(47), 10502–10508.
- McLean, A. D., & Yoshimine, M. (1967). Theory of molecular polarizabilities. *Journal of Chemical Physics*, 47(6), 1927–1935.
- Menon, M., & Subbaswamy, K. R. (1995). Structure and stability of Si₄₅ clusters: a generalized tight-binding molecular-dynamics approach. *Physical Review B*, 51(24), 17952–17956.
- Michalet, X., Pinaud, F. F., Bentolila, L. A., Tsay, J. M., Doose, S., Li, J. J., et al. (2005). Quantum dots for live cells, in vivo imaging, and diagnostics. *Science*, 307(5709), 538–544.
- Mitas, L., Grossman, J. C., Stich, I., & Tobik, J. (2000). Silicon clusters of intermediate size: energetics, dynamics, and thermal effects. *Physical Review Letters*, 84(7), 1479–1482.
- Murray, C. B., Kagan, C. R., & Bawendi, M. G. (2000). Synthesis and characterization of monodisperse nanocrystals and close-packed nanocrystal assemblies. *Annual Review of Materials Science*, 30, 545–610.
- Nagle, J. K. (1990). Atomic polarizability and electronegativity. *Journal of the American Chemical Society*, 112(12), 4741–4747.
- Nair, N. N., Bredow, T., & Jug, K. (2004). Molecular dynamics implementation in MSINDO: study of silicon clusters. *Journal of Computational Chemistry*, 25(10), 1255–1263.
- Nigam, S., Majumder, C., & Kulshreshtha, S. K. (2004). Structural and electronic properties of Si_n Si⁺_n, and AlSi⁻¹_n (n = 2–13) clusters: theoretical investigation based on ab initio molecular orbital theory. *Journal of Chemical Physics*, 121(16), 7756–7763.
- O'Brien, S. C., Liu, Y., Zhang, Q., Heath, J. R., Tittel, F. K., Curl, R. F., et al. (1985). Supersonic cluster beams of III–V semiconductors: GaAs_n. *Journal of Chemical Physics*, 84(7), 4074–4079.
- Papadopoulos, M. G., Reis, H., Avramopoulos, A., Erkoç, S., & Amirouche, L. (2005). A comparative study of the dipole polarizability of some Zn clusters. *Journal of Physical Chemistry B*, 109(40), 18822–18830.

- Papadopoulos, M. G., Reis, H., Avramopoulos, A., Erkoç, S., & Amirouche, L. (2006). Polarizabilities and second hyperpolarizabilities of Zn_mCd_n clusters. *Molecular Physics*, *104*(13–14), 2027–2036.
- Parr, R. G., & Chattaraj, P. K. (1991). Principle of maximum hardness. *Journal of the American Chemical Society*, *113*(5), 1854–1855.
- Pedroza, L. S., & Da Silva, A. J. R. (2007). Ab initio monte carlo simulations applied to Si_5 cluster. *Physical Review B - Condensed Matter and Materials Physics*, *75*(24), 245331/1–245331/10.
- Peng, X., Wickham, J., & Alivisatos, A. P. (1998). Kinetics of II-VI and III-V colloidal semiconductor nanocrystal growth: ‘Focusing’ of size distributions. *Journal of the American Chemical Society*, *120*(21), 5343–5344.
- Peng, X., Manna, L., Yang, W., Wickham, J., Scher, E., Kadavanich, A., et al. (2000). Shape control of CdSe nanocrystals. *Nature*, *404*(6773), 59–61.
- Pool, R. (1990). Clusters: strange morsels of matter. *Science*, *248*(4960), 1186–1188.
- Pouchan, C., Bégué, D., & Zhang, D. Y. (2004). Between geometry, stability, and polarizability: density functional theory studies of silicon clusters Si_n ($n = 3–10$). *Journal of Chemical Physics*, *121*(10), 4628–4634.
- Powell, G. D., Wang, J.-F., & Aspnes, D. E. (2002). Simplified bond-hyperpolarizability model of second harmonic generation. *Physical Review B - Condensed Matter and Materials Physics*, *65*(20), 205320/1–205320/8.
- Prinzbach, H., Weller, A., Landenberger, P., Wahl, F., Wörth, J., Scott, L. T., et al. (2000). Gas-phase production and photoelectron spectroscopy of the smallest fullerene. *C20 Nature*, *407*(6800), 60–63.
- Pushpa, R., Narasimhan, S., & Waghmare, U. (2004). Symmetries, vibrational instabilities, and routes to stable structures of clusters of Al, Sn, and As. *Journal of Chemical Physics*, *121*(11), 5211–5220.
- Raghavachari, K., & Logovinsky, V. (1985). Structure and bonding in small silicon clusters. *Physical Review Letters*, *55*(26), 2853–2856.
- Raghavachari, K., & Rohlfing, C. M. (1988). Bonding and stabilities of small silicon clusters: a theoretical study of $Si_7–Si_{10}$. *Journal of Chemical Physics*, *89*(4), 2219–2234.
- Raptis, S. G., Papadopoulos, M. G., & Sadlej, A. J. (1999). The correlation, relativistic, and vibrational contributions to the dipole moments, polarizabilities, and first and second hyperpolarizabilities of ZnS, CdS, and HgS. *Journal of Chemical Physics*, *111*(17), 7904–7915.
- Reis, H., Papadopoulos, M. G., & Boustani, I. (2000). DFT calculations of static dipole polarizabilities and hyperpolarizabilities for the boron clusters bn ($n = 3–8, 10$). *International Journal of Quantum Chemistry*, *78*(2), 131–135.
- Rohlfing, C. M., & Raghavachari, K. A. (1990). Theoretical study of small silicon clusters using an effective core potential. *Chemical Physics Letters*, *167*(6), 559–565.
- Roman, E., Yates, J. R., Veithen, M., Vanderbilt, D., & Souza, I. (2006). Ab initio study of the nonlinear optics of III-V semiconductors in the terahertz regime. *Physical Review B - Condensed Matter and Materials Physics*, *74*(24), 245204/1–245204/9.
- Sanville, E., Burnin, A., & BelBruno, J. J. (2006). Experimental and computational study of small ($n = 1–16$) stoichiometric zinc and cadmium chalcogenide clusters. *Journal of Physical Chemistry A*, *110*(7), 2378–2386.
- Schäfer, R., Schlecht, S., Woenckhaus, J., & Becker, J. A. (1996). Polarizabilities of isolated semiconductor clusters. *Physical Review Letters*, *76*(3), 471–474.
- Schaller, R. D., & Klimov, V. I. (2006). Non-poissonian exciton populations in semiconductor nanocrystals via carrier multiplication. *Physical Review Letters*, *96*(9), 1–4.
- Schlecht, S., Schäfer, R., Woenckhaus, J., & Becker, J. A. (1995). Electric dipole polarizabilities of isolated gallium arsenide clusters. *Chemical Physics Letters*, *246*(3), 315–320.
- Schnell, M., Herwig, C., & Becker, J. A. (2003). Analysis of semiconductor cluster beam polarization taking small permanent dipole moments into account. *Zeitschrift Fur Physikalische Chemie*, *217*(8), 1003–1030.

- Sen, S., & Chakrabarti, S. (2006). Frequency-dependent nonlinear optical properties of CdSe clusters. *Physical Review B - Condensed Matter and Materials Physics*, 74(20), 205435/1–205435/7.
- Sokolova, S., Lüchow, A., & Anderson, J. B. (2000). Energetics of carbon clusters C₂₀ from all-electron quantum monte carlo calculations. *Chemical Physics Letters*, 323(3–4), 229–233.
- Song, K. M., Ray, A. K., & Khowash, P. K. (1994). On the electronic structures of GaAs clusters. *Journal of Physics B: Atomic Molecular and Optical Physics*, 27(8), 1637–1648.
- Sun, Q., Wang, Q., Jena, P., Waterman, S., & Kawazoe, Y. (2003). First-principles studies of the geometry and energetics of the Si₃₆ cluster. *Physical Review A - Atomic, Molecular, and Optical Physics*, 67(6), 632011–632016.
- Swaminathan, P., Antonov, V. N., Soares, J. A. N. T., Palmer, J. S., & Weaver, J. H. (2006). Cd-based II-VI semiconductor nanostructures produced by buffer-layer-assisted growth: structural evolution and photoluminescence. *Physical Review B - Condensed Matter and Materials Physics*, 73(12), 1–8.
- Szabo, A., & Ostlund, N. S. (1989). *Modern quantum chemistry*. New York: MacMillan.
- Tekin, A., & Hartke, B. (2004). Global geometry optimization of small silicon clusters with empirical potentials and at the DFT level. *Physical Chemistry Chemical Physics*, 6(3), 503–509.
- Torrens, F. (2002). Fractal dimension of different structural-type zeolites and of the active sites. *Physica E (Amsterdam)*, 13, 67.
- Troparevsky, M. C., & Chelikowsky, J. R. (2001). Structural and electronic properties of CdS and CdSe clusters. *Journal of Chemical Physics*, 114(2), 943–949.
- Troparevsky, M. C., Kronik, L., & Chelikowsky, J. R. (2002). Ab initio absorption spectra of CdSe clusters. *Physical Review B - Condensed Matter and Materials Physics*, 65(3), 333111–333114.
- Vasiliev, I., Ögüt, S., & Chelikowsky, J. R. (1997). Ab initio calculations for the polarizabilities of small semiconductor clusters. *Physical Review Letters*, 78(25), 4805–4808.
- Vela, A., & Gázquez, J. L. (1990). A relationship between the static dipole polarizability, the global softness, and the fukui function. *Journal of the American Chemical Society*, 112(4), 1490–1492.
- Vijayalakshmi, S., Lan, A., Iqbal, Z., & Grebel, H. (2002). Nonlinear optical properties of laser ablated silicon nanostructures. *Journal of Applied Physics*, 92(5), 2490–2494.
- Wang, X. Q., Clark, S. J., & Abram, R. A. (2004). Ab initio calculations of the structural and electronic properties of Hg_mTe_n cluster. *Physical Review B - Condensed Matter and Materials Physics*, 70(23), 1–6.
- Wang, B.-C., Chou, Y.-M., Deng, J.-P., & Dung, Y.-T. (2008). Structural and optical properties of passivated silicon nanoclusters with different shapes: a theoretical investigation. *Journal of Physical Chemistry A*, 112(28), 6351–6357.
- Wang, J., Ma, L., Zhao, J., & Jackson, K. A. (2009). Structural growth behavior and polarizability of Cd_nTe_n (n = 1–14) clusters. *Journal of Chemical Physics*, 130(21), 214307/1–214307/8.
- Wei, S., Barnett, R. N., & Landman, U. (1997). Energetics and structures of neutral and charged sin (n ≤ 10) and sodium-doped Si_nNa clusters. *Physical Review B - Condensed Matter and Materials Physics*, 55(12), 7935–7944.
- Williams, R. E. (1992). The polyborane, carborane, carbocation continuum: architectural patterns. *Chemical Reviews*, 92(2), 177–207.
- Wolf, S. A., Awschalom, D. D., Buhrman, R. A., Daughton, J. M., Von Molnár, S., Roukes, M. L., et al. (2001). Spintronics: a spin-based electronics vision for the future. *Science*, 294(5546), 1488–1495.
- Wu, F., Lewis, J. W., Kliger, D. S., & Zhang, J. Z. (2003). Unusual excitation intensity dependence of fluorescence of CdTe nanoparticles. *Journal of Chemical Physics*, 118(1), 12–16.
- Xenides, D. (2006). (Hyper)polarizability dependence on the interatomic distance of N₆(T_d): fourth order polynomials and third order derivatives. *Journal of Molecular Structure: Theochem*, 764(1–3), 41–46.
- Xenides, D., & Maroulis, G. (2000). Basis set and electron correlation effects on the first and second static hyperpolarizability of SO₂. *Chemical Physics Letters*, 319(5–6), 618–624.

- Xenides, D., & Maroulis, G. (2006). Electric polarizability and hyperpolarizability of $\text{BrCl}(X1\Sigma^+)$. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 39(17), 3629–3638.
- Xiao, C., Hagelberg, F., & Lester, W. A., Jr. (2002). Geometric, energetic, and bonding properties of neutral and charged copper-doped silicon clusters. *Physical Review B - Condensed Matter and Materials Physics*, 66(7), 754251–7542523.
- Yoo, S., & Zeng, X. C. (2005). Structures and stability of medium-sized silicon clusters. III Reexamination of motif transition in growth pattern from Si_{15} to Si_{20} . *Journal of Chemical Physics*, 123(16), 1–6.
- Yoo, S., & Zeng, X. C. (2006). Structures and relative stability of medium-sized silicon clusters. IV. motif-based low-lying clusters Si_{21} – Si_{30} . *Journal of Chemical Physics*, 124(5), 1–6.
- Yoo, S., Zhao, J., Wang, J., & Xiao, C. Z. (2004). Endohedral silicon fullerenes $\text{Si}_n(27 \leq n \leq 39)$. *Journal of the American Chemical Society*, 126(42), 13845–13849.
- Yoo, S., Shao, N., & Zeng, X. C. (2008). Structures and relative stability of medium- and large-sized silicon clusters. VI. Fullerene cage motifs for low-lying clusters Si_{39} , Si_{40} , Si_{50} , Si_{60} , Si_{70} and Si_{80} . *Journal of Chemical Physics*, 128(10), 104316/1–104316/9.
- Yu, D. K., Zhang, R. Q., & Lee, S. T. (2002). Structural transition in nanosized silicon clusters. *Physical Review B - Condensed Matter and Materials Physics*, 65(24), 2454171–2454176.
- Zdetsis, A. D. (2001). The real structure of the Si_6 cluster. *Physical Review A. Atomic, Molecular, and Optical Physics*, 64(2), 023202/1–023202/4.
- Zdetsis, A. D. (2007a). Analogy of silicon clusters with deltahedral boranes: how far can it go? reexamining the structure of Si_n and Si_n^{2-} , $n = 5$ –13 clusters. *Journal of Chemical Physics*, 127(24), 244308/1–244308/6.
- Zdetsis, A. D. (2007b). Fluxional and aromatic behavior in small magic silicon clusters: a full ab initio study of Si_n , Si_{1-n} , Si_{2-n} , and Si_{1+n} $n = 6, 10$ clusters. *Journal of Chemical Physics*, 127(1), 014314/1–014314/10.
- Zdetsis, A. D. (2008). High-stability hydrogenated silicon-carbon clusters: a full study of $\text{Si}_2\text{C}_2\text{H}_2$ in comparison to Si_2C_2 , $\text{C}_2\text{B}_2\text{H}_4$, and other similar species. *Journal of Physical Chemistry A*, 112(25), 5712–5719.
- Zdetsis, A. D. (2009). Silicon-bismuth and germanium-bismuth clusters of high stability. *Journal of Physical Chemistry A*, 113(44), 12079–12087.
- Zhang, D. Y., Bégué, D., & Pouchan, C. (2004). Density functional theory studies of correlations between structure, binding energy, and dipole polarizability in Si_9 – Si_{12} . *Chemical Physics Letters*, 398(4–6), 283–286.
- Zhao, W., & Cao, P.-L. (2001). Study of the stable structures of Ga_6As_6 cluster using FP-LMTO MD method. *Physics Letters, Section A: General, Atomic and Solid State Physics*, 288(1), 53–57.
- Zhao, W., Cao, P.-L., Li, B.-X., Song, B., & Nakamatsu, H. (2000). Study of the stable structures of Ga_4As_4 cluster using FP-LMTO MD method. *Physical Review B - Condensed Matter and Materials Physics*, 62(24), 17138–17143.
- Zhao, J., Xie, R.-R., Zhou, X., Chen, X., & Lu, W. (2006). Formation of stable fullerene-like Ga_nAs_n clusters ($6 \leq n \leq 9$): gradient-corrected density-functional theory and a genetic global optimization approach. *Physical Review B - Condensed Matter and Materials Physics*, 74(3), 035319/1–035319/2.
- Zhou, R. L., & Pan, B. C. (2008). Low-lying isomers of Si_n^+ and Si_n^- ($n = 31$ –50) clusters. *Journal of Chemical Physics*, 128(23), 234302/1–234302/6.
- Zhu, X., & Zeng, X. C. (2003). Structures and stabilities of small silicon clusters: ab initio molecular-orbital calculations of Si_7 – Si_{11} . *Journal of Chemical Physics*, 118(8), 3558–3570.
- Zhu, X. L., Zeng, X. C., Lei, Y. A., & Pan, B. (2004). Structures and stability of medium silicon clusters. II Ab initio molecular orbital calculations of Si_{12} – Si_{20} . *Journal of Chemical Physics*, 120(19), 8985–8995.

Soohaeng Yoo and Sotiris S. Xantheas

Contents

Introduction	1140
Models of Intermolecular Interactions	1141
Quantum Models from Electronic Structure Calculations	1142
Global Minimum Structures of $n = 2-10$ Water Clusters	1146
The Global Minima of Medium-Sized Water Clusters in the Range $11 \leq n \leq 16$	1152
The Transition from “All-Surface” to “Internally Solvated” Clusters at $n = 17$	1155
Validation from Electronic Structure Calculations	1156
Spectroscopic Signature	1157
The Family of Minima for $(\text{H}_2\text{O})_{20}$	1157
The Four Major Families of Minima	1159
Vibrational Spectra	1160
The First Few Water Cages	1160
Significance	1160
The Pentagonal Dodecahedron (D-cage) $(\text{H}_2\text{O})_{20}$ Cluster	1162
The Tetrakaidecahedron (T-Cage) $(\text{H}_2\text{O})_{24}$ Cluster	1165
Outlook	1166
Bibliography	1168

Abstract

This chapter discusses the structures, energetics, and vibrational spectra of the first few ($n \leq 24$) water clusters obtained from high-level electronic structure calculations. The results are discussed in the perspective of being used to parameterize/assess the accuracy of classical and quantum force fields for water. To this end, a general introduction with the classification of those force fields is presented. Several low-lying families of minima for the medium cluster sizes

S. Yoo (✉) • S.S. Xantheas

Chemical and Material Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA

e-mail: sotiris.xantheas@pnl.gov

are considered. The transition from the “all surface” to the “fully coordinated” cluster structures occurring at $n = 17$ and its spectroscopic signature is presented. The various families of minima for $n = 20$ are discussed together with the low-energy networks of the pentagonal dodecahedron $(\text{H}_2\text{O})_{20}$ water cage. Finally, the low-energy networks of the tetrakaidecahedron (T-cage) $(\text{H}_2\text{O})_{24}$ cluster are shown and their significance in the construction of periodic lattices of structure I (sI) of the hydrate lattices is discussed.

Introduction

Water's function as a universal solvent and its role in mediating several biological functions that are responsible for sustaining life has created tremendous interest in the understanding of its structure at the molecular level (Ball 2008). Due to the size of the simulation cells and the sampling time needed to compute many macroscopic properties, most of the initial simulations are performed using a classical force field, whereas several processes that involve chemistry are subsequently probed with electronic structure based methods. A significant effort has therefore been devoted toward the development of classical force fields for water (Robinson et al. 1996). Clusters of water molecules are useful in probing the intermolecular interactions at the microscopic level as well as providing information about the subtle energy differences that are associated with different bonding arrangements within a hydrogen-bonded network. They moreover render a quantitative picture of the nature and magnitude of the various components of the intermolecular interactions such as exchange, dispersion, induction, etc. They can finally serve as a vehicle for the study of the convergence of properties with increasing size.

Over the last decade, there have been tremendous advances in the experimental (Cruzan et al. 1996a, b, 1997; Liu et al. 1994, 1996a, b, 1997a, b; Pugliano and Saykally 1992; Suzuki and Blake 1994; Viant et al. 1997) and theoretical (Xantheas 1994, 1996a; Xantheas and Dunning 1993a, b) studies of the structural, spectral, and energetic properties of water clusters, as well as refinements of the macroscopic structural experimental data for liquid water (Hura et al. 2000; Soper 2000; Sorenson et al. 2000). These developments have created a unique opportunity for incorporating molecular level information into classical force fields for water as well as using the available database of cluster energetics for the assessment of their accuracy. The process of developing interaction potentials is by no means straightforward, and the different approaches that are used to model the various components of the underlying physical interactions have their own advantages and shortcomings. Some of the outstanding issues that are associated with the development of models based on molecular level information are the shortage or scarcity of experimental data for clusters. To this end, an alternative approach based on the use of ab initio results for water clusters can be put into practice.

Assuming that the route of parameterization of an interaction potential for water from the cluster results is adopted, the question naturally arises of the level of accuracy that needs to be attained so the models can produce meaningful and

accurate results for the macroscopic properties of extended systems such as the bulk liquid and ice. In the subsequent sections, we will present the results of ab-initio electronic structure calculations for cluster sizes up to 24 water molecules and we will address some important features of their structural and spectral properties. We will finally present results on the first few water cages that are the constituents of the three-dimensional hydrate networks.

Models of Intermolecular Interactions

Intermolecular interactions Force fields describing intermolecular interactions can be in general classified into two categories: (1) classical and (2) quantum models, the characteristics of which are outlined in Fig. 1. Classical potentials consist of force fields that employ an analytic expression, in conjunction with several adjustable parameters, for the potential energy surface (PES) as a function of the coordinates of the atoms in the system. In contrast, for quantum models, the PES is computed by considering both the position of the nuclei and the distribution of the electrons of the system, and it is obtained by solving Schrödinger's equation.

The choice of the adjustable parameters used in conjunction with classical potentials can result to either effective potentials that implicitly include the nuclear quantization and can therefore be used in conjunction with classical simulations (albeit only for the conditions they were parameterized for) or transferable ones that attempt to best approximate the Born–Oppenheimer PES and should be used in nuclear quantum statistical simulations. Representative examples of effective force fields for water consist of TIP4P (Jorgensen et al. 1983), SPC/E (Berendsen et al. 1987) (pair-wise additive), and Dang-Chang (DC) (Dang and Chang 1997) (polarizable, many-body). The “polarizable” potentials contain – in addition to the pairwise additive term – a classical induction (polarization) term that explicitly

	Quantum		Classical			
Wave function	$\psi(\tilde{r}; \tilde{R})$		No electronic degrees of freedom			
Energy	$E = \langle \psi \hat{H} \psi \rangle$		$E = E(\tilde{R}; a_i)$			
Potential energy surface	Born-Oppenheimer		“Effective”		Born-Oppenheimer	
Zero-point energy	Explicitly included		Implicitly included		Explicitly included	
Simulation methodology	Classical	Quantum	Classical	Quantum	Classical	Quantum
Transferability across different environments	NO	YES	NO [§]	NO [¶]	NO	YES

[§] appropriate only for the environment that the ZPE has been effectively included

[¶] double-counting

Fig. 1 Characteristics of classical and quantum potential models

(albeit approximately) accounts for many-body effects to infinite order. These effective potentials are fitted to reproduce bulk-phase experimental data (i.e., the enthalpy at $T = 298$ K and the radial distribution functions at ambient conditions) in classical molecular dynamics simulations of liquid water. Despite their simplicity, these models describe some experimental properties of liquid water at ambient conditions quite well but usually fail for conditions and environments that are not parameterized for, and so they have limited predictive power. On the other hand, force fields such as the Effective Fragment Potential (EFP) (Adamovic et al. 2003; Day et al. 1996; Gordon et al. 2001; Jensen and Gordon 1998; Netzloff and Gordon 2004), the SAPT-5 (Bukowski et al. 2007; Szalewicz et al. 2009), and the family of Thole-type models (TTM (Burnham et al. 1999), TTM-2F (Burnham and Xantheas 2002a), TTM2.1-F (Fanourgakis and Xantheas 2006), TTM3-F (Fanourgakis and Xantheas 2008), and TTM4-F (Burnham et al. 2008)) are examples of transferable potentials that are fitted to high-level first principles electronic structure results of the PESs of small water clusters, usually the water dimer and trimer, and should be used in conjunction with nuclear quantum statistical simulations.

The availability of energetic information for water clusters (such as binding energies) is of paramount importance in both fitting transferable potentials as well as assessing their performance. Currently, binding energies of small water clusters, even the water dimer, are not available experimentally but can only be obtained from high-level electronic structure calculations.

Quantum Models from Electronic Structure Calculations

Figure 2 shows the spectrum of methods that are used to describe intermolecular interactions ranging from the classical descriptions, such as Molecular Mechanics (MM), up to the highly correlated ones, including Coupled Cluster with Single,

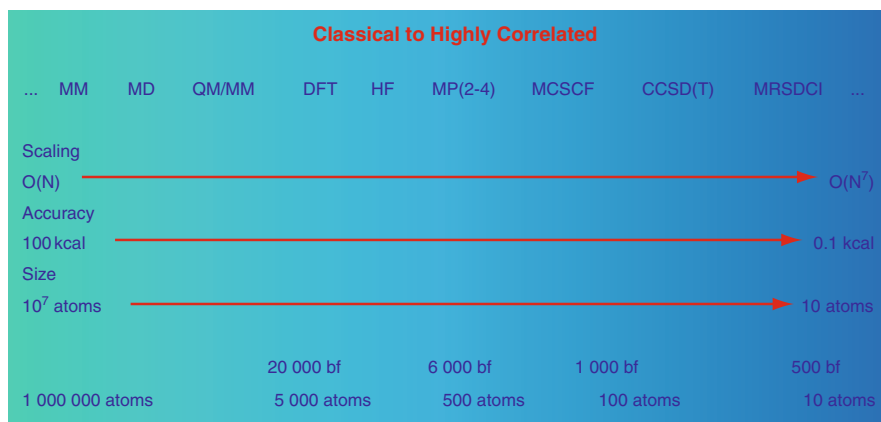


Fig. 2 Models of intermolecular interactions: from classical to highly correlated

Double, and perturbative estimates of Triple excitations [CCSD(T)] and Multi Reference Singles and Doubles Configuration Interaction (MRSDCI). Methods on the left part of the spectrum have the advantage of low formal scaling and as such allow for the treatment of large systems, but they are also associated with large inaccuracies. In contrast, methods to the right side of Fig. 2 have the disadvantage of high formal scaling (as much as $\sim O(N^9)$, where N is the system size) and are more appropriate for smaller systems for which high accuracy is achieved. Currently, the state of the art consists of the application of highly correlated methods such as CCSD(T) to a cluster of about 24 water molecules ($\sim 1,100$ basis functions and 270 electrons), a calculation that can sustain a performance of 1.4 Petaflop/s (Apra et al. 2009).

Electron Correlation and Orbital Basis Set

The cluster energies are obtained from the solution of the nonrelativistic Schrödinger equation for each system. The expansion of the trial many-electron wave function delineates the level of theory (description of electron correlation), whereas the description of the constituent orbitals is associated with the choice of the orbital basis set. A recent review (Dunning 2000) outlines a path, which is based on hierarchical approaches in this double expansion in order to ensure convergence of both the correlation and basis set problems. It also describes the application of these hierarchical approaches to various chemical systems that are associated with very diverse bonding characteristics, such as covalent bonds, hydrogen bonds and weakly bound clusters.

As regards the description of the electron correlation problem it has been recognized that the coupled cluster method (Bartlett and Purvis 1978; Cizek 1966, 1969; Coester 1958; Coester and Kümmel 1960; Kucharski and Bartlett 1992; Purvis and Bartlett 1982), which includes all possible single, double, triple, etc. excitations from a reference wave function, represents a viable route toward obtaining accurate energetics for hydrogen-bonded dimers (Halkier et al. 1999; Peterson and Dunning 1995). Among its variations, the CCSD(T) approximation (Raghavachari et al. 1989, 1990), which includes the effects of triple excitations perturbationally, represents an excellent compromise between accuracy and computational efficiency and has been recently been coined as the “gold standard of quantum chemistry.” The CCSD(T) approximation scales as N^6 for the iterative solution of the CCSD part, plus an additional single N^7 step for the perturbative estimate of the triple excitations. This represents a substantial additional expense over the widely used second-order perturbation level (Møller and Plesset 1934) of theory (MP2) which formally scales as N^5 .

For the basis set expansion, the correlation-consistent (cc-pV nZ) orbital basis sets of Dunning and coworkers (Dunning 1989; Dunning et al. 1998; Kendall et al. 1992), ranging from double to quintuple zeta quality ($n = D, T, Q, 5$), offer a systematic path in approaching the complete basis set (CBS) limit. These sets were constructed by grouping together all basis functions that contribute roughly equal amounts to the correlation energy of the atomic ground states. In this approach, functions are added to the basis sets in shells. The sets approach the complete basis

set (CBS) limit, for each succeeding set in the series provides an evermore accurate description of both the atomic radial and angular spaces. The extension of those sets to include additional diffuse functions for each angular momentum function present in the standard basis yields the augmented correlation consistent (aug-cc-pVnZ) sets. The exponents of those additional diffuse functions were optimized for the corresponding negative ions.

The Basis Set Superposition Error Correction

The use of a finite basis set in electronic structure calculations as been known to overestimate the interaction energy (Clementi 1967) (ΔE) of a weakly bound dimer complex, a quantity defined as

$$\Delta E = E_{AB}^{\alpha\cup\beta}(AB) - E_A^\alpha(A) - E_B^\beta(B), \quad (1)$$

where superscripts indicate the basis set; subscripts, the molecular system; and parentheses, the geometries for which the energies are obtained. In this notation, the term $E_{AB}^{\alpha\cup\beta}(AB)$ is the energy of the dimer (AB) at its equilibrium geometry AB with the basis set of the two fragments aUb. The overestimation comes from the fact that the basis functions that are centered on the one fragment help lower the energy of the other fragment and vice versa, a situation first termed “Basis Set Superposition Error” (BSSE) by Liu and McLean (1973). Boys and Bernardi (1970) originally proposed two approaches to circumvent this problem, introducing corrections via the “point counterpoise” (pCP) and the “function counterpoise” (fCP) methods, the latter being nowadays almost exclusively adopted in electronic structure calculations of intermolecular interaction energies (Chalasinski and Szczesniak 1994; van Duijneveldt et al. 1994).

For a dimer of two interacting moieties (extension to larger clusters is straightforward), the fCP correction is:

$$\Delta E \text{ (fCP)} = E_{AB}^{\alpha\cup\beta}(AB) - E_{AB}^{\alpha\cup\beta}(A) - E_{AB}^{\alpha\cup\beta}(B). \quad (2)$$

Equations 1 and 2 will not converge to the same result as the basis set increases toward the Complete Basis Set (CBS) limit since the reference energies of fragments A and B are computed at different geometries, viz. the dimer vs. the isolated fragments. Alternatively, if the BSSE correction is estimated via the equation (Xantheas 1996b)

$$\Delta E \text{ (BSSE)} = E_{AB}^{\alpha\cup\beta}(AB) - E_{AB}^{\alpha\cup\beta}(A) - E_{AB}^{\alpha\cup\beta}(B) + E_{\text{rel}}^\alpha(A) + E_{\text{rel}}^\beta(B) \quad (3)$$

where

$$E_{\text{rel}}^\alpha(A) = E_{AB}^\alpha(A) - E_A^\alpha(A) \quad (4)$$

$$E_{\text{rel}}^\beta(B) = E_{AB}^\beta(B) - E_B^\beta(B) \quad (5)$$

represent the energy penalty for distorting the fragments from their isolated geometries to the ones that adopt due to their interaction in the complex, then it is readily seen that after substitution of Eqs. 4 and 5 into Eq. 3 and collecting terms, the BSSE correction (Eq. 3) can be cast as:

$$\Delta E (\text{BSSE}) = \Delta E - \left\{ E_{\text{AB}}^{\alpha\cup\beta} (\text{A}) - E_{\text{AB}}^{\alpha} (\text{A}) \right\} - \left\{ E_{\text{AB}}^{\alpha\cup\beta} (\text{B}) - E_{\text{AB}}^{\beta} (\text{B}) \right\}. \quad (6)$$

Equations 1 and 6 do converge to the same result at the CBS limit since the terms in the brackets in Eq. 6 will numerically approach zero as the basis sets α and β tend toward CBS, viz.

$$\lim_{\alpha, \beta \rightarrow \text{CBS}} \Delta E (\text{BSSE}) = \lim_{\alpha, \beta \rightarrow \text{CBS}} \Delta E. \quad (7)$$

The importance of this correction, although previously recognized (Emsley et al. 1978; Smit et al. 1978; van Lenthe et al. 1987), has rarely been applied (Eggenberger et al. 1991; Kendall et al. 1989; Leclercq et al. 1983; Mayer and Surjan 1992; van Duijneveldt-van de Rijdt and van Duijneveldt 1992) until the problems arising from omitting it, especially for (1) strongly bound hydrogen-bonded complexes exhibiting large fragment relaxations and (2) calculations employing large basis sets, were noted. It should be emphasized that for some strongly hydrogen-bonded dimers such as $\text{F}^-(\text{H}_2\text{O})$ and $\text{OH}^-(\text{H}_2\text{O})$, the energy penalty term is quite large (≥ 5 kcal/mol) and its omission will lead to erroneous results as regards the magnitude and convergence of the uncorrected and BSSE-corrected values of the interaction energy as the basis set increases. This large energy penalty is caused by the strong interaction between the ion and water, a fact that results in a large elongation of the hydrogen-bonded OH bond causing the water molecule to adopt a configuration that is far from its isolated equilibrium geometry. A further consequence of this effect is the large red shift in the corresponding OH stretching vibration, a fact that has been experimentally verified.

Extrapolation to the Complete Basis Set Limit

Application of the family of these sets to a variety of chemical systems (see Reference Dunning (2000) and references therein) – ranging from the very weakly bound (by around 0.01 kcal/mol) rare gas diatomics, to intermediate strength hydrogen-bonded neutral (2–5 kcal/mol), singly charged ion-water (10–30 kcal/mol) clusters, and single-bond diatomics (50–100 kcal/mol), to very strong (>200 kcal/mol) multiply charged metal–water clusters and multiple bond diatomics – has permitted a heuristic extrapolation of the computed electronic energies and energy differences to the CBS limit. Among the various approaches that have been proposed (Bunge 1970; Fast et al. 1999; Feller 1992; Halkier et al. 1999; Klopper 1995; Martin 1996; Termath et al. 1991; Wilson and Dunning 1997; Xantheas and Dunning 1993c) in order to arrive at the CBS limit, we have relied on the following two:

1. A polynomial with inverse powers of 4 and 5 (4–5 polynomial):

$$\Delta E = \Delta E_{\text{CBS}} + \gamma/(\ell_{\text{max}} + 1)^4 + (\ell_{\text{max}} + 1)^5 \quad (8)$$

where ℓ_{max} is the value of the highest angular momentum function in the basis set and

- An exponential dependence on the cardinal number of the basis set n ($n = 2, 3, 4, 5$ for the sets of double through quintuple zeta quality, respectively):

$$\Delta E = \Delta E_{\text{CBS}} + \alpha \bullet \exp(-\beta \bullet n). \quad (9)$$

It should be noted that in nearly every case, this “extrapolation” procedure only accounts for a very small change when compared to the “best” computed quantity with the largest basis set [usually the (aug)-cc-pV5Z or, computer resources permitting, the (aug)-cc-pV6Z]. This result suggests that effective convergence of the respective properties (such as structure and relative energies) has already been achieved with the largest basis sets of this family.

Global Minimum Structures of then = 2–10 Water Clusters

Figure 3 shows the global minimum structures of the water clusters $n = 2-5$ and Fig. 4 shows the variation of their uncorrected and BSSE-corrected binding energies from MP2 calculations with the correlation-consistent basis sets (Xantheas et al. 2002). The CBS limit of -15.8 kcal/mol for the cyclic trimer was estimated from MP2 binding energies with basis sets up to aug-cc-pV5Z.

It is worth noting that extrapolating the uncorrected and BSSE-corrected binding energies using Eqs. 8 and 9 yields CBS limits that are within 0.01 kcal/mol from each other, while the extrapolation process itself results in minimal (<0.2 kcal/mol) changes with respect to the best computed numbers with the largest aug-cc-pV5Z set. For the cyclic water tetramer, the CBS limit estimated using the 4–5 polynomial (see Eq. 8) is -27.6 kcal/mol. The extrapolation of the uncorrected and BSSE-

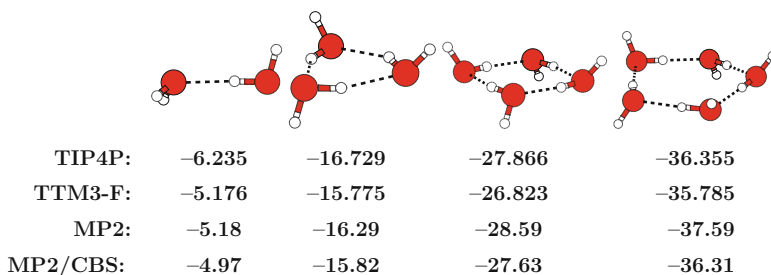
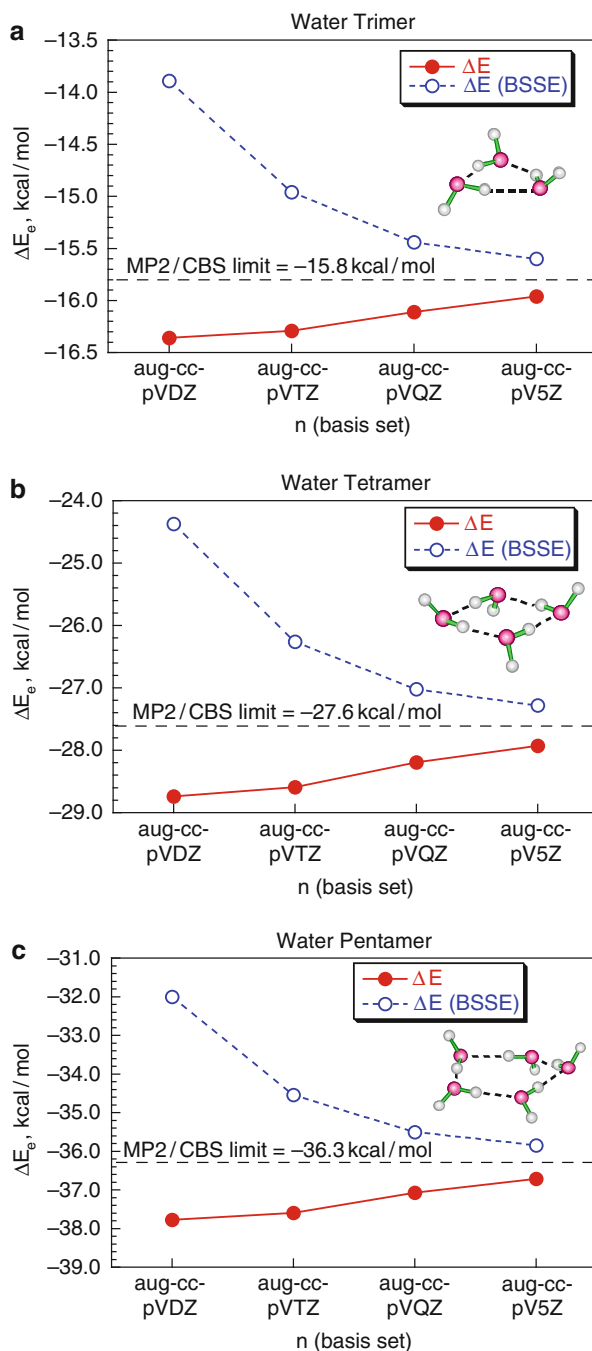


Fig. 3 Global minimum energy structures for $(\text{H}_2\text{O})_n$, $n = 1-5$ are the same for the TIP4P and TTM3-F potentials and the MP2 level of theory. The binding energies (D_e) at the MP2 level of theory (kcal/mol) were calculated with the aug-cc-pVTZ basis set

Fig. 4 The variation of the uncorrected (solid circles) and BSSE-corrected (open circles) binding energies from MP2 calculations with different basis sets for the global minimum structures of water clusters $n = 3-5$. The figure is reprinted with permission from reference Xantheas et al. (2002). © 2002 American Institute of Physics



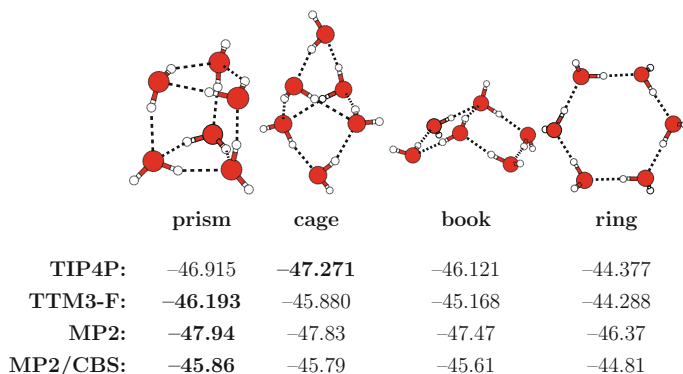
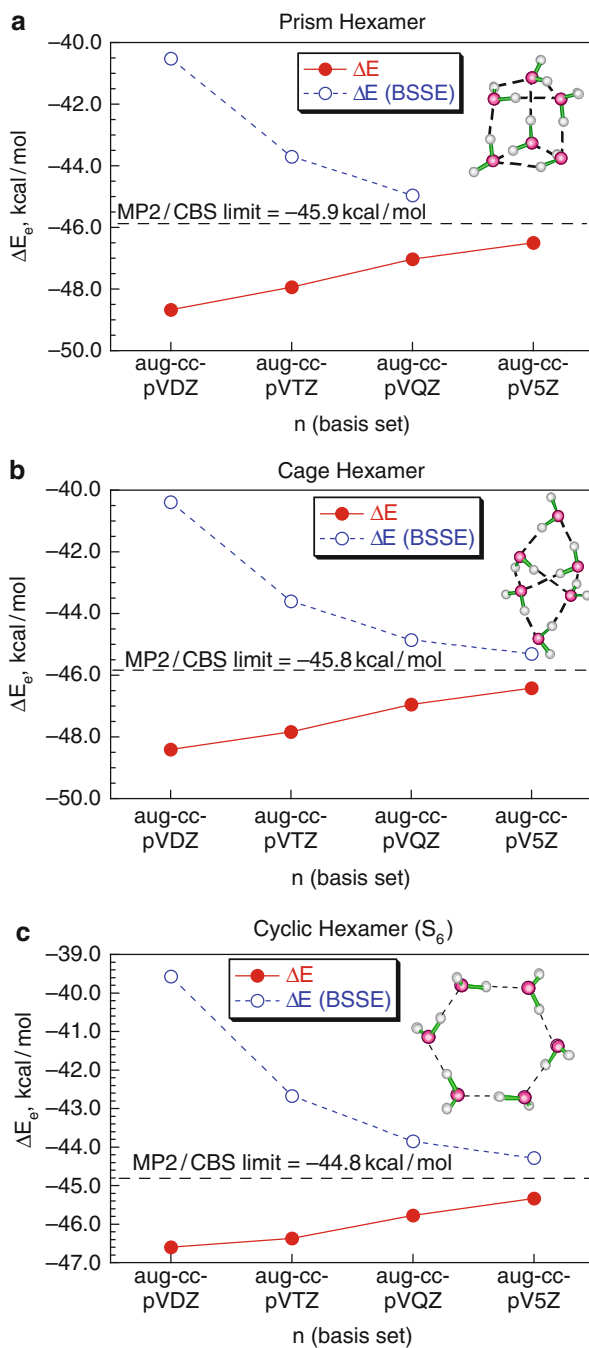


Fig. 5 Global minimum energy structures and binding energies (kcal/mol) for $(\text{H}_2\text{O})_6$. The TIP4P potential prefers the cage-like hexamer, while the TTM3-F potential the prism-like one. The prism-like hexamer becomes the global minimum structure at the MP2/aug-cc-pVTZ level of theory

corrected binding energies results in identical estimates for the CBS limit, which is just 0.3 kcal/mol away from the best computed value with the aug-cc-pV5Z set. The same trend is observed for the cyclic water pentamer with an estimated CBS limit of -36.3 kcal/mol. Both the effective pair-wise TIP4P (Wales and Hodges 1998) potential and the transferable, polarizable, many-body TTM3-F do yield these structures as the global minima, in agreement with the results of the MP2 calculations. The unique common feature of the global minimum structures of water clusters in the $n = 3-5$ size range is a cyclic (or ring) structure motif with the same direction of in-the-ring hydrogen bonding network in a homodromic direction (Xantheas 2000). These homodromic hydrogen bonding networks, i.e., those exhibiting donor-acceptor (da) arrangements between all water molecules, are associated with the largest non-additivities among other networks present in low-lying minima of small ($n = 3-5$) water clusters. These cyclic (or ring) structures, especially for $n = 4$ and 5 are of particular importance since they consist of the building blocks for the global minimum structures of water clusters in the size range $n = 6-20$.

The water hexamer has four distinct almost isoenergetic isomers that lie within 1 kcal/mol from each other (see Figs. 5 and 6) (Xantheas et al. 2002). Four distinct low-lying isomers are known as the prism, cage, book, and ring (S_6) structures. The prism isomer (bilayer of a trimer ring) is the global minimum at the MP2 level of theory with all basis sets, while the cage isomer becomes the global minimum when zero-point energy corrections are included. The MP2/CBS limits are estimated at -45.86 kcal/mol for the prism isomer and at -45.79 kcal/mol for the cage isomer, respectively. We have also estimated MP2/CBS limits of -45.61 kcal/mol for the book isomer and -44.81 kcal/mol for the ring isomer of S_6 symmetry. The calculations indicate that the “transition” from a quasi-planar, ring motif to a more compact, cage-like structure occurs for $n = 6$. A first difference among the results with classical water potentials was found for this cluster. The TIP4P potential

Fig. 6 The variation of the uncorrected (solid circles) and BSSE-corrected (open circles) binding energy from MP2 calculations with different basis sets for the water hexamer ($n = 6$). The figure is reprinted with permission from reference Xantheas et al. (2002). © 2002 American Institute of Physics



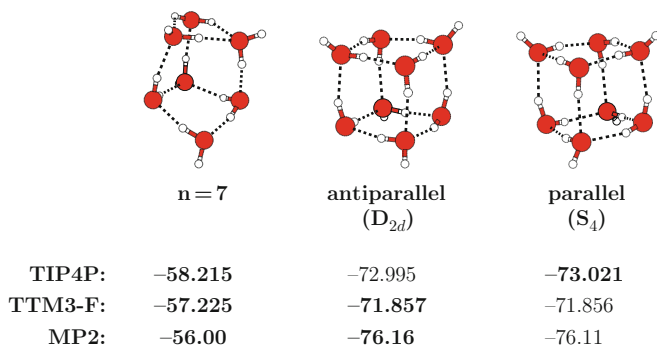
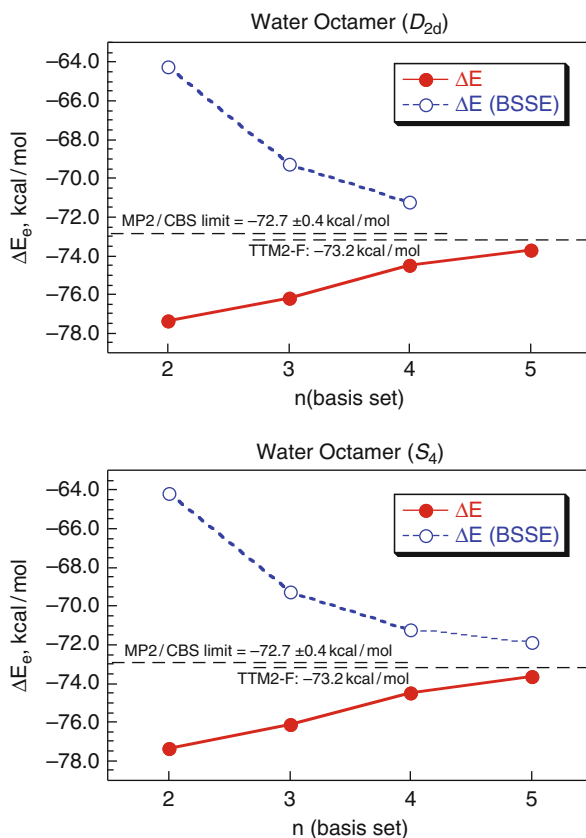


Fig. 7 The global minimum energy structures and binding energies (kcal/mol) for $(\text{H}_2\text{O})_7$ and $(\text{H}_2\text{O})_8$. For $n = 7$, the TIP4P and TTM3-F potentials have the same global minimum. For $n = 8$, the TIP4P potential favors the parallel circular arrangement of hydrogen bonds of two ring tetramers, while TTM3-F, the antiparallel arrangement. At the MP2/aug-cc-pVTZ level of theory, the antiparallel arrangement is preferred in agreement with the TTM3-F result

predicts the cage isomer as the global minimum (Wales and Hodges 1998), while the TTM3-F potential favors the prism isomer in agreement with the result of the MP2 calculations.

The water heptamer ($n = 7$) is an intermediate structure between the hexamer and the octamer. Note that water hexamer prefers to be a prism isomer (bilayer of cyclic ring trimer) and the water octamer a bilayer of the cyclic tetramer. The global minimum structure of the water heptamer consists of a cyclic trimer and cyclic tetramer on each layer as shown in Fig. 7. The water octamer ($n = 8$) has two nearly isoenergetic forms having D_{2d} and S_4 symmetry as the low-energy structures (see Figs. 7 and 8) (Xantheas and Aprà 2004). These are comprised of two stacked cyclic tetramers and the distinction between those two isoenergetic isomers lies in the direction of the two in-the-ring hydrogen bonding networks that can be either along the same (or parallel) for the S_4 isomer or opposite (or antiparallel) for the D_{2d} isomer. The variation of the MP2 binding energies with basis set is graphically illustrated in Fig. 8 for the two isomers. The best estimated BSSE-corrected MP2/CBS limits obtained with the 4–5 polynomial (cf. Eq. 8) are -72.94 kcal/mol and -72.78 kcal/mol for the antiparallel (D_{2d}) and the parallel (S_4) isomers, respectively. The corresponding BSSE-corrected MP2/CBS limits estimated with the exponential extrapolation (via Eq. 9) are -72.42 kcal/mol for the antiparallel and -72.35 kcal/mol for the parallel isomers, respectively. Note that the MP2 energy difference between both isomers is quite small of ~ 0.2 kcal/mol for both extrapolation methods. Given the heuristic nature of the two extrapolation methods to arrive at the CBS limit and these caveats, it is difficult and probably futile to assign a CBS limit for the binding energy with an “error bar” smaller than this quantity (~ 0.4 kcal/mol), which arises from the extrapolation procedure. As can be seen from Fig. 8, we suggest a value of -72.7 ± 0.4 kcal/mol for the MP2/CBS limit of the binding energy for both isomers of the water octamer, keeping in mind

Fig. 8 The variation of the uncorrected (solid circles) and BSSE-corrected (open circles) binding energy from MP2 calculations with different basis sets for the water octamer ($n = 8$). The figure is reprinted with permission from reference Xantheas and Aprà (2004). © 2004 American Institute of Physics



that the antiparallel isomer is always lower in energy by ~ 0.05 kcal/mol than the parallel one. As regards the results of classical potentials, TIP4P predicts the parallel isomer (Wales and Hodges 1998), while TTM3-F the antiparallel isomer as the global minimum structure in agreement with the results of the MP2 calculations.

Figure 9 shows the global minimum structures for $n = 9$ and 10. The TIP4P and TTM3-F potentials predict the same global minimum structure of $(\text{H}_2\text{O})_9$ consisting of the cyclic tetramer and the cyclic pentamer on each layer. The water decamer ($n = 10$) has two nearly isoenergetic isomers consisting of two stacked cyclic pentamers. These two isoenergetic isomers are distinguished by the direction of the in-the-ring hydrogen bonding networks as we have previously discussed for the two isomers of the water octamer. The antiparallel isomer of water decamer has the opposite direction of the in-the-ring hydrogen bonding networks on each layer, while the parallel isomer has the same direction. The TIP4P potential predicts the parallel isomer (Wales and Hodges 1998), whereas the TTM3-F potential predicts the antiparallel isomer as the global minimum structure that is consistent with the results of the MP2 calculations.

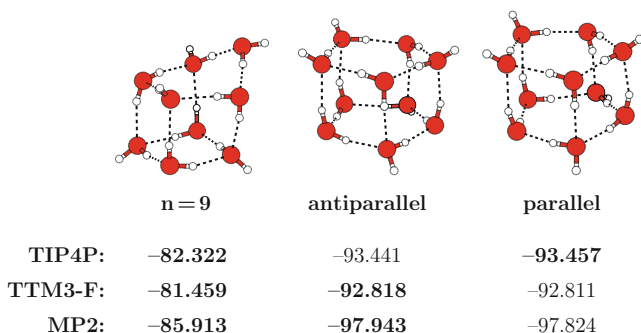


Fig. 9 The global minimum energy structures and binding energies (kcal/mol) of $(\text{H}_2\text{O})_9$ and $(\text{H}_2\text{O})_{10}$. The TIP4P and TTM3-F potentials predict the same global minimum structure for $(\text{H}_2\text{O})_9$, which is based on the antiparallel arrangement of hydrogen bonding in $(\text{H}_2\text{O})_8$. For $(\text{H}_2\text{O})_{10}$ (stacking of two pentagon rings), the TIP4P potential favors the parallel arrangement, while the TTM3-F potential the antiparallel arrangement. The binding energies at the MP2/aug-cc-pVTZ level of theory are listed

The Global Minima of Medium-Sized Water Clusters in the Range $11 \leq n \leq 16$

During the last 20 years, the search for the global minima of medium-sized water clusters has received much theoretical attention (Bulusu et al. 2006; Day et al. 2000; Fanourgakis et al. 2004, 2005; Hartke 2003; Kazimirski and Buch 2003; Koga et al. 2000; Lagutchenkov et al. 2005; Lee et al. 1995, 2001; Nigra and Kais 1999; Sadlej 2001; Tsai and Jordan 1993; Wales and Hodges 1998) and several previous ab initio calculations (Day et al. 2000; Fanourgakis et al. 2004, 2005; Koga et al. 2000; Lee et al. 1995, 2001; Sadlej 2001; Tsai and Jordan 1993) suggested that the global minima of medium-sized water clusters in the $n = 11$ –16 range mainly consist of the building blocks of the antiparallel octamer and the antiparallel decamer isomers. Figure 10 shows the global minimum structures for $n = 12$, 14, 15, and 16 at the MP2 level of theory. In detail, the global minima of $(\text{H}_2\text{O})_{12}$ and $(\text{H}_2\text{O})_{16}$ are stacked cube structures with the antiparallel arrangement of the in-the-ring hydrogen bonding networks that are built from the building blocks of the antiparallel octamer. For the global minimum of $(\text{H}_2\text{O})_{15}$, the antiparallel decamer was used as the building block to form a stacked pentagonal prism. In the case of the global minimum structure of $(\text{H}_2\text{O})_{14}$, two building blocks of the antiparallel octamer and decamer were used to obtain a fused square-pentagon prism structure. The TIP4P potential fails in predicting these arrangements as the global minima for $n = 12$, 15 and 16. It rather uses the parallel octamer and the parallel decamer isomers as starting points to build the global minimum structures of water clusters for $n = 12$, 15, and 16 since these were the most stable isomers for $n = 8$ and $n = 10$ (see Figs. 7 and 9) (Wales and Hodges 1998). In contrast, the TTM3-F potential again predicts the arrangements in agreement with the MP2 results.

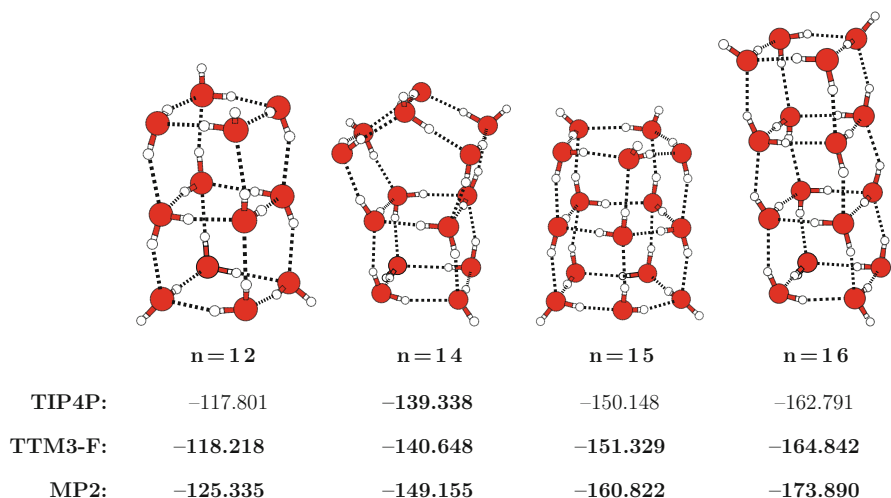


Fig. 10 The global minimum of $(\text{H}_2\text{O})_n$ for $n = 12, 14-16$. The TIP4P potential predicts a different global minimum for $n = 12, 15$ and $n = 16$. The MP2/aug-cc-pVDZ energies are listed

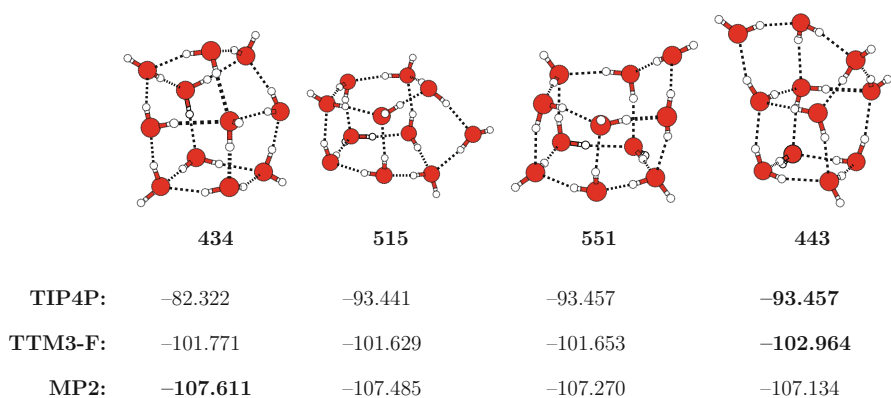


Fig. 11 The isomers of $(\text{H}_2\text{O})_{11}$. The (443) isomer is the global minimum structure with the TIP4P and TTM3-F potentials. The (434) isomer has the lowest binding energy (D_e) at the MP2/aug-cc-pVTZ level theory. Binding energies are shown in kcal/mol. The figure is reprinted with permission from reference Bulusu et al. (2006). © 2006 American Chemical Society

In the case of the odd-numbered water clusters ($n = 11$ and 13), it is not as easy to build the global minimum structures starting from most stable networks of the octamer and decamer clusters. Figure 11 shows the four distinct structural families of the low-lying isomers of $(\text{H}_2\text{O})_{11}$ (Bulusu et al. 2006). Two families labeled as (515) and (551) are built from the stacked pentagon $(\text{H}_2\text{O})_{10}$ (antiparallel decamer, the global minimum of $(\text{H}_2\text{O})_{10}$ as shown in Fig. 9). The (515) isomer can be viewed as constructed by adding one water molecule between the top and middle layers of

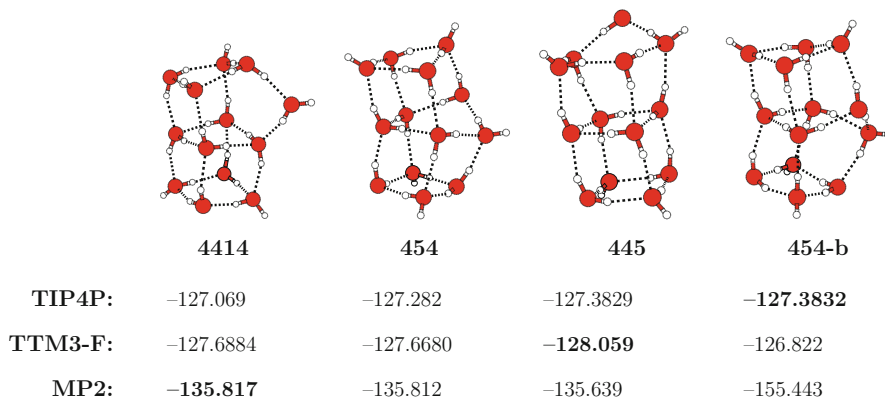


Fig. 12 The isomers of $(\text{H}_2\text{O})_{13}$. The MP2 energies were calculated with the aug-cc-pVDZ basis set. The figure is reprinted with permission from reference Bulusu et al. (2006). © 2006 American Chemical Society

the antiparallel decamer isomer, whereas the (551) isomer from inserting one water molecule into the top layer. The families labeled as (443) and (434) are built from the stacked cube $(\text{H}_2\text{O})_{12}$ isomer, which consists of three layers of cyclic tetramers: the (443) isomer can be built by removing one water molecule from the top layer of the global minimum of $(\text{H}_2\text{O})_{12}$ and the (434) isomer by removing one water molecule from the middle layer of the stacked cube structure of $(\text{H}_2\text{O})_{12}$. The MP2 results indicate that the low-lying isomers of the (434) and (515) families are the most probable candidates for the global minimum. Since the maximum MP2 energy difference for all four distinct almost isoenergetic isomers is very small (~ 0.5 kcal/mol), the inclusion of zero-point energy corrections is important to identify the global minimum. When zero-point energy corrections (albeit at the harmonic level) were included, the (515) isomer becomes the global minimum structure of $(\text{H}_2\text{O})_{11}$. Both the TIP4P and TTM3-F potentials predict the (443) isomer as the global minimum.

Figure 12 shows four distinct isomers for $(\text{H}_2\text{O})_{13}$ as the low-lying isomers of $(\text{H}_2\text{O})_{13}$ (Bulusu et al. 2006). Note that the three isomers (414), (454), and (445) can be built from the antiparallel global minimum of $(\text{H}_2\text{O})_{12}$, as shown in Fig. 10. However, the (454-b) isomer, which is the global minimum structure predicted by TIP4P, is built from the parallel $(\text{H}_2\text{O})_{12}$ in accordance with the fact that this is the global minimum structure of $(\text{H}_2\text{O})_{12}$ predicted by TIP4P. For completeness, the (4414) isomer can be built by adding one water molecule vertically between the top and middle layers of the antiparallel isomer of $(\text{H}_2\text{O})_{12}$. The (454) and (454-b) isomers can be built by inserting one water molecule into the middle layer of the antiparallel and parallel $(\text{H}_2\text{O})_{12}$ isomers, respectively. The (445) isomer can be constructed by inserting one water molecule into the top water layer. The MP2 calculations predict that the (4414) isomer is the most stable isomer of the $(\text{H}_2\text{O})_{13}$ cluster and the (454) isomer is the next lowest isomer. Since the maximum

MP2 energy difference for four distinct isoenergetic isomers of $(\text{H}_2\text{O})_{13}$ is very small (~ 0.4 kcal/mol), we furthermore estimate the effect of the zero-point energy correction to the relative stability of the isomers. The (4414) isomer remains as the global minimum structure of $(\text{H}_2\text{O})_{13}$ after including harmonic zero-point energy corrections. The TTM3-F potential predicts the (445) isomer as the global minimum structure of $(\text{H}_2\text{O})_{13}$.

The Transition from “All-Surface” to “Internally Solvated” Clusters at $n = 17$

So far we have discussed in section “[Global Minimum Structures of the \$n = 2\$ – \$10\$ Water Clusters](#)” that the structural “transition” from a quasi-planar ring motif to a more compact, cage-like structure, which occurs at $n = 6$. For the smaller clusters, the need to maximize hydrogen bonding results in all the atoms being on the surface of the cluster. However, as the cluster grows, there is a point where more compact, fully connected structures are formed, in which water molecules are fully coordinated (i.e., they have two donors and two acceptors of hydrogen bonds) like in an average arrangement in the liquid environment. Several experiments (Beuhler and Friedman 1982; Haberland 1984; Hermann et al. 1982; Lin 1973; Nagashima et al. 1986; Niedner-Schatteburg and Bondybey 2000; Schindler et al. 1996; Searcy and Fenn 1974; Stace and Moore 1983; Yang and Castleman 1989) such as expansion of ionized water vapor (Beuhler and Friedman 1982; Searcy and Fenn 1974), ion bombardment of ice surfaces (Haberland 1984), as well as black-body-radiation-induced dissociation studies (Niedner-Schatteburg and Bondybey 2000; Schindler et al. 1996) suggested a prevailed magic cluster of the $\text{H}_3\text{O}^+(\text{H}_2\text{O})_{20}$ cluster consisting of a clathrate-like pentagonal dodecahedron structure. This magic cluster of $\text{H}_3\text{O}^+(\text{H}_2\text{O})_{20}$ indicates the transition from the “all-surface” to “internally solvated” clusters. Note that this kind of transition for the neutral water clusters has not yet been identified experimentally. For a cluster of that size, the number of isomers increases dramatically with size. A detailed discussion of this problem will be presented in the following section “[The Family of Minima for \$\(\text{H}_2\text{O}\)_{20}\$](#) .” In brief, a global search algorithm (Hartke 2003; Li and Scheraga 1987; Wales and Scheraga 1999) has been previously used in order to identify low-energy structures of clusters of that size. That global search approach of an evolutionary algorithm combined with the TTM2-F potential predicted that for the neutral water clusters, the transition from “all-surface” to “internally solvated” occurs for $(\text{H}_2\text{O})_{17}$ (Hartke 2003). Figure 13 shows the alternation between the all-surface and interior minima at the $n = 17$ – 21 cluster regime predicted by the TTM2-F interaction potential. The solvated water molecule with a tetrahedral network is highlighted in green color. The TIP4P potential predicts that the transition from “all-surface” to “internally solvated” clusters occurs for $(\text{H}_2\text{O})_{19}$. Both the TIP4P and TTM3-F potentials predicted the same global minimum structures for $n = 18$ and 20. The lowest energy structure predicted with TIP4P has a slightly different hydrogen bonding orientation for $n = 19$ when compared to the global minimum structure predicted by TTM2-F.

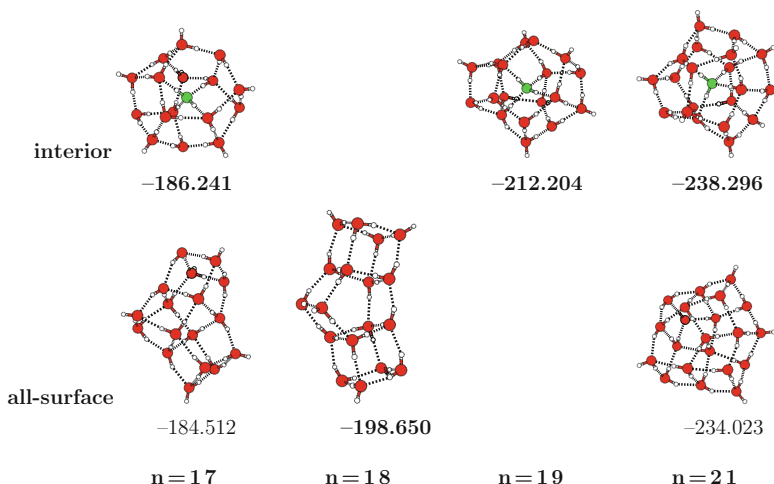


Fig. 13 The transition from “all-surface” to “internally solvated” clusters in the range of $n = 17$ – 21 . The binding energies were calculated at the MP2/aug-cc-pVDZ level of theory. The figure is reprinted with permission from reference Lagutchenkov et al. (2005). © 2005 American Institute of Physics

Validation from Electronic Structure Calculations

Since the TIP4P and TTM3-F classical water potentials predicted different global minimum structures for $n = 17$ and 21 , we concentrated our efforts in comparing the relative stabilities of the two different arrangements predicted by the two models from electronic structure calculations. The results (Lagutchenkov et al. 2005) obtained at the MP2 level of theory with the aug-cc-pVDZ and aug-cc-pVTZ basis sets are shown in Fig. 13. For the $n = 17$ cluster, MP2 calculations with both basis sets predicted the “interior” structure to be more stable, in agreement with the predictions of the TTM2-F potential. Inclusion of differential zero-point energy corrections (0.45 kcal/mol at B3LYP/TZVP level of theory) furthermore tends to increase the energy difference between the two isomers in favor of the “interior” configuration. Our best estimate for the energy difference between the “all-surface” and “interior” isomers of $n = 17$ is 1.43 kcal/mol (1.88 kcal/mol including zero-point energy corrections) at the MP2/aug-cc-pVTZ level of theory. For the $n = 21$ cluster, a similar behavior was found between the most stable “all-surface” and “interior” isomers. The basis set effect at the MP2 level of theory also results in decreasing the energy difference, where as the inclusion of zero-point energy corrections tends to stabilize the “interior” configuration with respect to the “all-surface” one. The TTM2-F potential predicts the correct ordering between the two different arrangements, yielding an energy difference that is almost half of the MP2/aug-cc-pVTZ value. At this level of theory, the “interior” configuration is stabilized by 3.71 kcal/mol (4.11 kcal/mol including zero-point energy corrections) with respect to the “all-surface” one for $n = 21$. Therefore,

the MP2 calculations provided a reliable benchmark that validated the predictions of the TTM2-F potential as regards the onset of the “all-surface” to “interior” configurations occurring at $n = 17$.

Spectroscopic Signature

As noted earlier, the *ab initio* investigation of the relative stability of the various hydrogen bonding networks in the $n = 17$ – 21 cluster regime has identified the existence of a transitional size regime where preferential stabilization alternates between “all-surface” (all atoms on the surface of a cluster) and “internally solvated” (one water molecule at the center of the cluster, fully solvated) configurations with the addition or the removal of a single water molecule.

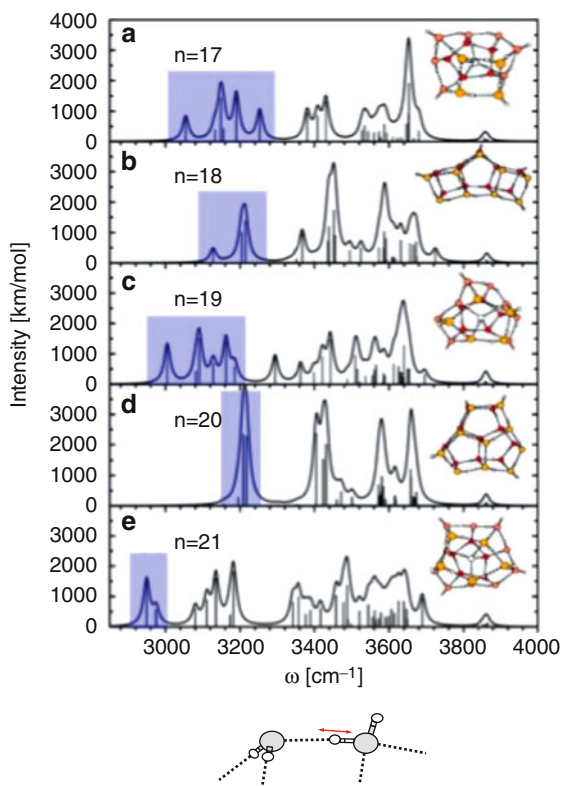
This behavior has been previously suggested based on the results of the TTM2-F interaction potential. It is qualitatively different from the picture that simple, pairwise-additive potentials like TIP4P suggest. The onset of the appearance of the first “interior” configuration in water clusters occurs for $n = 17$. The observed structural alternation between “interior” ($n = 17, 19, 21$) and “all-surface” ($n = 18, 20$) global minima in the $n = 17$ – 21 cluster regime is accompanied by a corresponding spectroscopic signature, namely the undulation in the position of the most red-shifted OH stretching vibrations (indicated by the shaded area in Fig. 14) according to the trend: “interior” configurations exhibit more red-shifted OH stretching vibrations than “all-surface” ones (see Fig. 14). These most red-shifted OH stretching vibrations form distinct groups in the intramolecular region of the spectra and correspond to localized vibrations of donor OH stretches that have the following characteristics: (1) the hydrogen-bonded OH group belongs to a water molecule that has a “free” (non-hydrogen-bonded) OH stretch and (2) the bound OH group acts as a donor to a neighboring water molecule via a water dimer-like hydrogen bond (i.e., a hydrogen bond in which the dihedral angle between the free OH of the donor molecule and the bisector of the HOH angle of the acceptor molecule is 180° , as seen in the lower panel of Fig. 14).

The Family of Minima for $(\text{H}_2\text{O})_{20}$

Interest in the family of water clusters for $n = 20$ stems from the fact that one of the families of minima is the dodecahedron which has been proposed as a model for inclusion compounds and constitutes a building block of type I ice clathrate (we will discuss this in detail in section “[The First Few Water Cages](#)”) The structure of $(\text{H}_2\text{O})_{20}$ has been the subject of debate due to the existence of various major families of minima which are energetically in close proximity and the disagreement as to which was the global minimum (Fanourgakis et al. 2004).

We have previously briefly discussed (see section “[The Transition from ‘All-Surface’ to ‘Internally Solvated’ Clusters at \$n = 17\$](#) ”) the issue regarding the number of isomers increasing dramatically with cluster size. The assignment of the structural

Fig. 14 Intramolecular spectra in the 3,000–4,000 cm^{-1} region associated with the hydrogen-bonded and free OH stretching vibrations for $n = 17$ –21 [(a)–(e)] at the B3LYP/TZVP level of theory. The stick diagram of the computed frequencies is also shown. Shaded area denotes the most red-shifted OH stretching vibrations. The figure is reprinted with permission from reference Lagutchenkov et al. (2005). © 2005 American Institute of Physics



motif of the neutral $(\text{H}_2\text{O})_{20}$ cluster represents a much more difficult task especially in the absence of experimental information (such as rotational constants or infrared spectra) that can yield definitive structural information. For typical clusters of that size, even for a fixed position of the Oxygen atom network, there exists a multitude of arrangements of the hydrogen atoms in a manner consistent with the cluster analog of the Bernal–Fowler ice rules (Hobbs 1974; Petrenko and Whitworth 1999). For instance, just for the dodecahedral arrangement of the oxygen atom network in $(\text{H}_2\text{O})_{20}$, the application of graph theoretical techniques (Kuo et al. 2001; McDonald et al. 1998) yields an estimate of 30,026 symmetry-distinct hydrogen bond topologies which are likely candidates for local minimum structures. For this reason, extensive sampling of the $n = 20$ cluster potential energy surface (PES) has been performed only with classical interaction potentials. Using the “basin-hopping” method combined with the TIP4P potential, Wales and Hodges (Wales and Hodges 1998) identified four major families of $(\text{H}_2\text{O})_{20}$ which, based on the arrangement of the oxygen atom network, they classified as (a) dodecahedron, (b) fused cubes, (c) face-sharing pentagonal prisms, and (d) edge-sharing pentagonal prisms. Representative candidates of these families of minima are illustrated in Fig. 15. The relative energetic order of the four families is predicted to be (d)-(b)-

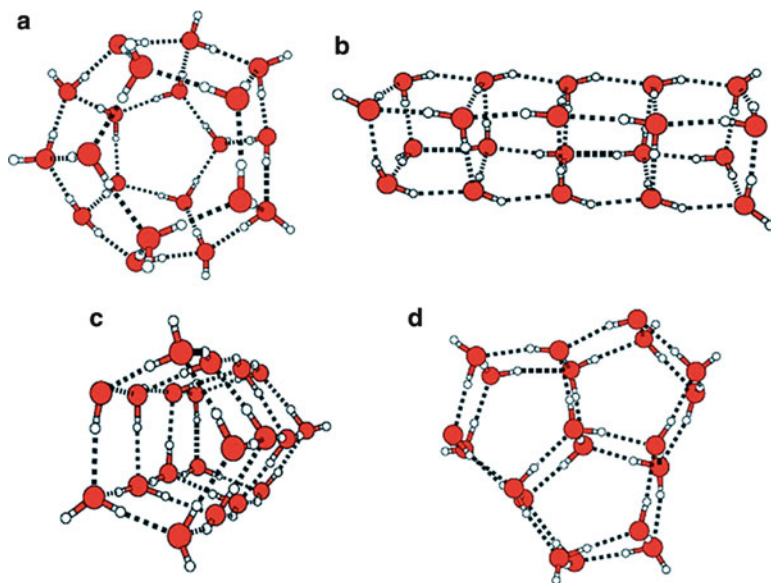


Fig. 15 The lowest energy isomers within the four-families of minima for $(\text{H}_2\text{O})_{20}$: (a) dodecahedron, (b) fused cubes, (c) face-sharing pentagonal prisms, and (d) edge-sharing pentagonal prisms. The figure is reprinted with permission from reference Fanourgakis et al. (2004). © 2004 American Institute of Physics

(c)-(a) with TIP4P. An evolutionary algorithm used by Hartke in conjunction with TTM2-F interaction potentials resulted in the edge-sharing pentagonal prism as the global minimum. The global minimum structure is in fact qualitatively identical for TTM2-F and TIP4P, and it is an all-surface one. The edge-sharing pentagonal prism (d) is lower by 11.2 (TIP4P) and by 13.98 kcal/mol (TTM2-F) than the dodecahedral family (a).

The Four Major Families of Minima

First principles electronic structure calculations at the MP2 level of theory with the family of augmented correlation-consistent basis sets up to quadruple zeta quality were used to estimate complete basis set (CBS) binding energies D_e for the lowest energy structures within each of the four low-lying families of minima as follows: (a) dodecahedron (−200.1 kcal/mol, 30 hydrogen bonds), (b) fused cubes (−212.6 kcal/mol, 36 hydrogen bonds), (c) face-sharing pentagonal prisms (−215.0 kcal/mol, 35 hydrogen bonds), and (d) edge-sharing pentagonal prisms (−217.9 kcal/mol, 34 hydrogen bonds) (Fanourgakis et al. 2004). Among the various interaction potentials, TTM2-F (Burnham and Xantheas 2002b) was found to predict the absolute cluster binding energies to within 1% from the corresponding

MP2/CBS values. These calculations also yielded the first harmonic vibrational spectra of the lowest energy isomers within each family as a means of providing the necessary information for their future spectroscopic identification.

Zero-point energy corrections do not change the relative order of the four different networks, although they reduce the separation between the highest-lying dodecahedron structure and the lowest-lying edge-sharing pentagonal prism isomer from 17.8 kcal/mol to 15.0 kcal/mol.

Vibrational Spectra

The IR spectra of the four low-lying families of minima of $(\text{H}_2\text{O})_{20}$ exhibit different features in the OH stretching region that can be used for their spectroscopic assignment. These are shown in Fig. 16. Some of these spectral features are reminiscent of their constituent water tetramer and pentamer fragments. The dodecahedral network is associated with the most red-shifted (by 700–800 cm^{-1} with respect to the monomer) OH stretching vibrations among the four isomers and shows the richest structure in the 3,000–4,000 cm^{-1} range. The lowest edge-sharing pentagonal prisms family of isomers displays IR intense bands that are red-shifted by $\sim 600 \text{ cm}^{-1}$ with respect to the average of the symmetric and antisymmetric OH stretching vibrations in the water monomer. A common feature of the vibrational spectra that is shared by all four networks is that the most red-shifted, IR-active OH stretching vibrations exhibit the same characteristics previously found for the most stable isomers of $n = 17\text{--}21$ (Fig. 14), namely that they belong to fragments that have a “free” OH stretch and act as donors to a neighboring water molecule along a hydrogen bond which resembles a gas-phase dimer arrangement. It is conceivable that the second characteristic directly affects the energetic stability of the isomer within the corresponding family.

The computed spectra can be used as guides for the spectroscopic assignment of the lowest energy edge-sharing pentagonal prism isomer in molecular beams or the higher-lying networks in other environments such as helium droplets or cryogenic para-hydrogen solids that can facilitate the formation of higher energy cluster isomers.

The First Few Water Cages

Significance

Gas hydrates are naturally occurring in situ in the deep oceans and permafrost regions of the earth (Mao et al. 2007; Sloan 1998). Their potential importance as inclusion compounds relevant to energy problems such as hydrogen storage has been recently highlighted. All of the clathrate hydrate forming natural gases adopt one of the following three crystal structures (Jeffrey et al. 1967; Ripmeester et al. 1987): (1) sI (cubic, $Pm\bar{3}n$), (2) sII (cubic, $Fd\bar{3}m$) and (3) sH (hexagonal, $P6/mmm$). These

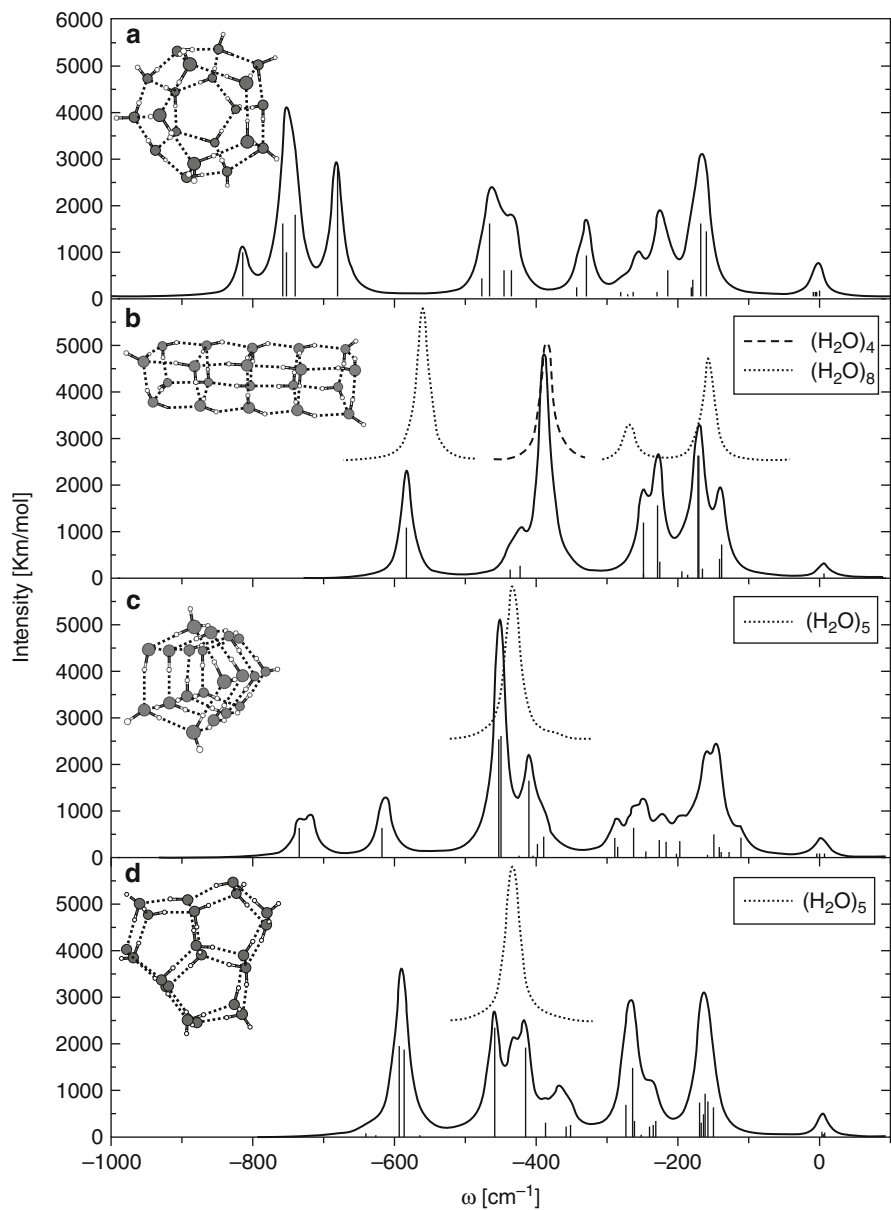


Fig. 16 Intramolecular spectra in the $3,000\text{--}4,000\text{ cm}^{-1}$ region associated with the OH stretching vibrations. The stick diagram of the frequencies is also shown. The IR spectra for the water ring tetramer, pentamer, and octamer cube are shown for comparison. The figure is reprinted with permission from reference Fanourgakis et al. (2004). © 2004 American Institute of Physics

three-dimensional lattices are constructed from the following water cages: (1) the pentagonal dodecahedron (D-cage), which consists of 20 water molecules that form 12 pentagonal faces (5^{12}); (2) the tetrakaidecahedron (T-cage), which consists of 24 water molecules that form 12 pentagonal and 2 hexagonal faces ($5^{12}6^2$); (3) the hexakaidecahedron (H-cage), which consists of 28 water molecules that form 12 pentagonal and 4 hexagonal faces ($5^{12}6^4$); (4) the irregular dodecahedron, which consists of 20 water molecules that form 3 tetragonal, 6 pentagonal, and 3 hexagonal faces ($4^35^66^3$); and (5) the icosahedron, that consists of 36 water molecules that form 12 pentagonal and 8 hexagonal faces ($5^{12}6^8$). In the previous notation, $p^n q^m$ identifies the underlying Oxygen atom network of the constituent cages, with p , q being the size of the polygons and n , m the number of faces; for instance, $5^{12}6^2$ denotes 12 pentagonal and 2 hexagonal faces. The unit cell of the three-dimensional (sI) hydrate periodic structure is made of 2 units of the 5^{12} cage and 6 units of the $5^{12}6^2$ cage. In a similar manner, the unit cell of the (sII) hydrate lattice is constructed from 16 units of the 5^{12} cage and 8 units of the $5^{12}6^4$ cage. This hydrate lattice (sII) has been recently considered for practical applications in energy storage as it was shown to meet current US Department of Energy's target densities for an onboard hydrogen storage system (Mao et al. 2006; Schüth 2005). Finally, the unit cell of the (sH) hydrate is more complex since it is constructed from 3 units of the 5^{12} cage, 2 units of the $4^35^66^3$ cage, and 1 unit of the $5^{12}6^8$ cage. The constituent cages of those three hydrate lattices are often used as models to probe the relevant guest/host interactions from electronic structure calculations to derive intermolecular potentials that can be then used for the full system of the periodic lattice and the molecular guests.

The Pentagonal Dodecahedron (D-cage)(H₂O)₂₀ Cluster

The analysis of the relative energetics of the various dodecahedron isomers is performed using the previous results (Wales and Hodges 1998) of Wales and Hodges as the starting point. These authors have used the basin-hopping Monte Carlo method (Li and Scheraga 1987; Wales and Scheraga 1999) in conjunction with the classical TIP4P (Jorgensen et al. 1983) potential in order to obtain the global and low-lying local minima of the pentagonal dodecahedron family of isomers. The list of the first 19 minima obtained with the TIP4P potential by Wales and Hodges were used as starting geometries for optimization with the TTM2-F (Burnham and Xantheas 2002b, c; Burnham et al. 1999) interaction potential (see Fig. 17). In that figure, the global minimum obtained with TIP4P for the dodecahedron corresponds to minimum #2 in the list and is indicated by an arrow. In general, there is a good correspondence between the TIP4P and TTM2-F relative isomer energetics, except for the isomer indicated as minimum "0" in the list. That isomer labeled by "0" is the one suggested as the global minimum from the combinatorial optimization with the "strong" and "weak" hydrogen bonds (SWB) model (Kirov 1996) when the intermolecular interactions were computed with the ST2 (Stillinger and Rahman 1974) interaction potential.

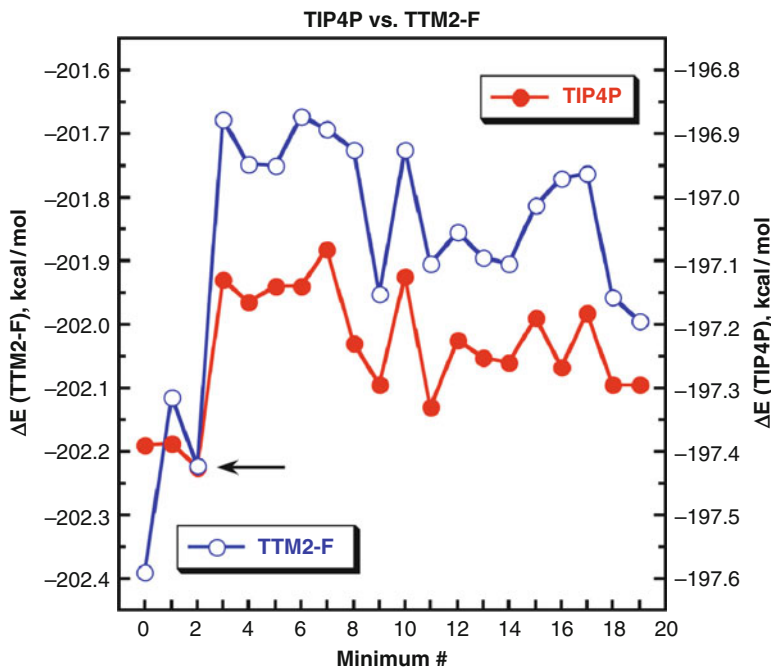


Fig. 17 Binding energies of the first 19 isomers of the pentagonal dodecahedron (H_2O)₂₀ cluster with the TIP4P (filled circles) and TTM2-F (open circles) potentials. The TIP4P global minimum corresponds to minimum #2 and is indicated with an arrow. Minimum “0” is the isomer predicted from the SWB model. The figure is reprinted with permission from reference Kirov et al. (2008). © 2008 Elsevier

In order to further assess the validity of the predictions of the empirical interaction potentials regarding the relative stability of isomers for the pentagonal dodecahedron, we have carried out electronic structure calculations for the 20 minima that are included in Fig. 17 (Kirov et al. 2008). The TTM2-F optimal structures (which are qualitatively similar to the ones obtained by TIP4P) for the 20 minima labeled 0–19 were used as starting points in the geometry optimization at both the density functional (DFT) (Hohenberg and Kohn 1964; Kohn and Sham 1965) and MP2 (Møller and Plesset 1934) levels. In the DFT calculations, Becke’s gradient-corrected exchange-correlation density functional (B3LYP) (Becke 1988, 1993) was used in conjunction with Ahlrichs’ polarized triple- ζ (TZVP) basis set (Schäfer et al. 1994). As for the previous cases, the MP2 optimizations were performed using Dunning’s (Dunning 1989; Kendall et al. 1992) augmented correlation-consistent basis sets of double- ζ (aug-cc-pVDZ) quality.

The binding energies of these 20 isomers of the dodecahedron (H_2O)₂₀ are shown in Fig. 18 at the B3LYP/TZVP (open circles) and MP2/aug-cc-pVDZ (filled squares) levels of theory (Kirov et al. 2008). The order of isomers has been maintained the same as in Fig. 16 for a direct comparison. We first note

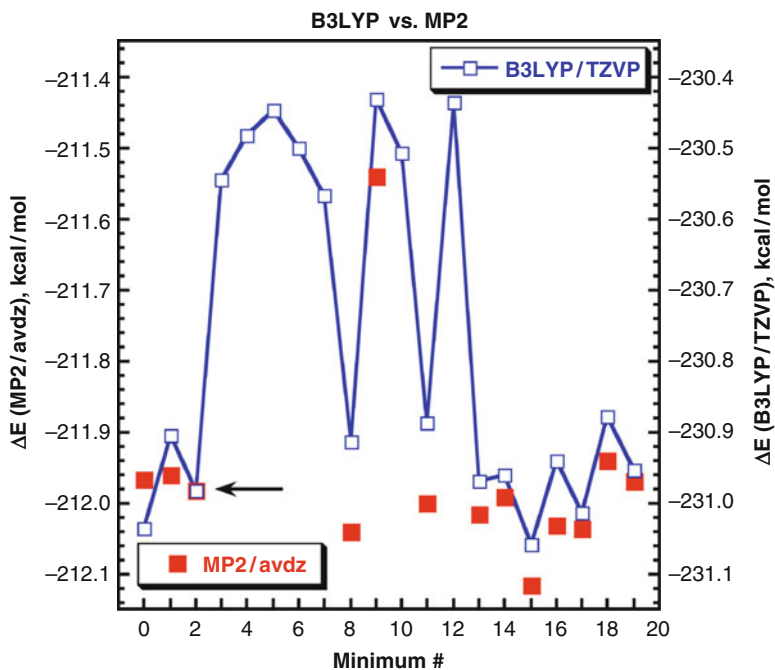


Fig. 18 Binding energies of the first 19 isomers of the pentagonal dodecahedron (H_2O)₂₀ cluster at the B3LYP/TZVP (open squares) and MP2/avg-cc-pVDZ (filled squares) levels of theory. The TIP4P global minimum corresponds to minimum #2 and is indicated with an arrow. Minimum “0” is the isomer predicted from the SWB model. The figure is reprinted with permission from reference Kirov et al. (2008). © 2008 Elsevier

that B3LYP/TZVP also predicts minimum “0” to be more stable than minimum #2 (the global dodecahedron minimum of TIP4P model) as was the case with TTM2-F potential. However, it yields an even lower energy structure (minimum #15). The energetic stabilization of minimum #15 is also confirmed during subsequent MP2/avg-cc-pVDZ optimizations. Based on the results of the DFT and MP2 calculations, we therefore suggest minimum #15 as the global minimum for the dodecahedron family of (H_2O)₂₀. Its binding energy (D_e) is -231.01 kcal/mol (B3LYP/TZVP) and -212.10 kcal/mol (MP2/avg-cc-pVDZ). The three-dimensional structure and its two-dimensional mapping onto the corresponding Schlegel diagram (Schlegel 1883) for the newly found global minimum are shown in Fig. 19. Schlegel diagrams assist in categorizing the various isomers by offering a helpful two-dimensional visual representation of the connectivity of the underlying hydrogen bonding network. Their construction is based on the principle that the nodes and lines of the graph correspond to the vertices and edges of the corresponding polyhedron. The hydrogen bonding network is naturally preserved upon this dimensional reduction from the 3D structure (left) to the 2D graph (right). In order to better indicate the correspondence between the two representations, we also indicate the numbering of the oxygen atoms in Fig. 19.

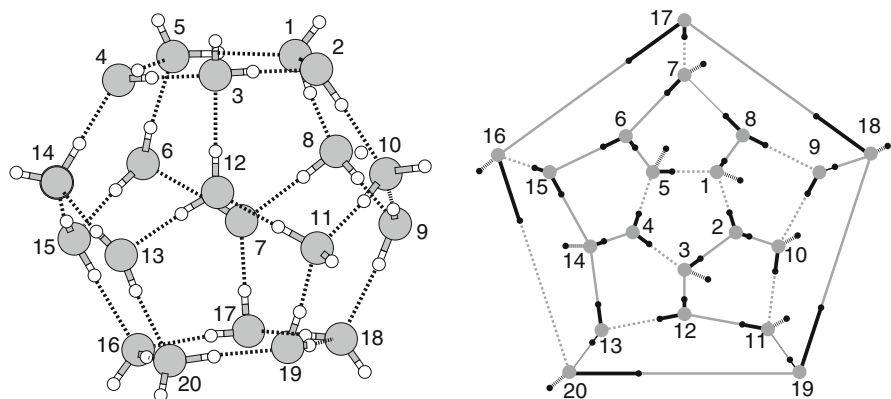


Fig. 19 The lowest energy isomer of the dodecahedron predicted at the MP2 level of theory. Left panel: 3D structure. Right Panel: Schlegel diagram indicating the “strong” (dimer-like, solid lines) and “weak” (dashed lines) nearest-neighbor (n–n) molecular arrangements. The figure is reprinted with permission from reference Kirov et al. (2008). © 2008 Elsevier

The Tetrakaidecahedron (T-Cage)(H₂O)₂₄ Cluster

For a fixed position of the Oxygen atoms in the $5^{12}6^2$ (H₂O)₂₄ cage cluster, there are 3,043,836 symmetry distinct configurations arising from the different hydrogen positions that are consistent with the ice rules. The application of the SWEB discrete model with subsequent refinement at the DFT and MP2 levels of theory yields the first few (~100) low-lying networks. The Schlegel diagrams (Schlegel 1883) (2D projections of the 3D structures) of the ten lowest minima of the T-cage (fully optimized at the MP2/aug-cc-pVDZ level of theory) are shown in Fig. 20 (Yoo et al. 2009). The various networks are labeled using the R or L designation, depending on whether the direction of the donor H atoms in the three concentric rings starting from the inner one is either clockwise (R) or counterclockwise (L). For example, the RLR isomer corresponds to an arrangement in which the H atoms are arranged in a clockwise (for the inner and outer circles) and a counterclockwise direction (for the middle circle). Primes are used to indicate different arrangements of the vertical H atoms that connect the three rings. Therefore, the L'LR network has the same connectivity and direction as LLR for the three rings but differs in the direction of the vertical H atoms. The ten lowest-lying minima of the T-cage can be furthermore grouped into three families (indicated as I, II and III in Fig. 20) depending on the donor–acceptor connectivity of the top hexagonal face. For this six-member hydrogen-bonded ring, isomers in family (I) isomers have only donor–acceptor (da) water molecules, whereas isomers in families (II) and (III) have both double donor (dd) and double acceptor (aa) water molecules. The (dd)/(aa) molecules are indicated with blue and gray colors in Fig. 20. It should be noted that all ten minima shown in Fig. 20 have a zero total dipole moment. Their relative order at the MP2/aug-cc-pVDZ level of theory with respect to the most stable LLR network of family (I) (I-LLR) is identical with the larger aug-cc-pVTZ basis set at the MP2 level of theory.

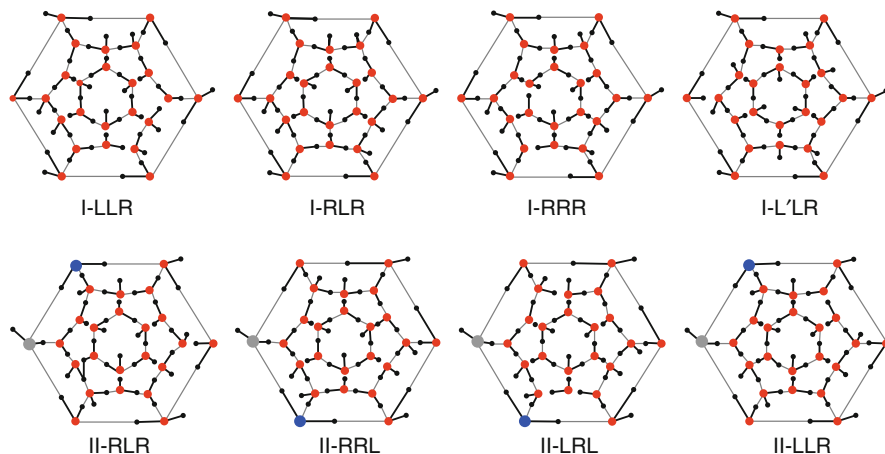


Fig. 20 Schlegel diagrams of the ten lowest minima of the T-cage $(\text{H}_2\text{O})_{24}$ cluster obtained at the MP2/aug-cc-pVDZ level of theory. The figure is reprinted with permission from reference Yoo et al. (2009). © 2009 American Chemical Society

Starting from any of those low-lying isomers of the T-cage the three-dimensional periodic lattice of the (sI) hydrate can now be constructed. In the 46-molecule unit cell of the (sI) lattice two adjacent T-cages are surrounded by six D-cages. When two T-cage isomers sharing a common hexagonal face are used as the building blocks, the positions of the hydrogen atoms of the rest of 36 molecules ($2 \times 24 - 6 - 6$) can be determined. In this manner, the possible networks of the (sI) hydrate unit cell are dramatically reduced. The proposed “bottom-up” approach, which starts from the constituent cages and builds up the three-dimensional lattice, can be used to construct three-dimensional unit cells of the (sI) hydrate from the low-energy networks of their constituent cages. In Fig. 21 the process of constructing a periodic unit cell of the (sI) hydrate lattice starting from the lowest-lying I-LLR isomer is illustrated.

Outlook

The study of small- to medium-sized water clusters offers an important, molecular-level insight into the fundamental interactions between water molecules. It provides information about the nature and magnitude of the cooperative effects that govern the structural and dynamic properties of the more complex aqueous environments. The structures of the first few water clusters are controlled by the need to maximize the amount of hydrogen bonding and as such they are inhomogeneous and under-coordinated, resulting in structures in which all atoms are on the surface of the cluster. This hydrogen bonding environment is very different than the on-the-average homogeneous, fully coordinated one that is found in aqueous systems.

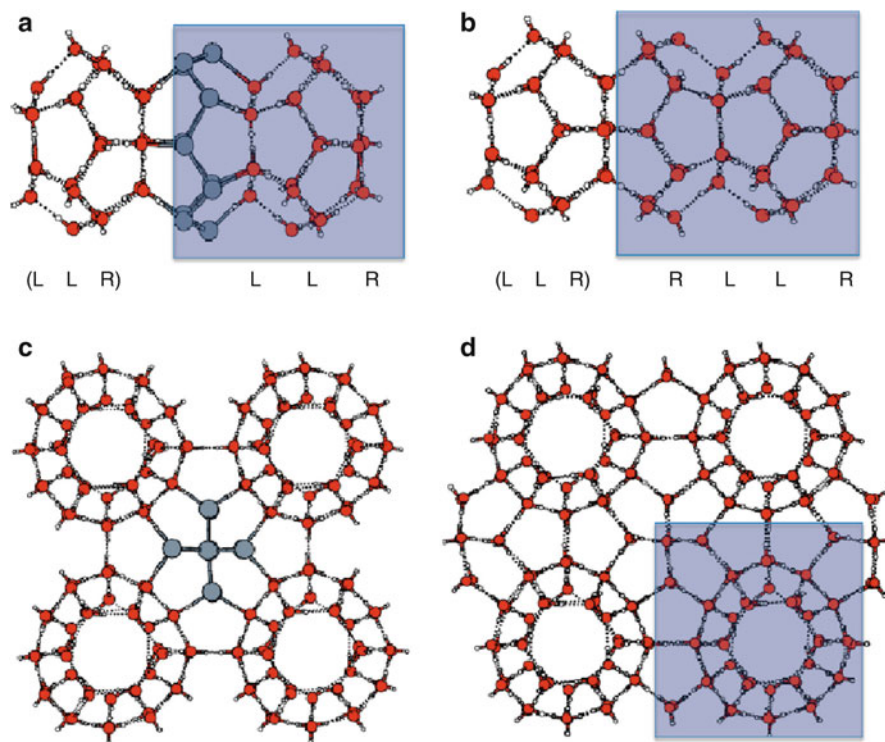


Fig. 21 Schematic approach for building the (sI) hydrate lattice from the low-lying networks of the T-cage. The connectivity of the O atoms between the two cages (without hydrogen atoms) is highlighted in gray. The blue box indicates the (sI) cubic unit cell (46 water molecules). The figure is reprinted with permission from reference Yoo et al. (2009). © 2009 American Chemical Society

As the cluster grows there is a point where fully coordinated molecules in the inside on the cluster are observed. The onset of this transition from the “all-surface” to “fully coordinated” molecules occurs at $n = 17$. Classical force fields for water are essential in probing the plethora of isomers especially for the larger ($n > 21$) clusters. They can be used in conjunction with new approaches for the efficient sampling of the configuration space in order to obtain the low-lying isomers of clusters in this size regime. The structures of the small- and medium-sized water clusters are essential in order to assess the accuracy of those force fields in developing transferable models that can be used in different environments such as liquid water, ice, at aqueous interfaces, or around charged species or hydrophobic surfaces. Given the fact that accurate energetics for the interaction even between two water molecules is not currently available experimentally, the use of high-level first principles electronic structure calculations to obtain the cluster energetics represents an indispensable and currently irreplaceable path toward understanding aqueous environments.

Acknowledgments This work was supported by the Division of Chemical Sciences, Geosciences and Biosciences, Office of Basic Sciences, U.S. Department of Energy. Battelle operates the Pacific Northwest National Laboratory for the U.S. Department of Energy. This research was performed in part using the Molecular Science Computing Facility (MSCF) in the Environmental Molecular Sciences Laboratory, a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research. Additional computer resources were provided by the Office of Basic Energy Sciences, US Department of Energy at the National Energy Research Scientific Computing Center, a U.S. Department of Energy's Office of Science user facility at Lawrence Berkeley National Laboratory.

Bibliography

- Adamovic, I., Freitag, M. A., & Gordon, M. S. (2003). Density functional theory based effective fragment potential method. *Journal of Chemical Physics*, *118*, 6725–6732.
- Apra, E., Rendell, A. P., Harrison, R. J., Tippraju, V., deJong, W. A., & Xantheas, S. S. (2009). Liquid water: Obtaining the right answer for the right reasons. In Proceedings of the conference on high performance computing networking, storage and analysis. Portland: ACM.
- Ball, P. (2008). Water: Water-an enduring mystery. *Nature*, *452*, 291–292.
- Bartlett, R. J., & Purvis, G. D. (1978). Many-body perturbation theory, coupled-pair many-electron theory, and the importance of quadruple excitation for the correlation problem. *International Journal of Quantum Chemistry*, *14*, 561–581.
- Becke, A. D. (1988). Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, *38*, 3098–3100.
- Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *Journal of Chemical Physics*, *98*, 5648.
- Berendsen, H. J. C., Grigera, J. R., & Straatsma, T. P. (1987). The missing term in effective pair potentials. *Journal of Physical Chemistry*, *91*, 6269–6271.
- Beuhler, R. J., & Friedman, L. (1982). A study of the formation of high molecular-weight water cluster ions ($m/e < 59000$) in expansion of ionized gas mixtures. *Journal of Chemical Physics*, *77*, 2549–2557.
- Boys, S. F., & Bernardi, F. (1970). Calculation of small molecular interactions by differences of separate total energies – some procedures with reduced errors. *Molecular Physics*, *19*, 553.
- Bukowski, R., Szalewicz, K., Groenenboom, G. C., & van der Avoird, A. (2007). Predictions of the properties of water from first principles. *Science*, *315*, 1249–1252.
- Bulusu, S., Yoo, S., Aprà, E., Xantheas, S., & Zeng, X. C. (2006). Lowest-energy structures of water clusters $(\text{H}_2\text{O})_{11}$ and $(\text{H}_2\text{O})_{13}$. *Journal of Physical Chemistry A*, *110*, 11781–11784.
- Bunge, C. F. (1970). Electronic wave functions for atom. II. Some aspects of convergence of configuration interaction expansion for ground states of He isoelectronic series. *Theoretica Chimica Acta*, *16*, 126.
- Burnham, C. J., & Xantheas, S. S. (2002a). Development of transferable interaction models for water. I. Prominent features of the water dimer potential energy surface. *Journal of Chemical Physics*, *116*, 1479–1492.
- Burnham, C. J., & Xantheas, S. S. (2002b). Development of transferable interaction models for water. IV. A flexible, all-atom polarizable potential (TTM2-F) based on geometry dependent charges derived from an ab initio monomer dipole moment surface. *Journal of Chemical Physics*, *116*, 1479–1492.
- Burnham, C. J., & Xantheas, S. S. (2002c). Development of transferable interaction models for water. III. Reparametrization of an all-atom polarizable rigid model (TTM2-R) from first principles. *Journal of Chemical Physics*, *116*, 1500–1510.
- Burnham, C. J., Li, J., Xantheas, S. S., & Leslie, M. (1999). The parametrization of a Thole-type all-atom polarizable water model from first principles and its application to the study of water clusters ($n = 2-21$) and the phonon spectrum of ice Ih. *Journal of Chemical Physics*, *110*, 4566–4581.

- Burnham, C. J., Anick, D. J., Mankoo, P. K., & Reiter, G. F. (2008). The vibrational proton potential in bulk liquid water and ice. *Journal of Chemical Physics*, *128*, 154519.
- Chalasinski, G., & Szczesniak, M. M. (1994). Origins of structure and energetics of van-der-waals clusters from ab-initio calculations. *Chemical Reviews*, *94*, 1723–1765.
- Cizek, J. (1966). On correlation problem in atomic and molecular systems. Calculation of wavefunction components in ursell-type expansion using quantum-field theoretical methods. *Journal of Chemical Physics*, *45*, 4256.
- Cizek, J. (1969). *Advances in Chemical Physics*, *14*, 35.
- Clementi, E. (1967). Study of electronic structure of molecules. II. Wavefunctions for $\text{NH}_3 + \text{HCl} \rightarrow \text{NH}_4\text{Cl}$ reaction. *Journal of Chemical Physics*, *46*, 3851–3880.
- Coester, F. (1958). Bound states of a many-particle system. *Nuclear Physics*, *7*, 421–424.
- Coester, F., & Kümmel, K. (1960). Short-range correlations in nuclear wave functions. *Nuclear Physics*, *17*, 477–485.
- Cruzan, J. D., Braly, L. B., Brown, M. G., Loeser, J. G., & Saykally, R. J. (1996a). Quantifying hydrogen bond cooperativity in water: VRT spectroscopy of the water tetramer. *Science*, *271*, 59–62.
- Cruzan, J. D., Brown, M. G., Liu, K., Braly, L. B., & Saykally, R. J. (1996b). The far-infrared vibration-rotation-tunneling spectrum of the water tetramer-d8. *Journal of Chemical Physics*, *105*, 6634–6644.
- Cruzan, J. D., Viant, M. R., Blake, G. A., & Saykally, R. J. (1997). Terahertz laser vibration-rotation tunneling spectroscopy of the water tetramer. *Journal of Physical Chemistry A*, *101*, 9022–9031.
- Dang, L. X., & Chang, T. M. (1997). Molecular dynamics study of water clusters, liquid, and liquid-vapor interface of water with many-body potentials. *Journal of Chemical Physics*, *106*, 8149–8159.
- Day, P. N., Jensen, J. H., Gordon, M. S., Webb, S. P., Stevens, W. J., Krauss, M., Garmer, D., Basch, H., & Cohen, D. (1996). An effective fragment method for modeling solvent effects in quantum mechanical calculations. *Journal of Chemical Physics*, *105*, 1968–1986.
- Day, P. N., Pachter, R., Gordon, M. S., & Merrill, G. N. (2000). A study of water clusters using the effective fragment potential and Monte Carlo simulated annealing. *Journal of Chemical Physics*, *112*, 2063–2073.
- Dunning, T. H., Jr. (1989). Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *Journal of Chemical Physics*, *90*, 1007–1023.
- Dunning, T. H., Jr. (2000). A road map for the calculation of molecular binding energies. *Journal of Physical Chemistry A*, *104*, 9062–9080.
- Dunning, T. H., Jr., Peterson, K. A., & Woon, D. E. (1998). *Encyclopedia of computational chemistry*. New York: Wiley.
- Eggenberger, R., Gerber, S., Huber, H., & Searles, D. (1991). Basis set superposition errors in intermolecular structures and force-constants. *Chemical Physics Letters*, *183*, 223–226.
- Emsley, J., Hoyte, O. P. A., & Overill, R. E. (1978). Ab initio calculations on very strong hydrogen-bond of biformate anion and comparative esterification studies. *Journal of the American Chemical Society*, *100*, 3303–3306.
- Fanourgakis, G. S., & Xantheas, S. S. (2006). The flexible, polarizable, thole-type interaction potential for water (TTM2-F) revisited. *Journal of Physical Chemistry A*, *110*, 4100–4106.
- Fanourgakis, G. S., & Xantheas, S. S. (2008). Development of transferable interaction potentials for water. V. Extension of the flexible, polarizable, Thole-type model potential (TTM3-F, v. 3.0) to describe the vibrational spectra of water clusters and liquid water. *Journal of Chemical Physics*, *128*, 074506.
- Fanourgakis, G. S., Aprà, E., & Xantheas, S. S. (2004). High-level ab initio calculations for the four low-lying families of minima of $(\text{H}_2\text{O})_{20}$. I. Estimates of MP2/CBS binding energies and comparison with empirical potentials. *Journal of Chemical Physics*, *121*, 2655–2663.
- Fanourgakis, G. S., Aprà, E., de Jong, W. A., & Xantheas, S. S. (2005). High-level ab initio calculations for the four low-lying families of minima of $(\text{H}_2\text{O})_{20}$. II. Spectroscopic signatures

- of the dodecahedron, fused cubes, face-sharing pentagonal prisms, and edge-sharing pentagonal prisms hydrogen bonding networks. *Journal of Chemical Physics*, *122*, 134304.
- Fast, P. L., Sanchez, M. L., & Truhlar, D. G. (1999). Infinite basis limits in electronic structure theory. *Journal of Chemical Physics*, *111*, 2921–2926.
- Feller, D. (1992). Application of systematic sequences of wave-functions to the water dimer. *Journal of Chemical Physics*, *96*, 6104–6114.
- Gordon, M. S., Freitag, M. A., Bandyopadhyay, P., Jensen, J. H., Kairys, V., & Stevens, W. J. (2001). The effective fragment potential method: A QM-based MM approach to modeling environmental effects in chemistry. *Journal of Physical Chemistry A*, *105*, 293–307.
- Haberland, H. (1984). *Electronic and atomic collisions*. New York: Elsevier.
- Halkier, A., Klopper, W., Helgaker, T., Jorgensen, P., & Taylor, P. R. (1999). Basis set convergence of the interaction energy of hydrogen-bonded complexes. *Journal of Chemical Physics*, *111*, 9157–9167.
- Hartke, B. (2003). Size-dependent transition from all-surface to interior-molecule structures in pure neutral water clusters. *Physical Chemistry Chemical Physics*, *5*, 275–284.
- Hermann, V., Kay, B. D., & Castleman, A. W., Jr. (1982). Evidence for the existence of structures in gas-phase homomolecular clusters of water. *Chemical Physics*, *72*, 185–200.
- Hobbs, P. V. (1974). *Ice physics*. Oxford: Clarendon.
- Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. *Physical Review*, *136*, B864.
- Hura, G., Sorenson, J. M., Glaeser, R. M., & Head-Gordon, T. (2000). A high-quality x-ray scattering experiment on liquid water at ambient conditions. *Journal of Chemical Physics*, *113*, 9140–1948.
- Jeffrey, G. A., Jordan, T. H., & McMullan, R. K. (1967). Clathrate hydrates of some amines. *Science*, *155*, 689.
- Jensen, J. H., & Gordon, M. S. (1998). An approximate formula for the intermolecular Pauli repulsion between closed shell molecules. II. Application to the effective fragment potential method. *Journal of Chemical Physics*, *108*, 4772–4782.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, *79*, 926–935.
- Kazimirski, J. K., & Buch, V. (2003). Search for low energy structures of water clusters (H₂O)_n, n = 20–22, 48, 123, and 293. *Journal of Physical Chemistry*, *107*, 9762–9775.
- Kendall, R. A., Simons, J., Gutowski, M., & Chalasinski, G. (1989). Ab initio energy and structure of H⁻(H₂)₂. *Journal of Physical Chemistry*, *93*, 621–625.
- Kendall, R. A., Dunning, T. H., Jr., & Harrison, R. J. (1992). Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *Journal of Chemical Physics*, *96*, 6796–6806.
- Kirov, M. V. (1996). Conformational combinatorial analysis of polyhedral water clusters. *Journal of Structural Chemistry*, *37*, 84–91.
- Kirov, M. V., Fanourgakis, G. S., & Xantheas, S. S. (2008). Identifying the most stable networks in polyhedral water clusters. *Chemical Physics Letters*, *461*, 108–188.
- Klopper, W. (1995). Limiting values for Møller-Plesset second-order correlation energies of polyatomic systems: A benchmark study on Ne, HF, H₂O, N₂, and He...He. *Journal of Chemical Physics*, *102*, 6168–6179.
- Koga, K., Parra, R. D., Tanaka, H., & Zeng, X. C. (2000). Ice nanotube: What does the unit cell look like? *Journal of Chemical Physics*, *113*, 5037–5040.
- Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Physical Review*, *140*, A1133.
- Kucharski, S. A., & Bartlett, R. J. (1992). The coupled-cluster single, double, triple, and quadruple excitation method. *Journal of Chemical Physics*, *97*, 4282–4288.
- Kuo, J.-L., Coe, J. V., & Singer, S. J. (2001). On the use of graph invariants for efficiently generating hydrogen bond topologies and predicting physical properties of water clusters and ice. *Journal of Chemical Physics*, *114*, 2527–2540.

- Lagutchenkov, A., Fanourgakis, G. S., Niedner-Schatteburg, G., & Xantheas, S. S. (2005). The spectroscopic signature of the "all-surface" to "internally solvated" structural transition in water clusters in the $n = 17 - 21$ size regime. *Journal of Chemical Physics*, *122*, 194310.
- Leclercq, J. M., Allavena, M., & Bouteiller, Y. (1983). On the basis set superposition error in potential surface investigations. I. Hydrogen-bonded complexes with standard basis set-functions. *Journal of Chemical Physics*, *78*, 4606-4611.
- Lee, C., Chen, H., & Fitzgerald, G. (1995). Chemical bonding in water clusters. *Journal of Chemical Physics*, *122*, 1266-1269.
- Lee, H. M., Suh, S. B., & Kim, K. S. (2001). Structures, energies, and vibrational spectra of water undecamer and dodecamer: An ab initio study. *Journal of Chemical Physics*, *114*, 10749-10756.
- Li, Z., & Scheraga, H. A. (1987). Monte-Carlo-minimization approach to the multiple-minima problem in protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, *84*, 6611-6615.
- Lin, S. S. (1973). Detection of large water clusters by a low re-quadrupole mass filter. *Review of Scientific Instruments*, *44*, 516.
- Liu, B., & McLean, A. D. (1973). Accurate calculation of attractive interaction of two ground-state helium-atoms. *Journal of Chemical Physics*, *59*, 4557-4558.
- Liu, K., Loeser, J. G., Elrod, M. J., Host, B. C., Rzepiela, J. A., Pugliano, N., & Saykally, R. J. (1994). Dynamics of structural rearrangements in the water trimer. *Journal of the American Chemical Society*, *116*, 3507-3512.
- Liu, K., Brown, M. G., Cruzan, J. D., & Saykally, R. J. (1996a). Vibration-rotation tunneling spectra of the water pentamer: Structure and dynamics. *Science*, *271*, 62-64.
- Liu, K., Brown, M. G., Carter, C., Saykally, R. J., Gregory, J. K., & Clary, D. C. (1996b). Characterization of a cage form of the water hexamer. *Nature*, *381*, 501-503.
- Liu, K., Brown, M. G., & Saykally, R. J. (1997a). Terahertz laser vibration rotation tunneling spectroscopy and dipole moment of a cage form of the water hexamer. *Journal of Physical Chemistry A*, *101*, 8995-9010.
- Liu, K., Brown, M. G., Cruzan, J. D., & Saykally, R. J. (1997b). Terahertz laser spectroscopy of the water pentamer: Structure and hydrogen bond rearrangement dynamics. *Journal of Physical Chemistry A*, *101*, 9011-9021.
- Mao, W. L., Mao, H.-K., Meng, Y., Eng, P. J., Hu, M. Y., Chow, P., Cai, Y. Q., Shu, J., & Hemley, R. J. (2006). X-ray-induced dissociation of H_2O and formation of an O_2-H_2 alloy at high pressure. *Science*, *314*, 636-638.
- Mao, W. L., Koh, C. A., & Sloan, E. D. (2007). Clathrate hydrates under pressure. *Physics Today*, *60*(10), 42.
- Martin, J. M. L. (1996). Ab initio total atomization energies of small molecules - Towards the basis set limit. *Chemical Physics Letters*, *259*, 669-678.
- Mayer, I., & Surjan, P. R. (1992). Monomer geometry relaxation and the basis set superposition error. *Chemical Physics Letters*, *191*, 497-499.
- McDonald, S., Ojamae, L., & Singer, S. J. (1998). Graph theoretical generation and analysis of hydrogen-bonded structures with applications to the neutral and protonated water cube and dodecahedral clusters. *Journal of Physical Chemistry A*, *102*, 2824-2832.
- Møller, C., & Plesset, M. S. (1934). Note on an approximation treatment for many-electron systems. *Physical Review*, *46*, 618-622.
- Nagashima, U., Shinohara, H., & Tanaka, H. (1986). Enhanced stability of ion clathrate structures for magic number water clusters. *Journal of Chemical Physics*, *84*, 209-214.
- Netzloff, H. M., & Gordon, M. S. (2004). The effective fragment potential: Small clusters and radial distribution functions. *Journal of Chemical Physics*, *121*, 2711-2714.
- Niedner-Schatteburg, G., & Bondybey, V. E. (2000). FT-ICR studies of solvation effects in ionic water cluster reactions. *Chemical Reviews*, *100*, 4059-4086.
- Nigra, P., & Kais, S. (1999). Pivot method for global optimization: A study of water clusters $(H_2O)_N$ with $2 \leq N \leq 33$. *Chemical Physics Letters*, *305*, 433-438.

- Peterson, K. A., & Dunning, T. H., Jr. (1995). Benchmark calculations with correlated molecular wave-functions. 7. Binding-energy and structure of the HF dimer. *Journal of Chemical Physics*, *102*, 2032–2041.
- Petrenko, V. F., & Whitworth, R. W. (1999). *Physics of ice*. New York: Oxford University Press.
- Pugliano, N., & Saykally, R. J. (1992). Measurement of quantum tunneling between chiral isomers of the cyclic water trimer. *Science*, *257*, 1937–1940.
- Purvis, G. D., & Bartlett, R. J. (1982). A full coupled-cluster singles and doubles model – The inclusion of disconnected triples. *Journal of Chemical Physics*, *76*, 1910–1917.
- Raghavachari, K., Trucks, G. W., Pople, J. A., & Head-Gordon, M. (1989). A 5th-order perturbation comparison of electron correlation theories. *Chemical Physics Letters*, *157*, 479–483.
- Raghavachari, K., Pople, J. A., Replogle, E. S., & Head-Gordon, M. (1990). 5th-order Møller-plesset perturbation-theory - comparison of existing correlation methods and implementation of new methods correct to 5th-order. *Journal of Physical Chemistry*, *94*, 5579.
- Ripmeester, J. A., Tse, J. S., Ratcliffe, C. I., & Powell, B. M. (1987). A new clathrate hydrate structure. *Nature*, *325*, 135–136.
- Robinson, G. W., Zhu, S.-B., Singh, S., & Evans, M. W. (1996). *Water in biology, chemistry and physics. Experimental overviews and computational methodologies*. Singapore: World Scientific.
- Sadlej, J. (2001). Theoretical study of structure and spectra of cage clusters (H₂O)_n, n = 11, 12. *Chemical Physics Letters*, *333*, 485–492.
- Schäfer, A., Huber, C., & Ahlrichs, R. (1994). Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *Journal of Chemical Physics*, *100*, 5829.
- Schindler, T., Berg, C., Niedner-Schatteburg, G., & Bondybey, V. E. (1996). Protonated water clusters and their black body radiation induced fragmentation. *Chemical Physics Letters*, *250*, 301–308.
- Schlegel, V. (1883). Verh. Kais. Leopold.-Carolin. Dtsch. Akad. Naturforsch, *44*, 343.
- Schüth, F. (2005). Technology: Hydrogen and hydrates. *Nature*, *434*, 712–713.
- Searcy, J. Q., & Fenn, J. B. (1974). Clustering of water on hydrated protons in supersonic free jet expansion. *Journal of Chemical Physics*, *61*, 5282–5288.
- Sloan, E. D., Jr. (1998). *Clathrate hydrates of natural gases* (2nd ed.). New York: Marcel Dekker.
- Smit, P. H., Derissen, J. L., & van Duijneveldt, F. B. (1978). Role of distortion energy in ab initio calculated dimerization energy of formic acid. *Journal of Chemical Physics*, *69*, 4241–4244.
- Soper, A. K. (2000). The radial distributions of water and ice from 220 to 673 K and at pressure up to 400 Mpa. *Chemical Physics*, *258*, 121–137.
- Sorenson, J. M., Hura, G., Glaeser, R. M., & Head-Gordon, T. (2000). What can x-ray scattering tell us about the radial distribution functions of water? *Journal of Chemical Physics*, *113*, 9149–1961.
- Stace, A. J., & Moore, C. (1983). A correlation between structure and reactivity in ion clusters. *Chemical Physics Letters*, *96*, 80–84.
- Stillinger, F. H., & Rahman, A. (1974). Molecular-dynamics study of liquid water under high pressure. *Journal of Chemical Physics*, *60*, 1545.
- Suzuki, S., & Blake, G. A. (1994). Pseudorotation in the D₂O trimer. *Chemical Physics Letters*, *229*, 499.
- Szalewicz, K., Leforestier, C., & van der Avoird, A. (2009). Towards the complete understanding of water by a first-principles computational approach. *Chemical Physics Letters*, *482*, 1–14.
- Termath, V., Klopper, W., & Kutzelnigg, W. (1991). Wave-functions with terms linear in the interelectronic coordinates to take care of the correlation cusp. II. second-order Møller-Plesset (MP2–R12) calculations on closed-shell atoms. *Journal of Chemical Physics*, *94*, 2002–2019.
- Tsai, C. J., & Jordan, K. D. (1993). Theoretical study of small water clusters: Low-energy fused cubic structures for (H₂O)_n, n = 8, 12, 16, and 20. *Journal of Physical Chemistry*, *97*, 5208–5210.
- van Duijneveldt, F. B., van Duijneveldt-van de Rijdt, J. G. C. M., & van Lenthe, J. H. (1994). State of the art in counterpoise theory. *Chemical Reviews*, *94*, 1873.

- van Duijneveldt-van de Rijdt, J. G. C. M., & van Duijneveldt, F. B. (1992). Convergence to the basis-set limit in abinitio calculations at the correlated level on the water dimer. *Journal of Chemical Physics*, *97*, 5019.
- van Lenthe, J. H., van Duijneveldt-van de Rijdt, J. G. C. M., & van Duijneveldt, F. B. (1987). *Advances in Chemical Physics*, *69*, 521.
- Viant, M. R., Cruzan, J. D., Lucas, D. D., Brown, M. G., Liu, K., & Saykally, R. J. (1997). Pseudorotation in water trimer isotopomers using terahertz laser spectroscopy. *Journal of Physical Chemistry A*, *101*, 9032–9041.
- Wales, D. J., & Hodges, M. P. (1998). Global minima of water clusters $(\text{H}_2\text{O})_n$, $n \leq 21$, described by an empirical potential. *Chemical Physics Letters*, *286*, 65–72.
- Wales, D. J., & Scheraga, H. A. (1999). Global optimization of clusters, crystals, and biomolecules. *Science*, *285*, 1368–1372.
- Wilson, A. K., & Dunning, T. H., Jr. (1997). Benchmark calculations with correlated molecular wave functions. X. Comparison with “exact” MP2 calculations on Ne, HF, H₂O, and N₂. *Journal of Chemical Physics*, *106*, 8718–8726.
- Xantheas, S. S. (1994). Ab initio studies of cyclic water clusters $(\text{H}_2\text{O})_n$, $n = 1-6$. II. Analysis of many-body interactions. *Journal of Chemical Physics*, *100*, 7523–7534.
- Xantheas, S. S. (1996a). Significance of higher-order many-body interaction energy terms in water clusters and bulk water. *Philosophical Magazine Part B*, *73*, 107–115.
- Xantheas, S. S. (1996b). On the importance of the fragment relaxation energy terms in the estimation of the basis set superposition error correction to the intermolecular interaction energy. *Journal of Chemical Physics*, *104*, 8821–8824.
- Xantheas, S. S. (2000). Cooperativity and hydrogen bonding network in water clusters. *Chemical Physics*, *258*, 225–231.
- Xantheas, S. S., & Aprà, E. (2004). The binding energies of the D_{2d} and S₄ water octamer isomers: High-level electronic structure and empirical potential results. *Journal of Chemical Physics*, *120*, 823–828.
- Xantheas, S. S., & Dunning, T. H., Jr. (1993a). The structure of the water trimer from ab initio calculations. *Journal of Chemical Physics*, *98*, 8037–8040.
- Xantheas, S. S., & Dunning, T. H., Jr. (1993b). Ab initio studies of cyclic water clusters $(\text{H}_2\text{O})_n$, $n = 1-6$. I. Optimal structures and vibrational spectra. *Journal of Chemical Physics*, *99*, 8774–8792.
- Xantheas, S. S., & Dunning, T. H., Jr. (1993c). Theoretical estimate of the enthalpy of formation of sulfhydryl radical (HSO) and HSO-SOH isomerization energy. *Journal of Physical Chemistry*, *97*, 18–19.
- Xantheas, S. S., Burnham, C. J., & Harrison, R. J. (2002). Development of transferable interaction models for water. II. Accurate energetics of the first few water clusters from first principles. *Journal of Chemical Physics*, *116*, 1493–1499.
- Yang, X., & Castleman, A. W. (1989). Large protonated water clusters $\text{H}^+(\text{H}_2\text{O})_n$ ($1 \leq n < 60$): The production and reactivity of clathrate-like structures under thermal conditions. *Journal of the American Chemical Society*, *111*, 6845–6846.
- Yoo, S., Kirov, M. V., & Xantheas, S. S. (2009). Low-energy networks of the T-cage $(\text{H}_2\text{O})_{24}$ cluster and their use in constructing periodic unit cells of the structure I (SI) hydrate lattice. *Journal of the American Chemical Society*, *131*, 7564–7566.

Fundamental Structural, Electronic, and Chemical Properties of Carbon Nanostructures: Graphene, Fullerenes, Carbon Nanotubes, and Their Derivatives

27

Tandabany C. Dinadayalane and Jerzy Leszczynski

Contents

Introduction to Carbon Nanostructures	1176
Graphene	1178
Fullerenes	1180
Natural Abundance of Fullerenes	1188
Fullerene Nano-capsules	1189
Isolated Pentagon Rule (IPR) in Fullerenes	1190
Common Defects in Fullerenes	1193
Carbon Nanotubes (CNTs)	1195
Various Defects in Carbon Nanotubes	1200
Computational Approaches Used to Study Carbon Nanostructures: An Overview	1202
Structural, Electronic, and Chemical Properties of Graphene, Fullerenes, and SWCNTs ..	1205
Graphene	1205
Hydrogenation of Graphene With and Without Defects	1206
Fullerenes	1210
Giant Fullerenes	1214
Local Strain in Curved Polycyclic Systems: POAV and Pyramidalization Angle	1215
Stone–Wales Defect in C ₆₀	1218
Computational Studies on Vacancy Defects in Fullerene C ₆₀	1219
Computational Studies of Single-Walled Carbon Nanotubes	1221
Covalent Functionalization of SWCNTs: H and F Atom Chemisorptions	1225
Theoretical Studies on Common Defects in SWCNTs	1231
Stone–Wales Defect	1232
Topological Ring Defects	1235
Single- and Di-vacancy Defects	1236

T.C. Dinadayalane (✉)

Department of Chemistry, Clark Atlanta University, Atlanta, GA, USA

e-mail: dina@icnanotox.org

J. Leszczynski

Department of Chemistry and Biochemistry, Interdisciplinary Center for Nanotoxicity, Jackson State University, Jackson, MS, USA

e-mail: jerzy@icnanotox.org

© Springer International Publishing Switzerland 2017

J. Leszczynski et al. (eds.), *Handbook of Computational Chemistry*,

DOI 10.1007/978-3-319-27282-5_22

1175

Outlook of Potential Applications of Carbon Nanostructures	1238
Summary and Outlook	1242
Bibliography	1243

Abstract

This chapter provides information on various carbon allotropes and in-depth details of structural, electronic, and chemical properties of graphene, fullerenes, and single-walled carbon nanotubes (SWCNTs). We have written an overview of different computational methods that were employed to understand various properties of carbon nanostructures. Importance of application of computational methods in exploring different sizes of fullerenes and their isomers is given. The concept of isolated pentagon rule (IPR) in fullerene chemistry has been revealed. The computational and experimental studies involving Stone–Wales (SW) and vacancy defects in fullerene structures are discussed in this chapter. The relationship between the local curvature and the reactivity of the defect-free and defective fullerene and single-walled carbon nanotubes has been revealed. We reviewed the influence of different defects in graphene on hydrogen addition. The viability of hydrogen and fluorine atom additions on the external surface of the SWCNTs is revealed using computational techniques. We have briefly pointed out the current utilization of carbon nanostructures and their potential applications.

Introduction to Carbon Nanostructures

Carbon is one of the first few elements known in antiquity. The pure forms of this element include diamond and graphite, which have been known for few thousand years (<http://www.nndc.bnl.gov/content/elements.html>; Pierson 1993; Wikipedia – <http://en.wikipedia.org/wiki/Carbon>). Both of these materials are of immense importance in industry and in everyday life. Diamond and graphite are termed as giant structures since, by means of a powerful microscope, one could see millions and millions of atoms, all connected together in a regular array. Diamond would appear as a rigid and rather complex system like some enormous scaffolding construction. Carbon is also the major atomic building block for life. All life-forms on Earth have carbon central to their composition. More than 10 million carbon-containing compounds are known. Compounds containing only carbon atoms, particularly nano-sized materials, are intriguing and attract attention of scientists working in various disciplines. Before 1985, scientists deemed that there were only three allotropes of carbon, namely, diamond, graphite, and amorphous carbon such as soot and charcoal. Soccer ball-shaped molecule comprising of 60 carbon atoms, C₆₀ buckyball named fullerene, was discovered in 1985, and it is another interesting carbon allotrope (Kroto et al. 1985). Carbon nanotubes (CNTs), a spin-off product of fullerene, were reported in 1991 by Iijima (1991). Important well-known carbon materials are depicted in Fig. 1. The publication of transmission electron microscope (TEM) images of CNTs by Iijima was a critical factor in

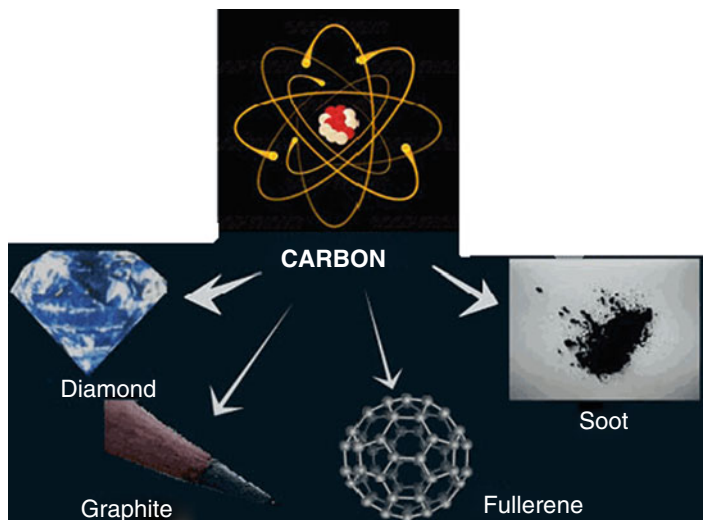


Fig. 1 Well-known carbon materials

convincing a broad community that “*there is plenty of room at the bottom*” and many new structures can exist at the nanoscale. Figure 2 shows eight allotropes of carbon. In addition to graphene, fullerenes, and carbon nanotubes, there are few other uncommon carbon nanostructures such as nanohorns (Iijima et al. 1999; Poonjarernsilp et al. 2009), nano-onions (Palkar et al. 2008; Zhou et al. 2009), nanobuds (He and Pan 2009; Nasibulin et al. 2007a, b; Wu and Zeng 2009), peapods (Launois et al. 2010; Li et al. 2009a; Smith et al. 1998), nanocups (Chun et al. 2009), and nanotori (Liu et al. 1997; Sano et al. 2001).

We performed a quick search in SciFinder on “fullerene,” “carbon nanotubes,” and “graphene” to reveal their importance and growth in current science, engineering, and technology. We received a total of nearly 60,000 references for the word “fullerene,” ~145,000 references for “carbon nanotubes,” and 110,000 for “graphene” when we searched these topics in November 2015. This is indicative that carbon nanomaterials have gained a momentum with the development of nanotechnology as the driving force of the modern science and engineering. Among various carbon nanostructures, CNTs play a special role in the nanotechnology era. The design and discovery of new materials is always exciting for the potential of new applications and properties (Cohen 1993; Serra et al. 1999). In this chapter, we aim to present an overview of carbon nanostructures, with a particular interest on structural, electronic, and chemical properties of graphene, fullerenes, and carbon nanotubes. Important topological defects in the graphene, fullerenes, and carbon nanotubes will be delineated. Thus, this chapter is intended to be an informative guide of carbon nanostructures and to provide description of current computational chemistry applications involving these species to facilitate the pursuit of both newcomers to this field and experienced researchers in this rapidly emerging area.

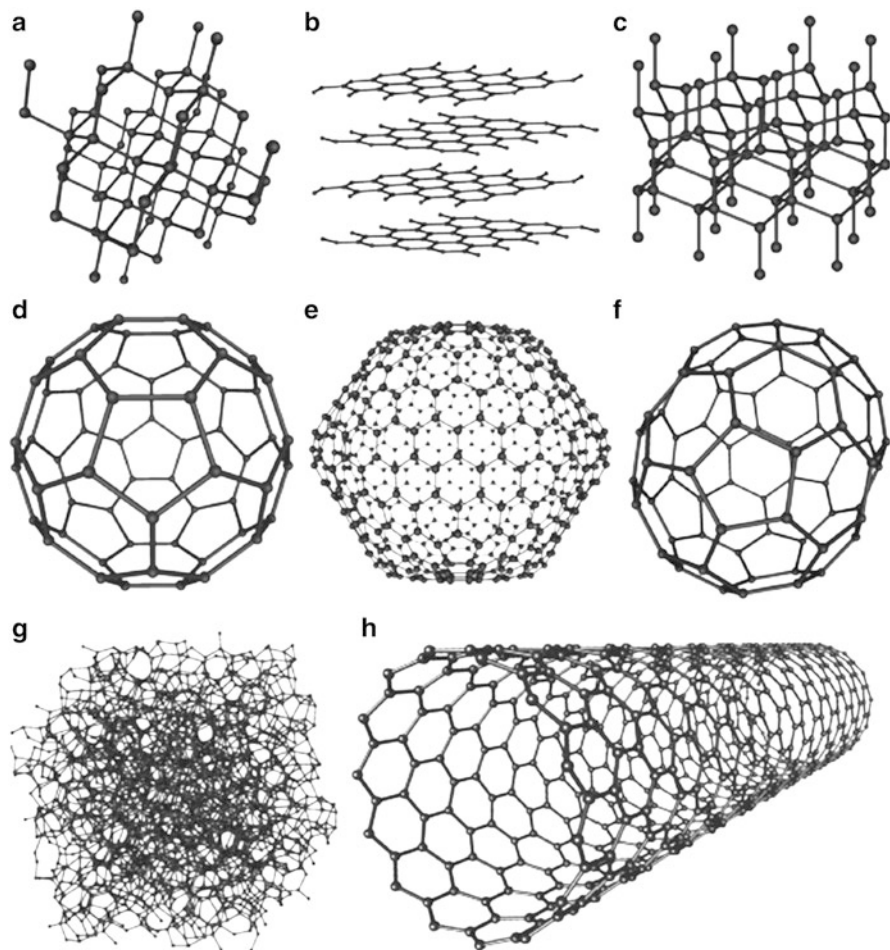


Fig. 2 Eight allotropes of carbon: (a) diamond, (b) graphite, (c) lonsdaleite, (d) C₆₀ (buckminsterfullerene or buckyball), (e) C₅₄₀, (f) C₇₀, (g) amorphous carbon, and (h) single-walled carbon nanotube or buckytube (The picture adopted from Wikipedia – http://en.wikipedia.org/wiki/Allotropes_of_carbon)

Graphene

Carbon displays a unique feature of making a chemically stable two-dimensional (2D), one-atom-thick membrane called graphene in a three-dimensional (3D) world. Each carbon atom in graphene is covalently bonded to three other carbon atoms with sp^2 hybridization. Graphene is the thinnest known material and in the same time is the strongest material ever to be measured. It can sustain current densities six orders of magnitude higher than that of copper. It has extremely high strength and very high thermal conductivity and stiffness and is impermeable to gases (Geim

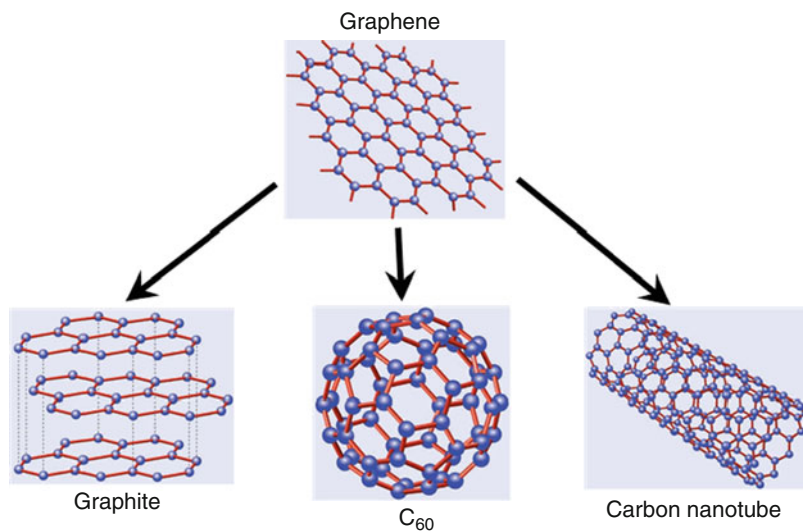


Fig. 3 Carbon-containing molecules (graphite, buckminsterfullerene (C₆₀), and carbon nanotube) derived from graphene

2009). There are many challenges and opportunities for graphene research because graphene is not a standard solid state material. It should be noted that electrons in graphene do not behave in the same way as in ordinary metals and semiconductors due to the unusual energy–momentum relation (Neto 2010). Well-known forms of carbon-containing molecules that derived from graphene are graphite, fullerene, and carbon nanotube, which are depicted in Fig. 3. Graphite consists of stacked layers of graphene sheets separated by 0.3 nm and is stabilized by weak van der Waals forces (He and Pan 2009). Buckminsterfullerene (C₆₀) is formed from graphene balled into a sphere by including some pentagons and hexagons into the lattice (Kroto et al. 1985). The combined experimental and computational study showed the direct transformation of graphene to fullerene (Chuvilin et al. 2010). Carbon nanotubes can be viewed as rolled-up cylinders of graphene. Therefore, graphene can be called “the mother” of all these three sp² carbon structures.

It was presumed that planar graphene cannot exist in free state since they are unstable compared to curved structures such as soot, nanotubes, and fullerenes. This presumption has changed since Novoselov et al. prepared graphitic sheets including single graphene layer and studied their electronic properties (Novoselov et al. 2004, 2005a). The detailed information of growth and isolation of graphene has been provided in the recent review by Geim (2009). Graphene is a prospective material for nanoelectronics. The electron transport in graphene is described by Dirac-like equation (Geim and Novoselov 2007; Novoselov et al. 2005b; Ponomarenko et al. 2008). The experimental realization of graphene motivates several studies focusing on fundamental physics, materials science, and device applications (Abanin et al. 2006; Geim and Novoselov 2007; Novoselov et al. 2004; 2005a, b; Pereira et al.

2009; Ponomarenko et al. 2008). The studies pertinent to the chemistry of graphene sheets have also been reported (Abanin et al. 2006; Avouris et al. 2007; Geim and Novoselov 2007, Neto et al. 2009; Pereira et al. 2009). Graphene research is a hot topic in this decade, thanks to the recent advances in technology for growth, isolation, and characterization of graphene.

Graphene sheets need not always be as perfect as one thinks. Various defects such as Stone–Wales (SW) (Stone and Wales 1986), vacancies (Carlson and Scheffler 2006), pore defects (Jiang et al. 2009), and substitution atoms (Miwa et al. 2008; Zhu et al. 2005) can occur in the thin graphene sheet. Like the creation of vacancies by knocking atoms out of the graphene sheet, surplus atoms can be found as adatoms on the graphene surface. Ad-dimer defect can be introduced to graphene and is characterized by two adjacent five-membered rings instead of two adjacent seven-membered rings in Stone–Wales defect. Therefore, ad-dimer defect is called inverse Stone–Wales (ISW) defect. Figure 4 depicts some of the common defects in graphene sheet.

Experimental observations of defects in graphene have been reported recently (Meyer et al. 2008; Wang et al. 2008). Zettl and coworkers showed the direct image of Stone–Wales defects in graphene sheets using transmission electron microscopy (TEM) and explored their real-time dynamics. They found that the dynamics of defects in extended, two-dimensional graphene membranes are different than in closed-shell graphenes such as nanotubes or fullerenes (Meyer et al. 2008). High-resolution transmission electron microscopy (HRTEM) and atomic force microscopy (AFM) have been useful in identifying various defects in graphene. AFM and HRTEM images of graphene sheet with different defects are shown in Fig. 5. The effect of various defects on the physical and chemical properties of graphene was studied theoretically (Boukhvalov and Katsnelson 2008; Carpio et al. 2008; Duplock et al. 2004; Lherbier et al. 2008; Li et al. 2005). The characteristics of typical defects and their concentrations in graphene sheets are unclear.

Computational and experimental studies concerning defects in graphene sheet are critically important for basic understanding of this novel system, and such understanding will be helpful for scientists who actively work on applications of graphene-based materials. Although the surface physics of graphene sheets is currently at the center of attention, its chemistry has remained largely unexplored. Like any other molecule, graphene can involve in chemical reactions. The chemical functionalization is probably one of the best approaches to detect imperfections in a graphene sheet (Boukhvalov and Katsnelson 2008). The functionalized graphene can be suitable for specific applications. Research on bended, folded, and scrolled graphene is rapidly growing now.

Fullerenes

Discovery of fullerene C_{60} and other fullerene molecules is discussed below.

The fullerene era started in 1985; Kroto and his colleagues obtained cold carbon clusters when they carried out an experiment to simulate the condition of red giant

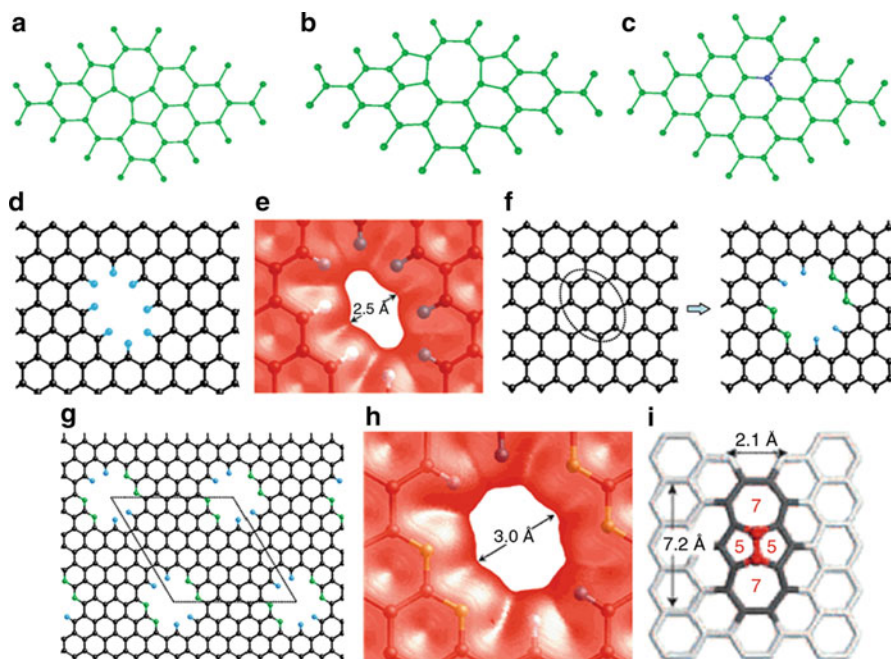


Fig. 4 Defects in graphene sheet; the segment of graphene containing (a) the Stone–Wales (SW) defect; (b) a bivacancy; (c) a nitrogen substitution impurity; (d) an all-hydrogen saturated pore in graphene; (e) the pore electron density isosurface of all-hydrogen passivated porous graphene; (f) creation of a nitrogen-functionalized pore within a graphene sheet: the carbon atoms in the dotted circle are removed, and four dangling bonds are saturated by hydrogen, while the other four dangling bonds together with their carbon atoms are replaced by nitrogen atoms; (g) the hexagonally ordered porous graphene. The *dotted lines* indicate the unit cell of the porous graphene; (h) the pore electron density isosurface of nitrogen-functionalized porous graphene; (i) an inverse Stone–Wales (ISW) defect. Color code for (d), (f), and (g): C black, N green, H cyan. Isosurface is at $0.02 e/\text{Å}^3$ (The pictures were reprinted with permission from references Jiang et al. (2009) and Boukhalov and Katsnelson (2008). Copyright 2008 and 2009 American Chemical Society)

star formation. With the use of the mass spectrometer, they found a large peak commensurate with 60 carbon atoms (Kroto et al. 1985). The molecule C_{60} was proposed to have a football structure, known to mathematicians as the truncated icosahedron. The shape is composed of 12 pentagons located around the vertices of an icosahedron and 20 hexagon rings placed at the centers of icosahedral faces. The C_{60} molecule was named “buckminsterfullerene” in honor of the renowned architect Buckminster Fuller, who designed geodesic domes based on similar pentagonal and hexagonal structures. The carbon atoms in C_{60} fullerene are arranged in exactly the same way, albeit much smaller, as the patches of leather found on the common football (Fig. 6a).

Since the remarkable discovery of fullerenes in 1985 (Kroto et al. 1985), these new carbon allotropes have received significant attention from the scientific

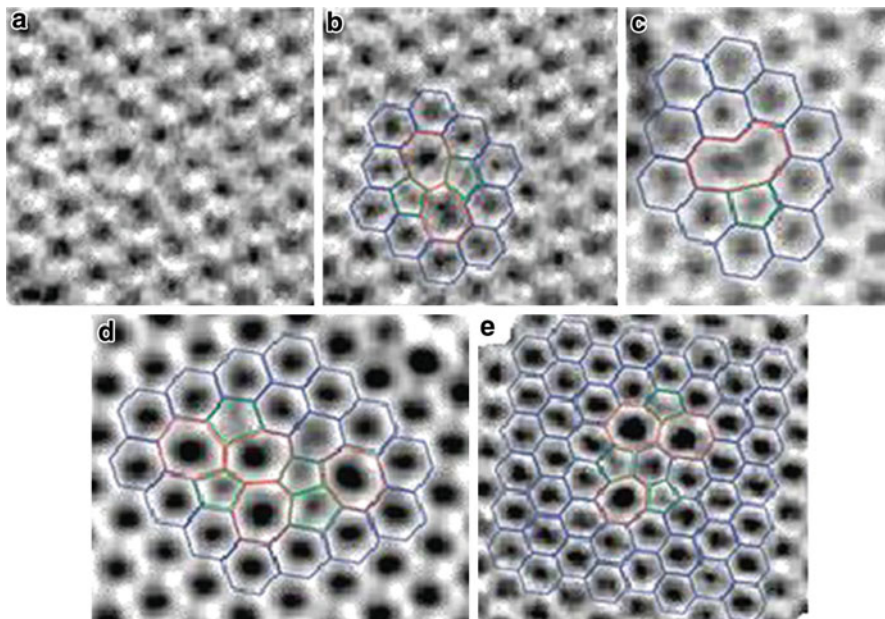


Fig. 5 (a) HRTEM image of a single graphene layer (atoms appear white). (b) Image of graphene with Stone–Wales defect (atomic configuration superimposed for easy recognition). (c) Image of vacancy defect with atomic configuration. (d) Defect image with atomic configuration consisting of four pentagons (*green*) and four heptagons (*red*). (e) Defect image with atomic configuration consisting of three pentagons (*green*) and three heptagons (*red*) (Pictures were reprinted with permission from reference Meyer et al. (2008). Copyright 2008 American Chemical Society)

community and still exhibit vast interest (Lu and Chen 2005; Thilgen and Diederich 2006). The 1996 Nobel Prize in Chemistry was awarded to Sir Harold W. Kroto, Robert F. Curl, and the late Richard E. Smalley for their discovery of fullerenes. Essentially, the most prominent representative of the fullerene family is C_{60} . In early 1990, a method was discovered for producing macroscopic amounts of this fascinating molecule (Krättschmer et al. 1990). This breakthrough allowed scientists to explore the properties of C_{60} and understand its chemistry. Krättschmer et al. characterized the fullerene C_{60} using mass spectroscopy, infrared spectroscopy, electron diffraction, and X-ray diffraction (Krättschmer et al. 1990). Both Kroto et al. (1985) and Krättschmer et al. (1990), by means of mass spectroscopy, also characterized the fullerene C_{70} . Pure C_{60} and C_{70} fullerenes were isolated and separated by Kroto and coworkers (Taylor et al. 1990). The stable fullerenes of C_{60} and C_{70} were reported in the ratio of approximately 5:1. ^{13}C nuclear magnetic resonance (NMR) spectroscopy was used to characterize the fullerenes (Taylor et al. 1990). These two molecules are members of a homologous series of hollow closed-cage molecules. The fullerene C_{70} belongs to a class of nonspherical fullerenes. It adopts an ellipsoidal shape (point group D_{5h}) and it looks like a “rugby ball” as

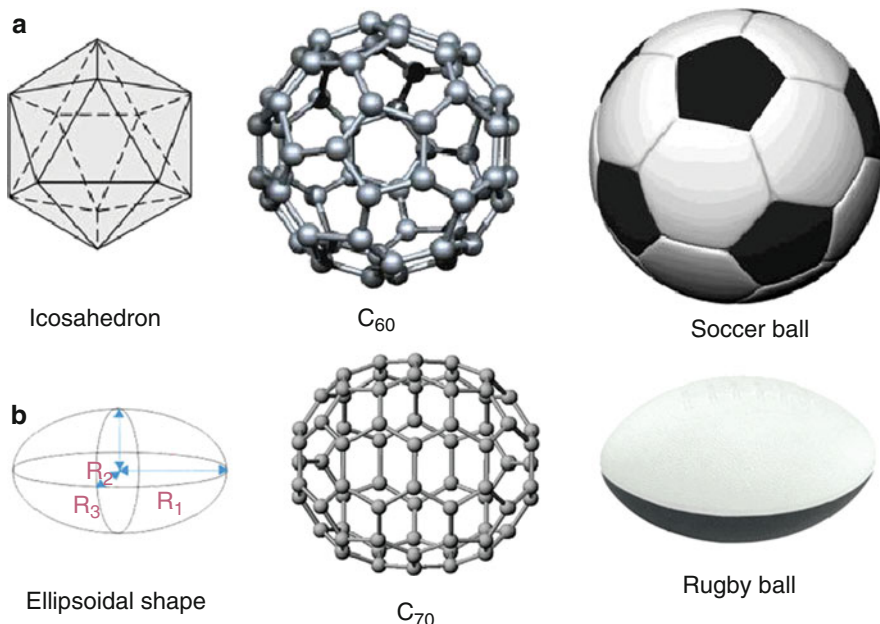


Fig. 6 The structures of fullerenes C_{60} and C_{70} and their familiar shapes. C_{60} and C_{70} are in icosahedron (I_h) and D_{5h} point groups

shown in Fig. 6b. Existence of C_{60} was predicted by Eiji Osawa (1970). However, his prediction did not reach Europe or America since it was published in Japanese magazine.

In the eighteenth century, the Swiss mathematician Leonhard Euler demonstrated that a geodesic structure must contain 12 pentagons to close into a spheroid, although the number of hexagons may vary. Later research by Smalley and his colleagues showed that there should exist an entire family of these geodesic-dome-shaped carbon clusters (Kroto et al. 1985). Fullerenes form with an even number $n \geq 20$ of three connected vertices, $3n/2$ edges, 12 pentagonal faces, and $(n-20)/2$ hexagonal faces (Fowler and Manolopoulos 1995; Kroto et al. 1985). Thus, C_{60} has 20 hexagons, whereas its “rugby ball”-shaped cousin C_{70} has 25 hexagons. As hexagons are added or removed, the molecule begins to lose its roundness. Giant fullerenes take on a pentagonal shape. Smaller fullerenes look like asteroids. One should note that all of the fullerenes have the same Gaussian curvature sign; therefore all of them have a convex surface. The buckminsterfullerene C_{60} , shown in Fig. 6, has a spherical-like shape and the full group of symmetry of the icosahedron (I_h), which means that it could be rotated by the angle of $2\pi/5$ around the center of each pentagon and reflected in the mirror located on each plane of its symmetry. Another class of spherical fullerenes like C_{140} and C_{260} lacks the mirror symmetry h ; hence, their maximum symmetry group is icosahedral (I) (Terrones et al. 2002).

The fullerenes C_{60} and C_{70} were identified in carbon flames, and their ratios depend on the temperature, pressure, carbon/oxygen ratio, and residence time in the flame (Howard et al. 1991). The molecular structure of C_{70} was deduced from electron diffraction using a simulated-annealing method (McKenzie et al. 1992). Scientists tried to understand the crystal structures of C_{60} and C_{70} using X-ray diffraction technique (David et al. 1991, 1992; Fischer et al. 1991; Valsakumar et al. 1993). At ambient temperature and pressure, C_{60} crystals have face-centered cubic (fcc) structure with a lattice constant of 14.17 Å (David et al. 1991), while the C_{70} crystals adopt to a hexagonal close-packed (hcp) structure with $a = 10.1$ Å and $c = 17.0$ Å (David et al. 1992). The average diameters of C_{60} and C_{70} fullerenes are about 7 and 9 Å, respectively. Since the discovery of C_{60} followed by C_{70} (Kroto et al. 1985; Taylor et al. 1990), different sizes of carbon cage fullerenes were revealed. In early 1990, the carbon cages of C_{76} , C_{84} , C_{90} , and C_{94} were characterized by mass spectrometry, ^{13}C NMR, electronic absorption (ultraviolet–visible), and vibrational (infrared) spectroscopy techniques (Diederich et al. 1991a). As compared with C_{60} and C_{70} , the isolation of higher fullerenes is really challenging, and their characterization is complicated by the presence of a varying number of isomers.

Fullerenes are generally represented by a formula C_n , where n is an even number and denotes the number of carbon atoms present in the cage. Theoretical calculations predicted that fullerenes larger than C_{76} should have at least two isomeric forms (Manolopoulos and Fowler 1991). For fullerenes C_{84} and C_{96} , 24 and 187 distinct isomers were predicted, respectively (Manolopoulos and Fowler 1991, 1992). Three isomers for C_{78} and two isomers for C_{84} were isolated and characterized by ^{13}C NMR spectroscopy (Kikuchi et al. 1992a). Some of the isomers of C_{78} and C_{82} proposed by experimental ^{13}C NMR spectroscopy are depicted in Fig. 7. Many of the unique properties of fullerenes originate from their unusual cage structures. Therefore, determining the ground-state geometries of the fullerenes was considered to be an important step in understanding their unusual properties.

Experimental works are very limited for higher fullerenes beyond C_{84} because such species are difficult to isolate in pure form in quantities suitable for comprehensive study. Synthesis of C_{60} , in isolable quantities, was achieved using flash vacuum pyrolysis (FVP) technique by Scott and coworkers by 12 steps in 2002, and no other fullerenes were formed as by-products (Scott et al. 2002). Larger fullerene C_{78} was synthesized using the same FVP technique used for C_{60} synthesis (Amsharov and Jensen 2008). This shows a promise for the synthesis of higher fullerenes. It is noteworthy to mention that there has been a lot of experimental and theoretical studies involving fragments of fullerenes, called “buckybowls” (Barth and Lawton 1966; Dinadayalane and Sastry 2001, 2002a, b; Dinadayalane et al. 2001, 2002, 2003, 2004; Mehta and Rao 1998; Mehta et al. 1997; Priyakumar and Sastry 2001a, b, c; Sakurai et al. 2003; Sastry and Priyakumar 2001; Sastry et al. 1993, 2000; Seiders et al. 1999; Sygula and Rabideau 1999; Wu and Siegel 2006). The smallest buckybowl “corannulene” was synthesized nearly 20 years prior to the discovery of fullerene C_{60} (Barth and Lawton 1966). Another small fragment of C_{60} called

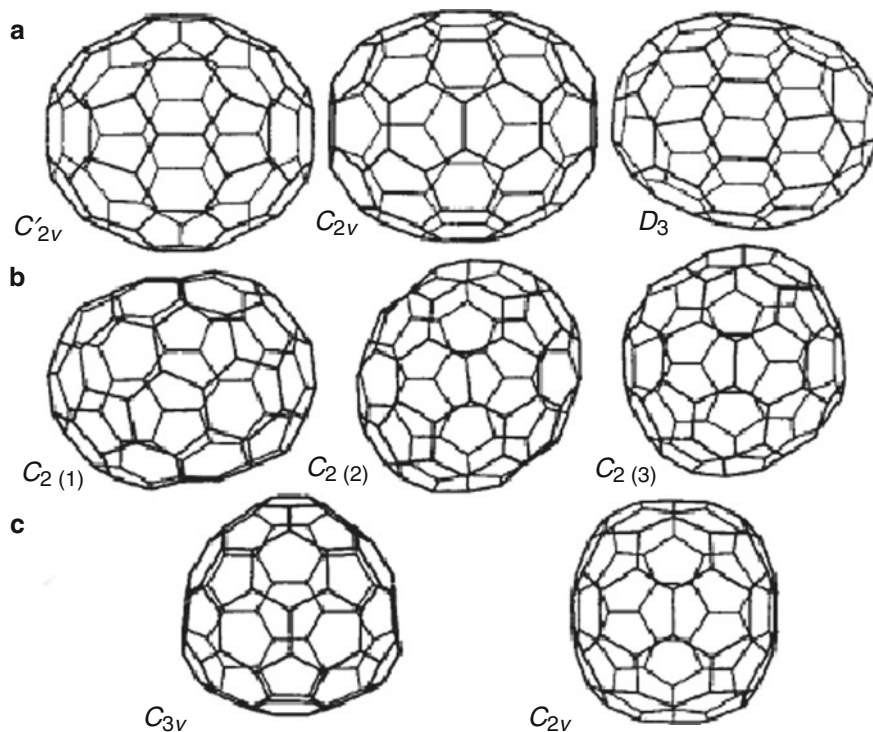


Fig. 7 The structures of fullerene isomers suggested by the ^{13}C NMR measurements. (a) Three isomers of C_{78} fullerene with C_{2v}' , C_{2v} , and D_3 point group. (b) Three structural candidates for C_{82} fullerene with C_2 symmetry. (c) Structures of C_{3v} - and C_{2v} - C_{82} isomers. (The picture was reprinted with permission from Macmillan Publishers Ltd.: Nature, reference Kikuchi et al. (1992a), Copyright 1992)

“sumanene” was successfully synthesized in 2003 after so many futile attempts by different groups (Sakurai et al. 2003).

Fullerene can be classified into (a) classical fullerene and (b) nonclassical fullerene. The former one is a closed carbon cage containing 12 pentagons and any number of hexagons, while a nonclassical fullerene can have heptagons, octagons, and an additional number of pentagons or squares. Growing classical fullerenes from nonclassical fullerenes, for example, from C_{50} to C_{60} by the dimer addition, was proposed (Hernández et al. 2001). However, there is no clear experimental evidence for fullerene formation through this route. The experimental evidence was reported for the formation of fullerenes by collisional heating of carbon rings in the gas phase (Helden et al. 1993). Various mechanisms have been proposed so far for the formation of fullerenes. They can be divided into two major models: the pentagon road (PR) model (Klein and Schmalz 1990; Maruyama and Yamaguchi 1998; Smalley 1992) and the fullerene road (FR) model (Heath 1991).

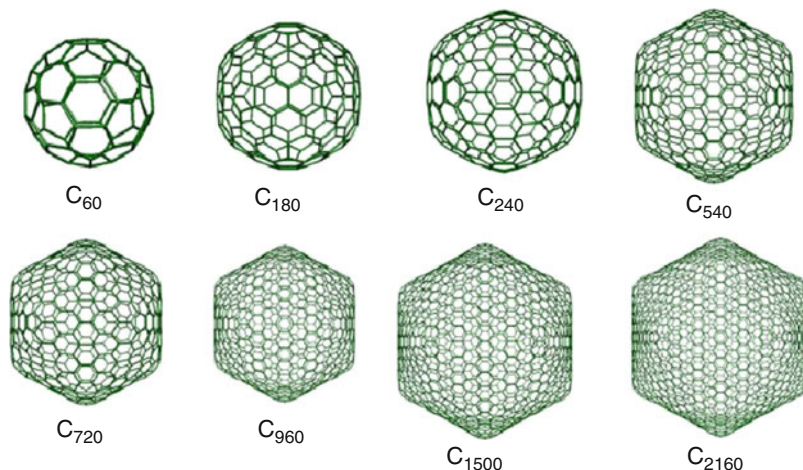


Fig. 8 The structures of giant fullerenes (Reprinted with permission from Zope et al. (2008). Copyright 2008 by the American Physical Society)

Since the discovery of C_{60} , scientists showed vast interest in larger fullerenes. Therefore, the family of fullerenes increased, and now it also includes C_{70} , C_{76} , C_{78} , C_{82} , C_{84} , C_{86} , C_{88} , C_{90} , C_{92} , C_{94} , and C_{96} (Diederich et al. 1991a, b; Kikuchi et al. 1992b; Kimura et al. 1995; Miyake et al. 2000; Mizorogi and Aihara 2003; Taylor et al. 1992, 1993). The fundamental understanding of the size dependence of the closed carbon cage structures is important for tailoring these systems for possible nanotechnology applications. Larger fullerenes that have an icosahedral symmetry can also be constructed. This procedure generates 12 pentagons positioned around vertices of an icosahedron, while all other carbon rings are hexagonal. In general, there are two kinds of fullerenes with I_h symmetry, one being $n = 60 k^2$ and the other $n = 20 k^2$, where n is the number of carbon atoms and k is any positive integer (Miyake et al. 2000). Figure 8 depicts some of the giant fullerene structures, where C_{180} and C_{720} belong to $180 k^2$ family of icosahedral fullerenes, but all other structures belong to $60 k^2$ family of icosahedral fullerenes. For more than a decade, these giant fullerenes have been fascinating molecules for theoreticians (Calaminici et al. 2009; Dulap and Zope 2006; Dunlap et al. 1991; Gueorguiev et al. 2004; Lopez-Urias et al. 2003; Tang and Huang 1995; Tang et al. 1993; Zope et al. 2008).

The closed carbon cages smaller than C_{60} consist of adjacent pentagons. Such smaller fullerenes are predicted to have unusual electronic, magnetic, and mechanical properties that arise mainly from the high curvature of their molecular surface (Kadish and Ruoff 2002). A dodecahedron consisting of 20 carbon atoms with only pentagon rings is topologically the smallest possible fullerene. The well-known isomers of C_{20} are cage, bowl, and ring as shown in Fig. 9. The bowl-shaped isomer is reminiscent of corannulene. The realization of the smallest carbon closed-cage C_{20} , which exclusively contains 12 pentagons, was doubtful

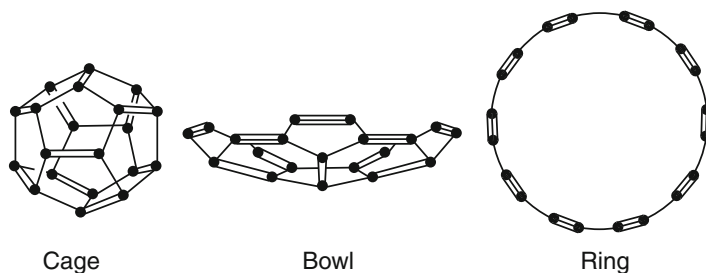


Fig. 9 Isomers of C_{20} : cage-, bowl-, and ring-shaped structures (Reprinted with permission from Macmillan Publishers Ltd.: Nature, reference Prinzbach et al. (2000), Copyright 2000)

until 2000. The C_{20} closed cage has extreme curvature and high reactivity, which led to doubts about its existence and stability (Wahl et al. 1993). Prinzbach et al. produced the smallest fullerene C_{20} from its perhydrogenated form in the gas phase and also obtained the bowl- and ring-shaped isomers for comparison purposes (Prinzbach et al. 2000). All these structures were characterized by photoelectron spectroscopy (PES) and their electron affinities vary significantly (Prinzbach et al. 2000). Theoretical calculations at different levels predicted dissimilar energetic ordering for these three isomers. However, all revealed very small relative energies of isomers. MP2 method predicted the fullerene to be the most stable followed by the bowl and then the ring, and this prediction is very similar to the calculations of density functional theory (DFT) using the local density approximation (LDA). Complete reversal of the stability ordering was obtained in the calculations with Becke–Lee–Yang–Parr (BLYP) functional. Some other DFT functionals predicted the bowl to be the most stable structure, closely followed by the fullerene isomer (Scuseria 1996). Hybrid density functional theory and time-dependent DFT formalism validated the synthesis of the smallest cage fullerene C_{20} by comparing the computed photoelectron spectra with the experimental results (Saito and Miyamoto 2001).

Closed-cage structure of C_{36} was detected by mass spectroscopy in very early days of fullerene science (Kroto 1987; Rohlffing et al. 1994). Zettl's group claimed the first preparation of C_{36} in the solid form (Piskoti et al. 1998). However, the existence of C_{36} has not been fully confirmed to date. C_{36} has 15 conventional fullerene isomers, out of which the D_{6h} and D_{2d} have a minimal number of pentagons (Fowler and Manolopoulos 1995). Therefore, these two are potential candidates for the most stable structure. In general, the number of isomers increases as the carbon cage size increases for these small fullerenes as shown in Fig. 10.

Fullerenes from C_{20} to C_{58} have been extensively studied by theoreticians (Fowler and Manolopoulos 1995; Scuseria 1996; Shao et al. 2007). They have been predicted to have narrow HOMO–LUMO gaps and high reactivity. The structures, aromaticity, reactivity, and other properties of fullerenes smaller than 60 carbon atoms were reviewed (Lu and Chen 2005). Readers may refer to the review by

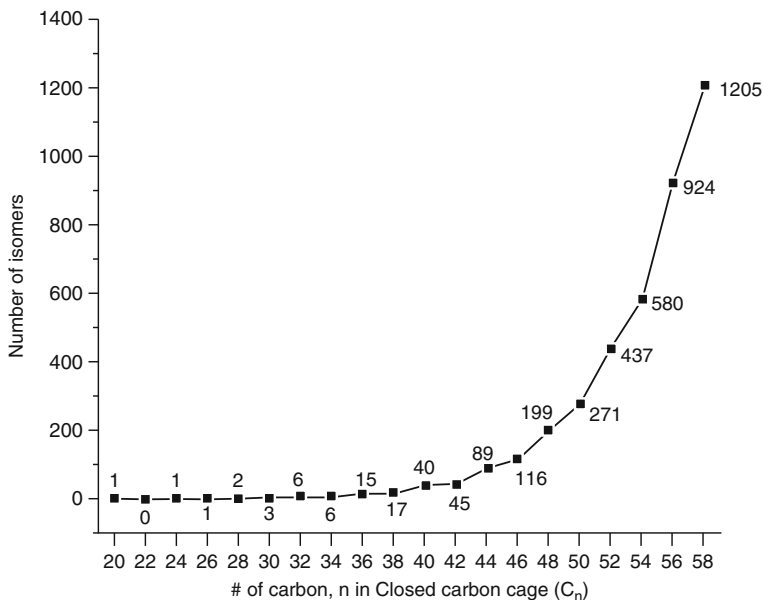


Fig. 10 Number of isomers for the closed carbon cages from C_{20} to C_{58} (Data was taken from reference Fowler and Manolopoulos 1995)

Lu and Chen and the references therein for detailed understanding and further knowledge if required (Lu and Chen 2005). Selected structures of smaller fullerenes and their isomers are depicted in Fig. 11. Schlegel diagram is commonly used by scientists to sketch the fullerenes in planar view, which is very helpful in identifying atoms and the C–C bonding networks (Fowler and Heine 2001; Thilgen and Diederich 2006; Troyanov and Tamm 2009). Figure 12 depicts the Schlegel diagram for fullerenes C_{20} , C_{36} , C_{60} , and C_{70} .

Natural Abundance of Fullerenes

Scientists discovered the presence of natural fullerenes on Earth. Interestingly, occurrence of fullerenes such as C_{60} and C_{70} was reported in shungite, a meta-anthracite coal from a deposit near Shunga, Russia (Buseck et al. 1992). The presence of C_{60} at very low concentrations in Cretaceous–Tertiary boundary sites in New Zealand was published (Heymann et al. 1994). Fullerenes (C_{60} and C_{70}) were found in a unit of shock-produced impact breccias (Onaping Formation) from the Sudbury impact structure in Ontario, Canada (Becker et al. 1994). The abundance of naturally occurring fullerenes was found in carbon materials, for example, coal, rocks, interstellar media, and even dinosaur eggs (Heymann et al. 2003).

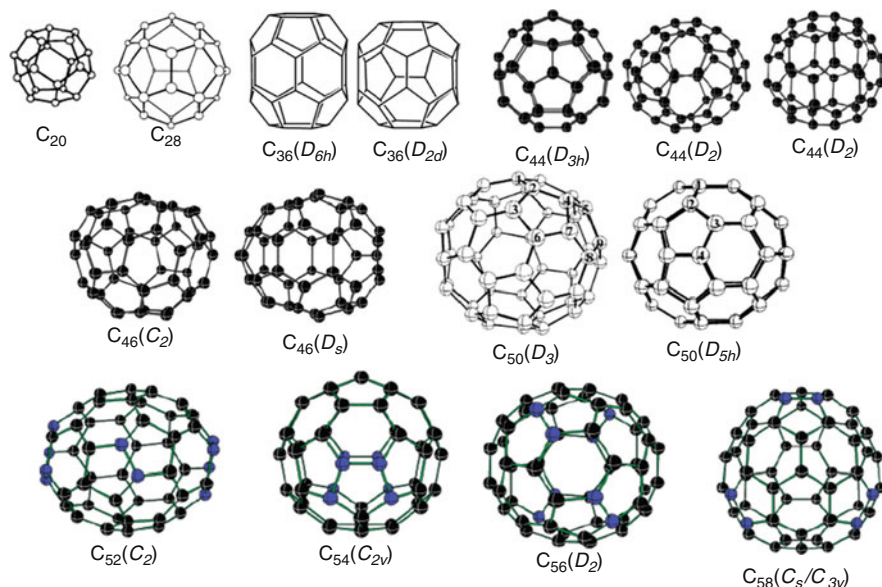


Fig. 11 Representative structures of smaller fullerenes and low-energy isomers are given for some of them. The pentagon–pentagon fusions are highlighted in blue color only for C_{52} , C_{54} , C_{56} , and C_{58} . The point groups are given in parentheses (Reprinted with permission from reference Lu and Chen (2005). Copyright 2005 American Chemical Society)

Fullerene Nano-capsules

In the area of fullerene science, one should not forget to mention an interesting property of holding the atoms or ions or molecules inside the fullerene cage (Thilgen and Diederich 2006). Fullerenes are potential nano-capsules. Experimental detection of the nano-capsules of fullerenes such as $\text{La}@C_{60}$, $\text{La}@C_{70}$, $\text{La}@C_{74}$, $\text{La}@C_{76}$, $\text{La}@C_{78}$, $\text{La}@C_{82}$, and $\text{Ce}_2@C_{80}$ was reported (Kessler et al. 1997; Kubozono et al. 1996; Moro et al. 1993; Saunders et al. 1993; Thilgen and Diederich 2006; Yamada et al. 2005, 2010). Fullerenes are known in the field of radioactive chemistry/physics. Radioactive nuclear materials can be stored by encapsulating inside the fullerenes. $\text{U}@C_{28}$, $\text{Gd}@C_{60}$, and $\text{Gd}@C_{82}$ are few examples of encapsulated radioactive materials (Guo et al. 1992; Kubozono et al. 1996). The stability of these metallic fullerenes could bring the new effective solution of the radioactive waste elimination. For the substance enclosed in the fullerene nano-capsule, carbon atoms act like a defense shield, and the fullerene containers are good for protecting their contents from water and acid. The structures, stabilities, and reactivities of encapsulated fullerenes (nano-capsules) and doped fullerenes have been the subject of theoretical interest (Guo et al. 1992; Lu et al. 2000; Park et al. 2005; Simeon et al. 2005; Wang et al. 2003; Wu and Hagelberg 2008; Zhao and Pitzer 1996). The closed-cage “fullerenes” or “heterofullerenes” can be placed inside the

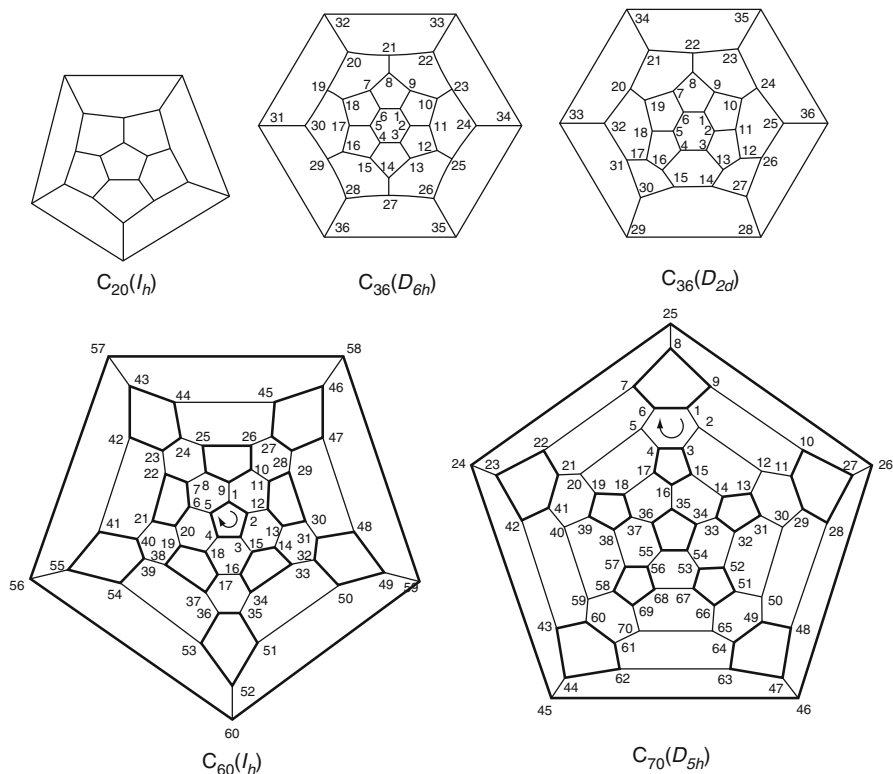


Fig. 12 Schlegel diagrams of C_{20} , C_{36} , C_{60} , and C_{70} fullerenes. The point groups are given in parentheses, systematic numbering recommended by IUPAC (Pictures were reprinted with permission from reference Thilgen and Diederich (2006). Copyright 2006 American Chemical Society. Reference Fowler and Heine (2001). Copyright 2001 Royal Society of Chemistry)

single-walled carbon nanotubes, for example, C_{60} @SWCNT (Hirahara et al. 2000; Okada 2007; Smith et al. 1999). Leszczynski and coworkers have explored the mechanism of the catalytic activity of fullerene derivatives using reliable computational methods (Sulman et al. 1999; Yanov et al. 2004). Fullerenes are certainly worthy of scientific study because of their unique shape and intriguing properties.

Isolated Pentagon Rule (IPR) in Fullerenes

A wide range of methods available for producing fullerenes concluded that C_{60} is the most abundant and is followed by C_{70} (Kadish and Ruoff 2002). The pristine C_{60} (I_h) contains two different C–C bonds: the one at the junction of two six-membered rings and the other one at the junction of a five- and a six-membered rings. These two bonds are usually labeled as a [6,6] and [5,6] C–C bonds, respectively. The pristine

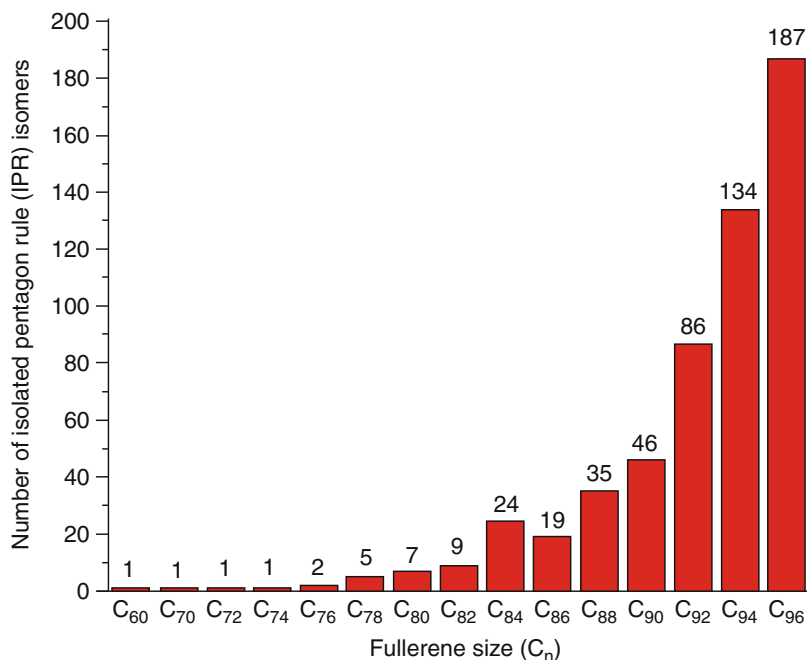


Fig. 13 Fullerene size versus number of isolated pentagon rule (IPR) isomers (The data was taken from reference Manolopoulos and Fowler (1992))

C_{70} (D_{5h}) has eight distinguishable C–C bonds. It has been known to chemists that energetically it is not favorable to have two pentagons sharing the same C–C bond. There are 1812 mathematical ways of forming a closed cage with 60 carbon atoms (isomers), but the buckminsterfullerene C_{60} (I_h) is the most special and stable because all of its pentagons are isolated by hexagons. This condition is called the “isolated pentagon rule” (IPR), which tends to make fullerenes more stable (Fowler and Manolopoulos 1995).

In fact, C_{60} is the smallest fullerene cage that obeys the isolated pentagon rule. Fullerenes C_{62} , C_{64} , C_{66} , and C_{68} do not satisfy the IPR. The next fullerene, which follows the IPR, is C_{70} (it has 8149 possible isomers). Also, C_{72} has an IPR structure. Most of the higher fullerenes have proven to follow IPR (Kroto 1987). Only one IPR-obeying isomer exists for C_{60} and for C_{70} (Fowler and Manolopoulos 1995), while the number of possible IPR isomers increases rapidly with increase in the size of the fullerenes as shown in Fig. 13. Fullerene C_{78} has five isomers that satisfy the IPR. Three isomers (two with C_{2v} and one with D_3 symmetry) out of these five were identified and characterized using ^{13}C NMR spectra (Kikuchi et al. 1992a). The fourth isomer (D_{3h} - C_{78}) has been recently separated and characterized in the form of $C_{78}(\text{CF}_3)_{12}$ (Shustova et al. 2006, 2007), and the last one has been synthesized using FVP technique (Amsharov and Jensen 2008). Theoretical study predicted that

C_{82} has nine IPR-satisfying isomers (Manolopoulos and Fowler 1991), out of which three isomers with C_2 symmetry were experimentally characterized using ^{13}C NMR spectroscopy, which gave 41 NMR lines with nearly equal intensity (Kikuchi et al. 1992a).

There are 24 geometric isomers satisfying IPR and 51,568 non-IPR isomers are possible for fullerene C_{84} (Fu et al. 2009). Earlier experimental ^{13}C NMR spectroscopy studies characterized two IPR isomers with D_2 and D_{2d} point groups (Kikuchi et al. 1992a; Taylor et al. 1993). Third isomer was also identified and reported (Achiba et al. 1995; Crassous et al. 1999). Pure D_2-C_{84} was synthesized (Dennis and Shinohara 1998). A theoretical study revealed that C_{84} cage is special in the family of fullerenes from C_{60} to C_{90} since the number of preferable isomers is larger than in the case of other fullerenes. The geometries of all of the 24 IPR isomers of fullerene C_{84} along with their point groups were reported (Okada and Saito 1996). Two IPR isomers of C_{86} out of possible 19 isomers were separated using multistage HPLC (high-performance liquid chromatography) (Miyake et al. 2000), and these two isomers were characterized to have C_2 and C_s point groups by ^{13}C NMR spectroscopy (Taylor et al. 1993). Burda et al. showed the experimental evidence for the photoisomerization of higher fullerenes. They confirmed the theoretical prediction that C_{86} has less number of IPR-satisfying isomers (19 isomers) than C_{84} (24 isomers) (Fig. 14; Burda et al. 2002). Experimental studies based on ^{13}C NMR spectroscopy revealed that fullerenes C_{88} and C_{92} possess 35 and 86 IPR isomers, respectively, and HPLC was used to separate the isomers (Achiba et al. 1996; Miyake et al. 2000; Tagmatarchis et al. 2002). Computational methods

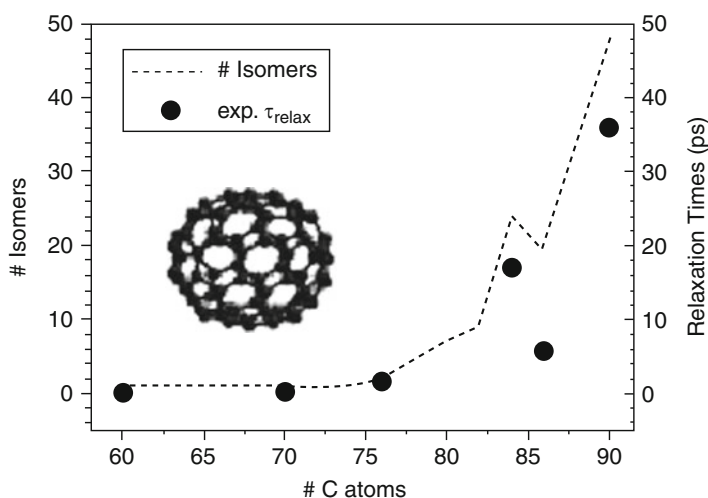


Fig. 14 Correlation between the number of isomers (dotted line as a visual guide) for each fullerene and the time constant for the formation of the lowest excited singlet state monitored at 550 nm (dots) (Reprinted with permission from reference Burda et al. (2002). Copyright 2002 American Chemical Society)

were used to calculate relative energies of the IPR isomers and the ^{13}C NMR spectra of fullerenes (Shao et al. 2006; Slanina et al. 2000a, b; Sun 2003). The computed ^{13}C NMR spectra were interpreted with the available experimental data (Beavers et al. 2006; Chaur et al. 2009a, b; Melin et al. 2007; Rojas et al. 2007; Scheina and Friedrich 2008; Shao et al. 2006; Slanina et al. 2000a, b; Sun 2003; Xie et al. 2004). Theoretical calculations predicted that C_{88} (**33**) is one of the most stable IPR isomers (Shao et al. 2006). The number provided in the parenthesis is the isomer number. In consistent to the theoretical prediction, Troyanov and Tamm reported the isolation and X-ray crystal structures of trifluoromethyl derivatives of C_{88} (**33**) and C_{92} (**82**) fullerene isomers complying with the isolated pentagon rule (Troyanov and Tamm 2009).

Rojas et al. showed the experimental evidence of the decreasing trend in the gas-phase enthalpy of formation and strain energy per carbon atom as the size of the cluster increases. Thus, the fullerenes become more stable as they become larger in size (Rojas et al. 2007). Interestingly, molecules encapsulated inside the carbon cages stabilize the fullerene isomers that violate IPR (Beavers et al. 2006; Fu et al. 2009; Thilgen and Diederich 2006). Several IPR and non-IPR endohedral fullerenes (single metal, di-metal, or tri-metal nitride encapsulated fullerenes) were isolated and characterized experimentally (Beavers et al. 2006; Chaur et al. 2009a, b; Fu et al. 2009; Melin et al. 2007; Thilgen and Diederich 2006), and their isolation motivated significant theoretical interest (Fu et al. 2009; Park et al. 2005; We and Hagelberg 2008). The structures and relative energies of the IPR isomers of buckybowls were examined using computational methods (Dinadayalane and Sastry 2003). Head-to-tail exclusion rule was proposed in explaining the stability of carbon cage structures that obey the IPR (Scheina and Friedrich 2008).

Fullerenes with less than 60 carbon atoms cannot have isolated pentagons and therefore they should be highly unstable and reactive. Xie et al. synthesized non-IPR $D_{5h}\text{-C}_{50}$ fullerene, which is a little sister of C_{60} , by introduction of chlorine atoms at the most reactive pentagon–pentagon vertex fusions. They confirmed the $D_{5h}\text{-C}_{50}$ structure by mass spectrometry, infrared, Raman, ultraviolet–visible, and fluorescence spectroscopic techniques (Xie et al. 2004). The report of novel small cage “Saturn-shaped” $\text{C}_{50}\text{Cl}_{10}$ structure encourages the possibility of obtaining other small non-IPR fullerenes and their derivatives. The investigations of the properties and applications of small fullerenes and their derivatives are now open.

Common Defects in Fullerenes

Stone and Wales examined rotation of C–C bonds in various fullerene structures using approximate Huckel calculations. The 90° rotation of C–C bond in fullerene is called Stone–Wales (SW) or “pyraclyene” rearrangement (Fig. 15; Stone and Wales 1986). Austin et al. reported that 94 % of all fullerene C_{60} isomers can rearrange to buckminsterfullerene by SW transformation (Austin et al. 1995). The C_{78} cage represents the smallest fullerene in which SW rearrangement can give

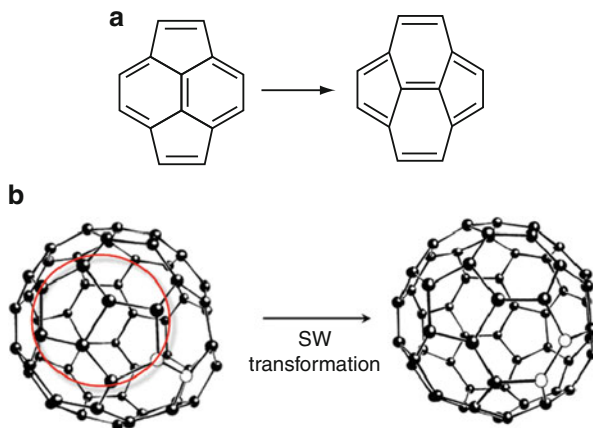


Fig. 15 (a) Stone–Wales or “pyracylene” transformation in fullerenes interchanges pentagons and hexagons; (b) Stone–Wales transformation of C_{2v} isomer of C_{60} with two adjacent pentagons gives the most stable I_h buckminsterfullerene; the C–C bond involved in 90° rotation is highlighted, and the two adjacent pentagons are marked in the fullerene structure in left-hand side

stable IPR isomers, $C_{78}:5$ (D_{3h}) \leftrightarrow $C_{78}:3$ (C_{2v}) \leftrightarrow $C_{78}:2$ (C_{2v}) \leftrightarrow $C_{78}:4$ (D_{3h}), where the numbers 5, 3, 2, and 4 indicate the isomer numbers (Austin et al. 1995). In case of higher fullerenes, the number of IPR isomers that can be transformed one into another by SW rearrangement considerably increases. For example, the SW transformation gives 9 and 21 stable IPR isomers for C_{82} and C_{84} , respectively (Fowler and Manolopoulos 1995). The SW transformation is usually thought to be the possible mechanism for achieving fullerene isomers (Austin et al. 1995; Fowler and Manolopoulos 1995; Stone and Wales 1986).

It was proposed that fullerenes can have seven-membered rings in addition to five- and six-membered rings (Taylor 1992). Troshin et al. isolated and characterized the C_{58} fullerene derivatives in which the cage structure contains the seven-membered ring. The structures were characterized using mass spectrometry, IR, and NMR spectroscopy (Troshin et al. 2005). Smalley and coworkers found that laser irradiation can fragment C_{60} into C_{58} , C_{56} , C_{54} , and other smaller cages with even number of carbon atoms via losing C_2 fragments (O’Brien et al. 1988). The formation of seven-membered rings was considered to play an important role in the fragmentation process of fullerenes (Murry et al. 1993). The laser desorption/ionization of products generated from the reactions of C_{60} with O_3 gives the odd-numbered clusters such as C_{59} , C_{57} , C_{55} , and C_{53} (Christian et al. 1992; Deng et al. 1993). Vacancy defects destroy the original topology of five- and six-membered rings in fullerenes (Christian et al. 1992; Deng et al. 1993; Hu and Ruckenstein 2003, 2004; Lee and Han 2004; Murry et al. 1993; O’Brien et al. 1988). They generate various sizes of rings such as four-, seven-, eight-, and nine-membered rings and also produce the new five- and six-membered rings depending on the number of carbon atom vacancies (Fig. 16; Christian et al. 1992; Deng et al.

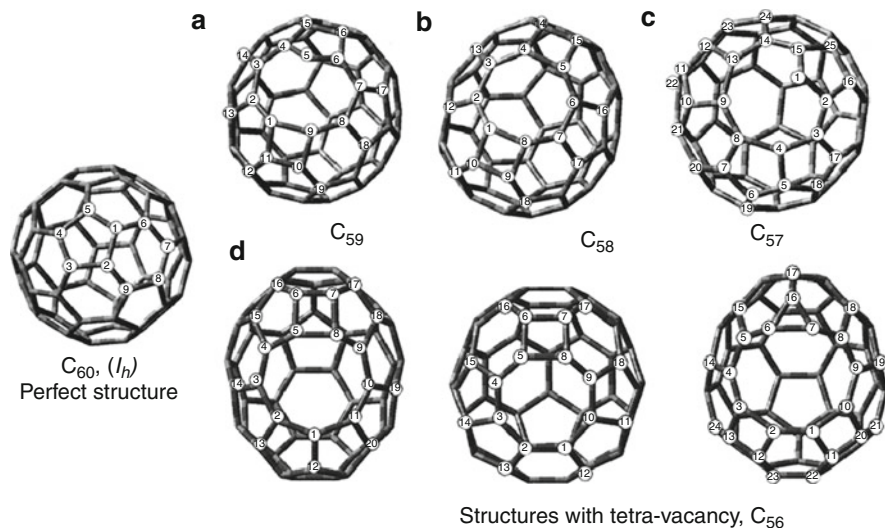


Fig. 16 Buckminsterfullerene (perfect structure); vacancy defect fullerenes generated from C_{60} : (a) mono-vacancy, (b) di-vacancy, (c) tri-vacancy, (d) three different structures of tetra-vacancy (Reprinted with permission from Hu and Ruckenstein (2004). Copyright 2004, American Institute of Physics)

1993; Hu and Ruckenstein 2003, 2004; Lee and Han 2004; Murry et al. 1993; O'Brien et al. 1988).

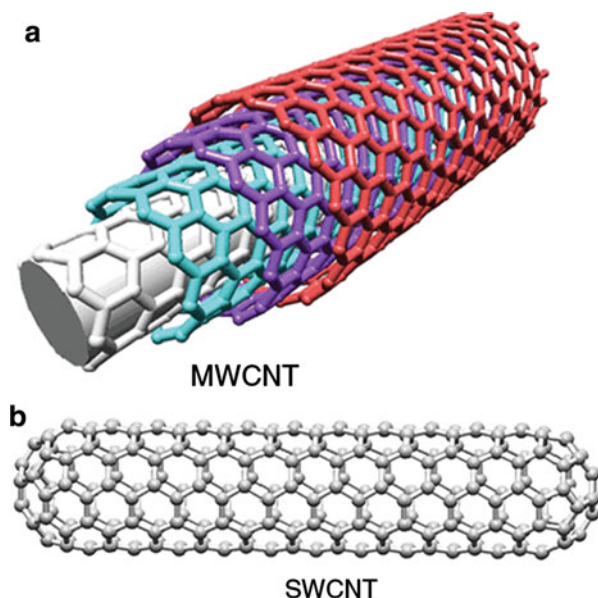
Carbon Nanotubes (CNTs)

Discovery and Classification of CNTs

Modern “nanotechnology revolution” was flourished by the discovery of fullerenes and has been escalating since the isolation of multi- and single-walled carbon nanotubes. The detection of carbon nanotubes by Iijima in 1991 is one of the landmarks in nanotechnology (Iijima 1991). In the interview to Nature Nanotechnology, Iijima told that the discovery of carbon nanotubes was unexpected but not entirely accidental because he had accumulated a lot of experience in looking at short-range order in carbon species such as amorphous carbon and very thin graphite sheets (Iijima 2007). The discovery of buckminsterfullerene by Kroto, Curl, Smalley, and coworkers motivated Iijima’s interest in finding out new carbon allotropes (Kroto et al. 1985).

There are two structural forms of carbon nanotubes: multi-walled carbon nanotubes (MWCNTs) and single-walled carbon nanotubes (SWCNTs). The former one was reported by Iijima (1991). The SWCNTs were reported independently by Iijima as well as Bethune groups (Bethune et al. 1993; Iijima and Ichihashi 1993). The existence of carbon nanotubes was reported as early as 1952 and also in 1976

Fig. 17 Representative structures of (a) multi-walled and (b) single-walled carbon nanotubes



(Oberlin et al. 1976; Radushkevich and Lukyanovich 1952). However, those reports did not reach the wide range of scientific community because they were published in unpopular journals, and at that time, no fabrication process was known that would lead to the synthesis of macroscopic amounts of carbon nanotubes (Oberlin et al. 1976; Radushkevich and Lukyanovich 1952). Monthioux and Kuznetsov documented the history of carbon nanotube since 1952 (Monthioux and Kuznetsov 2006). Synthesis of carbon nanotube using coal pyrolysis was reported (Moothi et al. 2015).

High-resolution electron microscopy (HREM) images of the CNTs showed the resemblance of a “Russian doll” structural model that is based on hollow concentric cylinders capped at both ends. The model structures of multi-walled and single-walled carbon nanotubes are shown in Fig. 17. A wide range of methods, such as arc evaporation of graphite, laser ablation, chemical vapor deposition (CVD), vapor phase decomposition or disproportionation of carbon-containing molecules, etc., have been reported for the synthesis of multi-walled and single-walled carbon nanotubes (Dresselhaus et al. 2001). It remains unclear whether SWCNTs and MWCNTs are formed via the same mechanism. It is also unclear whether various methods used to produce carbon nanotubes are mechanistically consistent (Dresselhaus et al. 2001). For the transformation pathway, fullerenes are known to be a suitable carbon source for MWCNT growth under certain conditions (Suchanek et al. 2001). An ideal MWCNT consists of cylindrical tubes in which the neighboring tubes are weakly bonded through van der Waals forces. The MWCNT is incommensurate when each of its walls has its own chirality independent of other walls.

SWCNT, which is a one-dimensional (1D) system, can be considered as the conceptual rolling of a section of two-dimensional (2D) graphene sheet into a seamless cylinder forming the nanotube. The structure of SWCNT is uniquely described by two integers (n, m), which refer to the number of \vec{a}_1 and \vec{a}_2 unit vectors of the 2D graphene lattice that are contained in the chiral vector, $C_h = n\vec{a}_1 + m\vec{a}_2$. The chiral vector determines whether the nanotube is a semiconductor, metal, or semimetal. From the (n, m) indices, one can calculate the nanotube diameter (d_t), the chirality or chiral angle (θ), the electronic energy bands, and the density of electronic states. The nanotube diameter (d_t) determines the number of carbon atoms in the circular cross section of the nanotube shell, one atom in thickness (Saito et al. 1998). The tube diameter and chiral angle can be written in terms of (n, m) as

$$\text{Tube diameter, } d_t = \left(\sqrt{3}/\pi\right) a_{cc} \left(\sqrt{m^2 + mn + n^2}\right).$$

$$\text{Chiral angle, } \theta = \tan^{-1} \left\{ \sqrt{3} \frac{m}{2n + m} \right\},$$

where a_{cc} is the nearest-neighbor carbon atom distance of 1.421 Å.

Among the large number of possible C_h vectors, there are two inequivalent high-symmetry directions. These are termed “zigzag” and “armchair” and are designated by ($n,0$) and (n,n), respectively. Either achiral (armchair and zigzag) or chiral SWCNTs can be constructed depending on the orientation of the six-membered rings with respect to the nanotube axis. Schematic representation of the structures of armchair, zigzag, and chiral SWCNTs is shown in Fig. 18. Theoretical studies in 1992 predicted that the electronic properties of “ideal” SWCNTs depend on the width and chirality of the tubes (Hamada et al. 1992; Mintmire et al. 1992; Saito et al. 1992). The electronic properties of an SWCNT vary in a periodic way between being metallic and semiconductor. SWCNTs are metals if $(n-m)/3$ represent an integer; otherwise, they are called semiconductors (Dresselhaus et al. 2002).

Several metallic (n, m) nanotubes have almost the same diameter d_t (from 1.31 to 1.43 nm), but have different chiral angles: $\theta = 0, 8.9, 14.7, 20.2, 24.8,$ and 30.0° for nanotubes (18, 0), (15, 3), (14, 5), (13, 7), (11, 8), and (10, 10), respectively (Dresselhaus et al. 2002). Few people realize that CNTs constitute a large family with a wide variety of sizes and properties, which are determined by their structure and composition, including chirality, number of walls, ordering of the wall, defects, surface functionalization, and other features.

Strano sorted out chiral SWCNTs into left-handed and right-handed tubes (Strano 2007). Significant progress has been made in the area of carbon nanotubes. Scientists are able to disperse, identify, sort, and now also isolate various types of carbon nanotubes (Arnold et al. 2006; Peng et al. 2007; Strano 2003, 2007). Specific methods were found to grow long SWCNTs and control the nanotube diameters (Lu et al. 2008; Zhang et al. 2008). Controlled synthesis of nanotubes

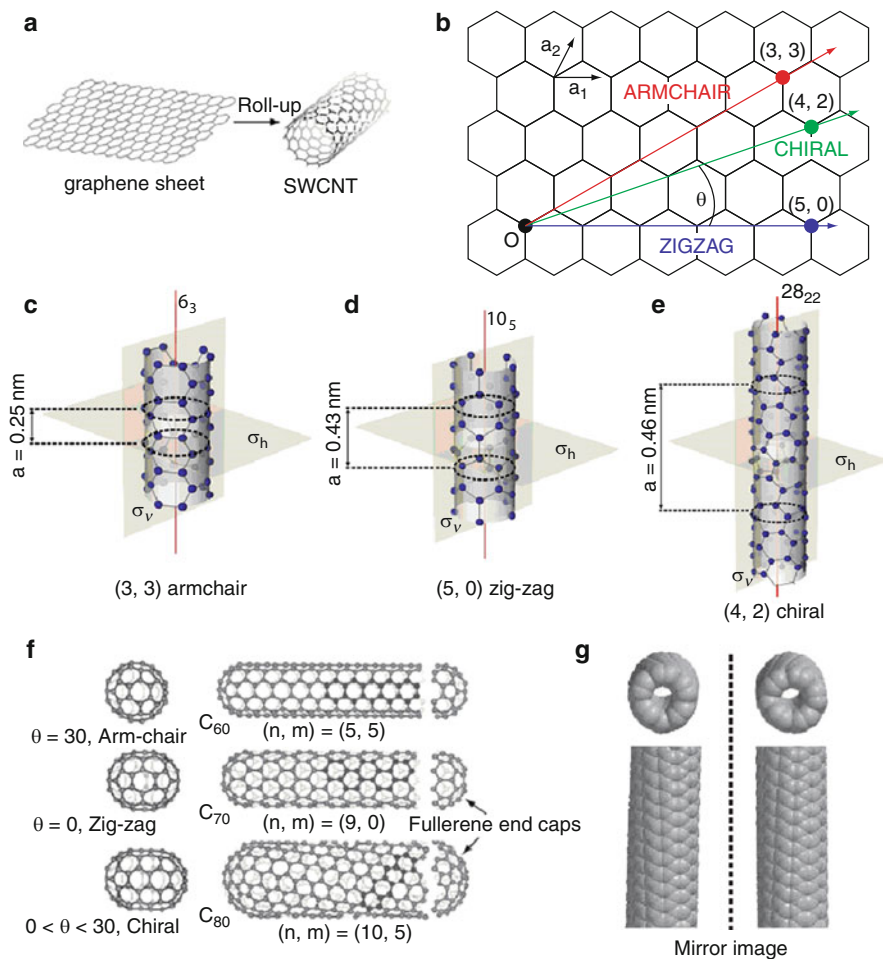


Fig. 18 (a) The roll-up of graphene sheet into SWCNT; (b) picture shows how to roll up graphene sheet to generate three different types of SWCNTs; (c) (3,3) armchair SWCNT; (d) (5,0) zigzag SWCNT; (e) (4,2) chiral SWCNT; (f) three types of SWCNTs (armchair, zigzag, and chiral) with fullerene end caps. These can be viewed as the growth of SWCNTs by adding several layers of hexagonal rings at the middle of different fullerenes; (g) mirror image of the chiral SWCNT; the structures (c), (d), and (e) are given exactly same types of SWCNTs that are mentioned in (b). In (c)–(e), σ_v and σ_h indicate the vertical plane of symmetry and horizontal plane of symmetry, respectively. Further, in these three structures, the red line is the axis of rotation; the distance of one unit cell for these three types of SWCNTs is provided; the number of carbon atoms (x) in each layer and the number of layers (y) required for one unit cell is given as x_y , for example, -6_3 means six carbon atoms in each layer and three layers required for one unit cell of (3,3) armchair SWCNT (Pictures (b) and (g) were reprinted with permission from Macmillan Publishers Ltd.: Nature Nanotech., reference Strano (2007), Copyright 2007)

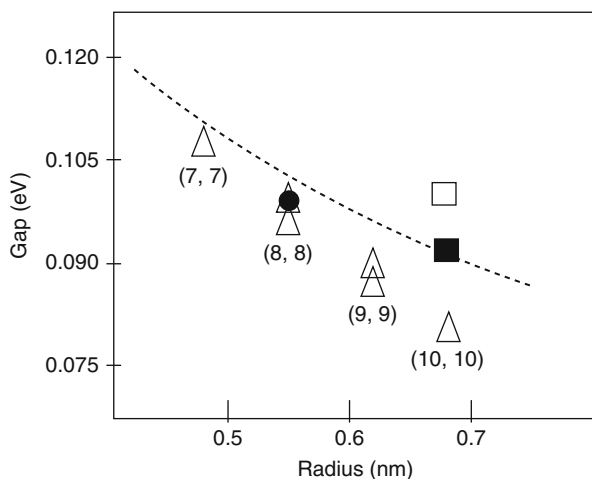


Fig. 19 Tube radius versus observed band gap. Each experimental data point (Δ) represents an average gap value measured on a distinct (n,n) tube. Theoretical results (*solid square*, *open square*, and *solid circle*) are also shown for comparison (Reprinted with permission from reference Ouyang et al. (2002). Copyright 2002 American Chemical Society)

opens up exciting opportunities in nanoscience and nanotechnology (Dai 2002). A range of methods was found for effective separation of metallic and semiconducting SWCNTs. Although some synthetic procedures have been known, they are not easy methods for synthesizing bulk quantities of metallic and semiconducting SWCNTs (Zhang et al. 2008). Scientists succeeded the preferential growth of SWCNTs with metallic conductivity (Rao et al. 2009a).

Raman and electronic spectroscopy techniques are useful in characterizing metallic and semiconducting SWCNTs. The radial breathing mode (RBM) in Raman spectra of SWCNTs is helpful in determining the diameter and chiral indices (n, m) of the nanotubes (Dresselhaus et al. 2002, 2005, 2007; Harutyunyan et al. 2009; Rao et al. 2009a). Experimental results pointed out decreasing band gap with increasing radius of the armchair SWCNTs (Fig. 19; Ouyang et al. 2002). Another breakthrough in carbon nanotube chemistry was accounted by Zhang and Zuo, who have determined a quantitative atomic structure of MWCNT containing five walls with diameter ranging from 17 to 46 Å and the C–C bond lengths of individual SWCNTs using electron diffraction technique (Zhang and Zuo 2009). Their results indicate that there are three different bond lengths in chiral walls and two different bond lengths in achiral walls (Zhang and Zuo 2009). Electron diffraction technique was used in determination of atomic structure of SWCNTs and the chiral indices (n,m) of CNTs (Jiang et al. 2007; Qin 2007). Rao et al. have revealed the efficient growth of SWCNTs of diameter 1–3 nm from diamond nanoparticles and fullerenes (Rao et al. 2009b). Recently, experimental reports are available to create three-dimensional graphene-CNT hollow fibers with radially aligned CNTs seamlessly sheathed by a cylindrical graphene layer through a

one-step chemical vapor deposition using an anodized aluminum wire template for efficient energy conversion and storage (Xue et al. 2015; Yu et al. 2014; Zhu et al. 2012).

SWCNTs have stimulated vast interest due to their unique structural, mechanical, electronic, thermal, and chemical properties and their potential applications in diversified areas. There has been enormous growth in patents related to carbon nanotubes, fuelled by predictions that the market for nanotubes will be \$9 billion by 2020. Between 1994 and 2006, it was estimated that 1865 nanotube-related patents were issued in the USA. Still, there is a cumulative backlog of more than 4500 patent applications relevant to CNT as reported in 2008 (MacKenzie et al. 2008).

Since there are reports of the natural abundance of fullerenes (Becker et al. 1994; Buseck et al. 1992; Heymann et al. 1994, 2003), the issue of the natural occurrence of carbon nanotubes has also attracted the attention of researchers. In 2004, TEM images that appear to be MWCNTs isolated from a Greenland ice core were reported (Esquivel and Murr 2004). The images of hollow carbon fibers from oil-well samples were reported (Velasco-Santos et al. 2003). However, we do not have any evidence for naturally occurring SWCNTs.

Various Defects in Carbon Nanotubes

Carbon nanotubes are not as perfect as they were thought to be earlier. Defects such as pentagons, heptagons, Stone–Wales defects, vacancies, adatoms, and dopants can occur in the nanotube during the growth or in processing and handling of the CNTs (Charlier 2002). Figure 20 depicts different types of defects in SWCNTs. Heptagon defects are found to play a crucial role in the topology of nanotube-based molecular junctions, for making X and Y type nanotube connections (Menon and Srivastava 1997). Long ago, theoretical studies proposed that pentagon–heptagon pair can be found in the intramolecular junctions of two SWCNT segments of different chiralities (Fig. 20d; Charlier et al. 1996; Chico et al. 1996). Experimental study revealed that ion irradiation-induced defects in the SWCNTs and the dangling bonds produced by irradiation are rapidly saturated (Chakraborty et al. 2007).

Low-energy electron and photon also induce damage in SWCNTs. The defect formation in SWCNTs is strongly dependent on the nanotube diameter, suggesting that the curvature-induced strain energy plays a crucial role in the damage (Suzuki and Kobayashi 2007). The defect formation and healing are reversible processes (Berthe et al. 2007; Suzuki and Kobayashi 2007). The defects in the SWCNTs affect their electronic, optical, and chemical properties. A competition between the defect formation and healing at room temperature or even below was reported. Raman spectroscopy, electrical measurements, and photoluminescence (PL) spectroscopy were used to examine the defect formation. However, the type of defects was not confirmed. Chemically stable topological defect, Stone–Wales defect, was ruled out because the activation energy for the defect healing was quite small (1 eV). Low-energy electron and photon can break C–C bonds in SWCNTs, as it was concluded

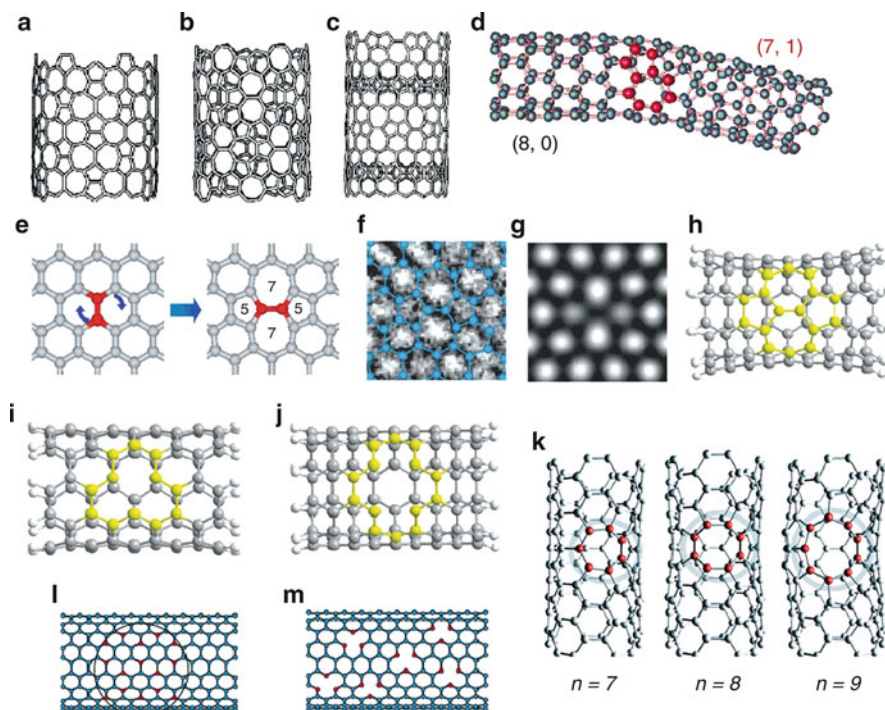


Fig. 20 (a) to (c) Nonchiral Haeckelite nanotubes of similar diameter; (b) nanotube segment containing only heptagons and pentagons paired symmetrically. (b) Nanotube segment exhibiting repetitive units of three agglomerated heptagons, surrounded by alternating pentagons and hexagons. (c) Nanotube segment containing pentalene and heptalene units bound together and surrounded by six-membered rings. (d) Atomic structure of an (8,0)-(7,1) intermolecular junction; the large red balls denote the atoms forming the pentagon-heptagon pair. (e) The SW transformation leading to the 5-7-7-5 defect, generated by rotation of a C-C bond in a hexagonal network. (f) HRTEM image obtained for the atomic arrangement of the SW defect. (b) Simulated HRTEM image for the model shown in (f). (h) (5,5) armchair SWCNT with a Stone-Wales defect. (i) Ideal single-vacancy defect in (5,5) armchair SWCNT. (j) Ideal double-vacancy defect in (5,5) armchair SWCNT. (k) Defect (5,5) SWCNT with seven ($n = 7$)-, eight ($n = 8$)-, and nine ($n = 9$)-membered rings. (l) SWCNT doped with boron [B atoms are bonded to three C atoms; B in red spheres and C in blue spheres]. (m) SWCNT doped with nitrogen [N atoms are bonded to two C atoms; N in red spheres and C in blue spheres] (Pictures (a)-(d), (l), and (m) were reprinted with permission from reference Charlier (2002). Copyright 2002 American Chemical Society. Pictures (e)-(g) were reprinted with permission from Macmillan Publishers Ltd.: Nature, reference Suenaga et al. (2007), copyright 2007. Pictures (h-j) were reprinted with permission from reference Yang et al. (2006a). Copyright 2006 American Chemical Society. Picture (k) was reprinted with permission from Nishidate and Hasegawa (2005). Copyright 2005 by the American Physical Society)

based on energetic criterion. Thus, the experimental study proposed that the defects may be a vacancy and an adatom (Suzuki and Kobayashi 2007).

The Stone-Wales defect is one of the important defects in carbon nanotubes. Stone and Wales showed that a dipole consisting of a pair of five- and seven-membered rings could be created by 90° rotation of a C-C bond in a hexagonal

network (Stone and Wales 1986). Such a dipole was later called Stone–Wales defect. SW transformation is thought to play an important role during the growth of carbon nanotubes. Miyamoto et al. reported an unambiguous identification of SW defect in carbon and boron nitride nanotubes using photoabsorption and vibrational spectroscopy (Miyamoto et al. 2004). Experimental vibrational frequency of 1962 cm^{-1} was reported to be a signature in identifying SW defect in carbon nanotube (Miyamoto et al. 2004). Identifying and characterizing topological defects in SWCNTs are highly challenging tasks. A powerful microscope with high resolution and high sensitivity is required for characterizing the topological defects in CNTs. Using HRTEM, the first direct image of the pentagon–heptagon pair defect (Stone–Wales defect) in the SWCNT was reported (Suenaga et al. 2007). Computational studies examined the structures and defect formation energies of the SWCNTs with defects containing different sizes of rings (seven-, eight-, and nine-membered rings) (Nishidate and Hasegawa 2005) and different types of defects (Amorim et al. 2007; Dinadayalane and Leszczynski 2007a, b; Ding 2005; Wang et al. 2006; Yang et al. 2006a, b).

Computational Approaches Used to Study Carbon Nanostructures: An Overview

Theory and computation play an important role in understanding structures and reactivity of carbon nanosystems such as graphene, fullerenes, and carbon nanotubes. Computational nanoscience often complements the experiments and is very useful for the design of novel carbon nanomaterials as well as predicting their properties. Theory is helpful in obtaining knowledge on the mechanism of reactions and fragmentations of carbon clusters. Thus, we provide an overview of the computational approaches employed to study various carbon nanostructures in this chapter. Carbon nanostructures are very large systems. Hence, performing very high-level quantum chemical calculations is not possible even when using modern supercomputers.

Many-body empirical potentials, empirical tight-binding molecular dynamics, and local density functional (LDF) methods were used to perform electronic structure calculations of carbon nanosystems including fullerenes and model CNTs in early of the last decade (Robertson et al. 1992; Zhang et al. 1993). In the mid-1990s, electronic structure calculations for large fullerenes with I_h point group were performed using Huckel approximation (Tang and Huang 1995). In the late 1990s, scientists performed geometry optimizations for large fullerenes using molecular mechanics (MM3), semiempirical methods (MNDO (Dewar and Thiel 1977), AM1 (Dewar et al. 1985) and PM3 (Stewart 1989)), and Semi-Ab Initio Model 1 (SAM1) (Dewar et al. 1993). The single-point energy calculations were affordable at that time using ab initio Hartree–Fock (HF) method in combination with small basis sets such as 3-21G and 4-31G (Slanina et al. 1997). The computing power has been tremendously increasing since 2000. Thus, currently theoreticians enjoy

investigating medium-sized molecules using reliable quantum chemical methods and exploring carbon nanoclusters beyond molecular mechanics and semiempirical methods.

The popular B3LYP functional, which is a combination of Becke's three-parameter (B3) (Becke 1993) hybrid functional incorporating exact exchange with Lee, Yang, and Parr's (LYP) (Lee et al. 1988) correlation functional, has been employed with small- and medium-sized basis sets like STO-3G, 3-21G, 4-31G, and 6-31G(d) for calculations on fullerenes and carbon nanotubes (Bettinger et al. 2003; Dinadayalane and Leszczynski 2007b; Feng et al. 2005; Matsuo et al. 2003; Yumura et al. 2005a, b; Zhou et al. 2004). Computational studies indicate that the B3LYP functional can yield reliable answers for the properties of carbon compounds and carbon nanostructures (Bettinger et al. 2003; Dinadayalane and Leszczynski 2007b; Feng et al. 2005; Matsuo et al. 2003; Yumura et al. 2005a, b; Zhou et al. 2004). PBE1PBE/6-311G(d) level has been used for calculating relative energies and ^{13}C NMR spectra of fullerene isomers (Shao et al. 2006, 2007). The PBE1PBE functional was concluded to be very reliable DFT functional since it yields the same relative energy ordering as the high-level coupled cluster calculations for the top three isomers of C_{20} (cage, bowl, and ring isomers) (An et al. 2005).

In comparison with *ab initio* MP2 or CCSD methods, DFT is less time-consuming and computationally feasible for large carbon nanosystems. For studying chemical reactivity in fullerenes and nanotubes, ONIOM approach is more cost-effective than treating the whole molecule with DFT. ONIOM is a hybrid methodology in which the molecule is partitioned into two or more fragments. The most important part (one fragment) of the molecule is treated with high-level method, and the other parts are treated with low-level methods (Maseras and Morokuma 1995; Morokuma et al. 2006; Osuna et al. 2009). The performance of ONIOM approach by taking different density functional theory levels was examined against the experimental results for the Diels–Alder reaction between cyclopentadiene and C_{60} (Osuna et al. 2009). Two-layer ONIOM approach ONIOM(B3LYP/6-31G(d):AM1), where B3LYP/6-31G(d) and AM1 are used for high and low layers, was utilized to study chemisorption of alkoxide ions with the perfect and Stone–Wales defective armchair (5,5) SWCNTs of cap-ended and H-terminated structures (Wanbayor and Ruangpornvisuti 2008).

Independent theoretical studies considered DFT methods in investigating the structures and properties of SWCNTs (Akdim et al. 2007; Amorim et al. 2007; Andzelm et al. 2006; Bettinger 2005; Dinadayalane and Leszczynski 2007a, b; Govind et al. 2008; Lu et al. 2005; Nishidate and Hasegawa 2005; Robertson et al. 1992; Wang et al. 2006; Yang et al. 2006a, b; Zhang et al. 1993). The B3LYP functional with double- ζ basis set was often employed to investigate the electronic structures of pristine and defective SWCNTs and also the influence of defects on functionalization of SWCNTs (Akdim et al. 2007; Andzelm et al. 2006; Dinadayalane and Leszczynski 2007b; Govind et al. 2008; Lu et al. 2005). Sometimes, more than one basis set was utilized for exploration of defective SWCNTs and the viability of metal adsorption in the defect tubes (Yang et al. 2006b). DFT with periodic boundary condition (PBC) as implemented in Gaussian

03 program package (Frisch et al. 2003) was used to examine the reactivity of Stone–Wales defect in (5,5) and (10,0) SWCNTs (Bettinger 2005). The unit cell should be carefully chosen for the calculations involving PBC in order to simulate the tubes of infinite length. Popular DFT methods fail to provide reliable answers for π – π interactions involving fullerenes and other carbon clusters (Cuesta et al. 2006; Kar et al. 2008; Shukla and Leszczynski 2009). Although calculations at the MP2 and CCSD(T) levels are required to obtain very reliable results for π – π stacking interactions (Dinadayalane et al. 2007a; Lee et al. 2007; Sinnokrot and Sherrill 2004), they are not possible for such large systems with current computational facilities. Recently developed meta-hybrid density functional (M06-2X) has been reported to be a promising functional to calculate the binding energies for π – π interactions involving large carbon nanostructures (Zhao and Truhlar 2007, 2008).

Using powerful supercomputers, performing static and dynamic calculations at high-level ab initio and DFT methodologies is affordable for graphene and carbon nanotubes. Very recently, density functional theory (PBE functional, Perdew et al. 1996) calculations with plane-wave basis sets and periodic boundary conditions (PBCs) were employed to understand small-molecule interactions with the defective graphene sheets (Jiang et al. 2009). Vienna ab initio simulation package (VASP) has been used in several studies to perform static and dynamic calculations (Kresse and Furthmuller 1996a, b). Theoretical calculations are helpful to understand the electronic structure of graphene sheets and SWCNTs and their viability as ion separation systems and gas sensors (Jiang et al. 2009; Li et al. 2009b; Nishidate and Hasegawa 2005). The Stone–Wales defect formation energy for graphene and CNTs was calculated using DFT, invoking the local density approximation to the exchange–correlation potential as implemented in VASP (Ertekin et al. 2009).

The mechanical properties of CNTs have been investigated by theoreticians for the last two decades (Avila and Lacerda 2008; Chandra et al. 2004; Dereli and Sungu 2007; Yakobson et al. 1996). An array of methods has been employed for computing the Young's modulus of MWCNTs and different types of SWCNTs (armchair, zigzag, chiral). A wide range of Young's modulus values has been reported in the literature (Avila and Lacerda 2008; Chandra et al. 2004; Dereli and Sungu 2007; Mielke et al. 2004; WenXing et al. 2004; Yakobson et al. 1996). Most of the molecular dynamics methods used so far are classical or tight binding (Avila and Lacerda 2008; Chandra et al. 2004; Dereli and Sungu 2007; WenXing et al. 2004; Yakobson et al. 1996). Quantum chemical calculations on mechanical properties of carbon nanotubes or graphene sheets are scarce since they are still highly time-consuming (Mielke et al. 2004). It is not of our interest to discuss mechanical properties in this chapter since there are many papers and some of classic reviews on this subject available (Avila and Lacerda 2008; Chandra et al. 2004; Dereli and Sungu 2007; Mielke et al. 2004; WenXing et al. 2004; Yakobson et al. 1996). DFT and DFT-D approaches were used to study single and multiple Na adsorption and diffusion on graphene (Malyi et al. 2015). By using DFT and time-dependent DFT methods, one can obtain IR, Raman, NMR, and UV spectra. Recent advances in computer hardware and ab initio electronic structure methods have brought a substantial improvement in the capabilities of quantum chemists to predict and study

the properties of carbon nanostructures. However, the application of state-of-the-art quantum chemical methods to study the structures and properties of large carbon nanoclusters (graphenes, fullerenes, and CNTs) is still a great challenge.

Structural, Electronic, and Chemical Properties of Graphene, Fullerenes, and SWCNTs

Graphene

An experimental investigation of mechanical properties of monolayer graphene reported a breaking strength of 40 N/m and the Young's modulus of 1.0 TPa (Lee et al. 2008). Graphene displays a thermal conductivity of $5000 \text{ W m}^{-1} \text{ K}^{-1}$ at room temperature (Balandin et al. 2008). Graphene chemistry is expected to play an important role in producing graphene-based materials. Computational study using the density functional theory with generalized gradient approximation revealed the cooperative effects of degenerate perturbation and uniaxial strain on band gap opening in graphene. Furthermore, the band gap width could be continuously tuned by controlling the strain (Jia et al. 2016). In this chapter, we outline the Stone–Wales defect in graphene, chemisorption process (covalent functionalization) on graphene, and the influence of defects on chemisorption; particular interest is given to hydrogen chemisorption. Chemical functionalization in graphene should produce new 2D systems with distinct electronic structures and different electrical, optical, and chemical properties. Chemical changes can probably be induced even locally. The first known example of hydrogenated graphene is graphane, which is a 2D hydrocarbon with one hydrogen atom attached to every site of the honeycomb lattice (Elias et al. 2009; Sofo et al. 2007).

Stone–Wales defect is expected to enhance the tendency of graphitic layers to roll up into other carbon nanostructures such as fullerenes and nanotubes. Therefore, in-depth understanding of Stone–Wales defect in graphene is required. It is known that pentagons and heptagons induce curvature in graphitic materials. In perfect graphene, the equilibrium C–C bond length is reported as 1.42 Å using PBE functional with the plane-wave code CPMD (Hutter et al.). Further details of the calculations can be obtained from the paper of Ma et al. (2009). The C–C bond shared by two heptagons of the SW defect in graphene is compressed to 1.32 Å using the same method. Density functional theory and quantum Monte Carlo simulations reveal that the structure of the SW defect in graphene is not simple. Ma et al. systematically studied the polycyclic hydrocarbon size dependence on the structural distortion caused by the Stone–Wales defect formation. They considered different systems ranging from the smallest analog of SW defect, azulene ($\text{C}_{16}\text{H}_{10}$), to 1D tape-like structure of $\text{C}_{50}\text{H}_{28}$ and, finally, to 2D planar cluster of $\text{C}_{228}\text{H}_{38}$. As known earlier, azulene is planar. The optimized bond length of the C–C bond at the center of azulene is 1.38 Å, which is longer than the corresponding C–C bond length of 1.32–1.33 Å observed for the SW defect in graphene (Ma et al. 2009).

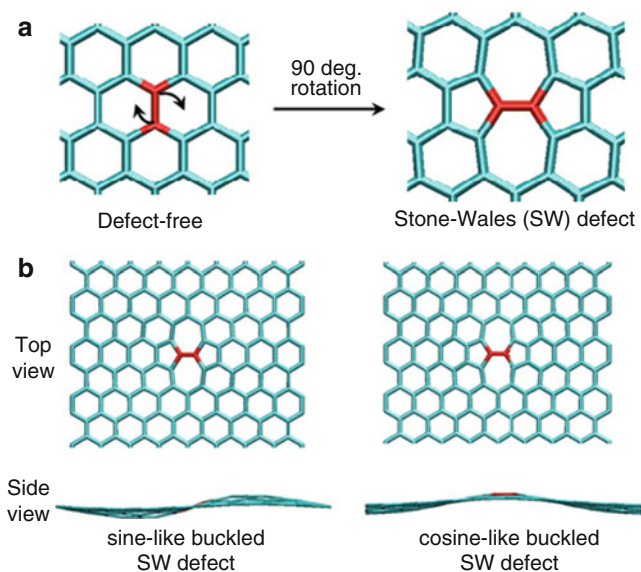


Fig. 21 (a) Stone–Wales transformation by 90° rotation of C–C bond in graphene sheet; (b) *top* and *side* views of sine-like and cosine-like buckled SW defect graphene sheets (The pictures were reprinted with permission from Ma et al. (2009). Copyright 2009 by the American Physical Society)

Large carbon clusters exhibit a tendency to buckle upon the creation of SW defects. Vibrational frequency calculations of the flat graphene sheet with the Stone–Wales defect reveal that the structure is not a local minimum, but instead has two imaginary frequencies. The true minimum is a sine-like structure in which the C–C bond at the defect core is 0.01 \AA longer than in the flat defect. Furthermore, many C–C bonds are slightly elongated in the buckled structure compared to the flat defect structure. The cosine-like SW defect structure was obtained as a transition state connecting to sine-like SW defect structure. The optimized structures of sine-like and cosine-like SW defect graphenes are depicted in Fig. 21. Vibrational frequencies also revealed that the maximum phonon frequencies corresponding to the stretch of the rotated C–C bond for the flat and buckled SW structure are 1880 and 1774 cm^{-1} , respectively. The corresponding frequency computed for perfect graphene is 1612 cm^{-1} (Ma et al. 2009). Theoretical study pointed out that for a graphene sheet of $\text{C}_{228}\text{H}_{38}$ containing a SW defect, the sine-like buckled structure becomes more stable (by $\sim 10 \text{ meV}$) than the flat SW defect (Ma et al. 2009).

Hydrogenation of Graphene With and Without Defects

Chemical modification of graphene has been less explored (Geim and Novoselov 2007). Attachment of atomic hydrogen to each site of the graphene lattice to create graphane is an elegant idea (Sofo et al. 2007). As a result, the hybridization of

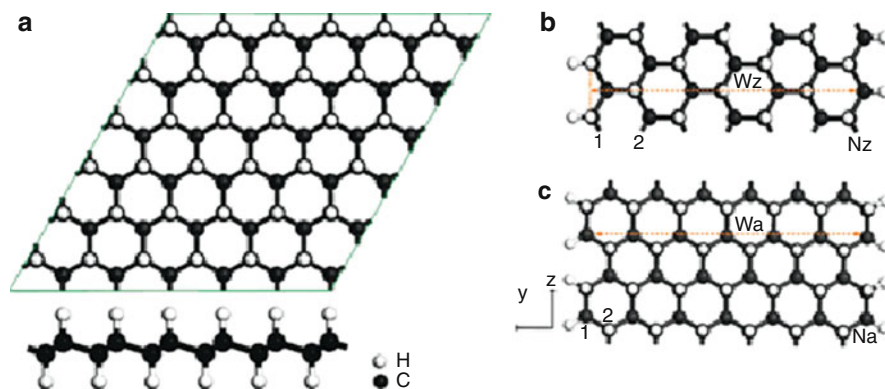


Fig. 22 (a) Top (*upper*) and side (*lower*) view of a 2D graphene layer. Geometric structures of the (b) 7 zigzag and (c) 13 armchair graphene nanoribbons. The ribbons are periodic along the z direction. The ribbon widths are denoted by W_z and W_a , respectively (Reprinted with permission from reference Li et al. (2009b). Copyright 2009 American Chemical Society)

carbon atoms changes from sp^2 to sp^3 , thus removing the conducting p-bands and opening energy gap (Boukhvalov et al. 2008; Sofo et al. 2007). In experiment, the fully hydrogenated graphene called “graphane” was produced by exposing graphene to hydrogen plasma discharge. Raman spectroscopy and transmission electron microscopy confirmed the reversible hydrogenation of single-layer graphene (Elias et al. 2009).

The structural and electronic properties of graphane were investigated using DFT PW91 functional with plane-wave basis set applying periodic boundary conditions as implemented in VASP (Li et al. 2009b). Computations revealed that hydrogenation of graphene nanoribbon is experimentally viable and the electronic properties of graphane are completely different from graphene nanoribbons. Figure 22 depicts the structures of graphane. Two types of graphene nanoribbons (zigzag and armchair edge) can be obtained by cutting the optimized graphene layer. The edge carbon atoms were all saturated with H atoms to avoid the effects of dangling bonds. The bond lengths of edge C–C and C–H bonds are almost as the inner C–C (1.52 Å) and C–H (1.11 Å) bonds. The calculated C–C bond length is similar to the bond length of 1.53 Å in diamond (sp^3 carbon atoms) and is longer than 1.42 Å characteristic of sp^2 carbon in graphene. Both spin-unpolarized and spin-polarized computations yielded same energy for ground-state graphane nanoribbons (Li et al. 2009b).

Figure 23a shows that computed band gap decreases monotonically with increasing ribbon width for both zigzag and armchair nanoribbons. Graphane nanoribbons are semiconductors. The formation energy increases with increasing ribbon width (Fig. 23b) irrespective of the type, indicating that narrow ribbons are more likely to form than the wider ribbons (Li et al. 2009b). Sofo et al. investigated the structures, formation energies, and vibrational frequencies of graphane using DFT with plane-wave basis set (Sofa et al. 2007). They found two favorable conformations of

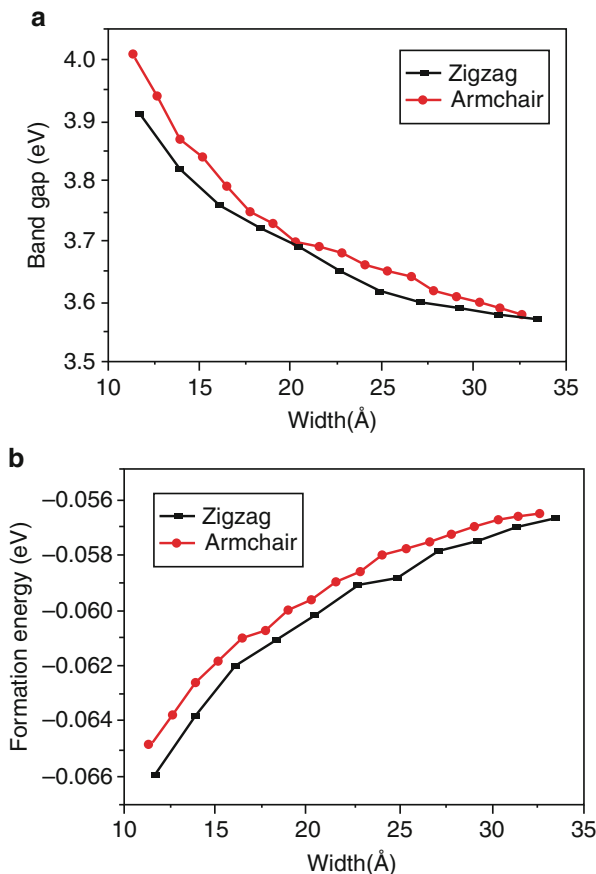


Fig. 23 Variation of the band gap (**a**) and the formation energy (**b**) of zigzag ($6 \leq N_z \leq 16$) and armchair ($10 \leq N_c \leq 27$) graphane nanoribbons as a function of ribbon width. N is the number of zigzag chains for a zigzag ribbon and the number of dimer lines along the ribbon direction for an armchair ribbon (Reprinted with permission from reference Li et al. (2009b). Copyright 2009 American Chemical Society)

graphane: chair-like conformer with the hydrogen atoms alternating on both sides of the plane and the boat-like conformer with the hydrogen atoms alternating in pairs. Chair conformer has one type of C–C bond (1.52 Å), while boat conformer possesses two different types of C–C bonds (bond lengths of 1.52 Å and 1.56 Å). The boat conformer is less stable than the chair conformer due to the repulsion of the two hydrogen atoms bonded to first-neighbor carbon atoms on the same side of the sheet. This repulsion results in slightly longer C–C bonds in boat conformer. Calculated C–H bond stretching frequencies are 3026 cm^{-1} and 2919 cm^{-1} for the boat and chair conformers, respectively. These C–H stretching modes are IR active, and they should be useful in characterizing these two types of conformers of graphane (Sofa et al. 2007). Molecular dynamics (MD) simulations using adaptive

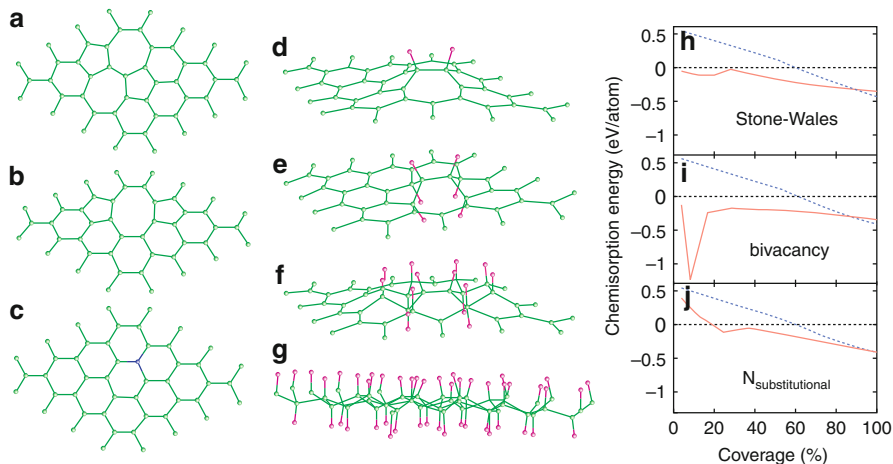


Fig. 24 Optimized geometric structures for graphene supercell containing (a) the Stone–Wales defect, (b) a bivacancy, and (c) a nitrogen substitution impurity. Optimized structures for the Stone–Wales (SW) defect functionalized by (d) 2, (e) 6, and (f) 14 hydrogen atoms and (g) completely covered by hydrogen. *Green circles* represent carbon atoms, *violet circles* represent hydrogen atoms, and *blue circle* represents nitrogen atom. Hydrogen atom chemisorption energy per atom as a function of coverage for a graphene sheet containing (h) a Stone–Wales (SW) defect, (i) a bivacancy, and (j) a nitrogen substitution impurity. The *blue dashed* line represents the results for the ideal infinite graphene sheet (Reprinted with permission from reference Boukhvalov and Katsnelson (2008). Copyright 2008 American Chemical Society)

intermolecular reactive empirical bond order (AIREBO) force field in LAMMPS package revealed the wrinkling characteristics in hydrogenated graphene annulus under circular shearing at the inner edge. Such hydrogenation-induced changes in topological and mechanical characteristics of graphene will be useful to develop novel graphene-based devices (Li et al. 2015).

Using density functional calculations, Boukhvalov and Katsnelson have studied hydrogenation of graphene sheets with defects such as Stone–Wales (SW), bivacancies, nitrogen substitution impurities, and zigzag edges. They performed calculations for chemisorptions of hydrogen atoms on the defects in the graphene from low to high coverage. The optimized geometries of the graphene supercells with various types of defects as well as their hydrogenated structures are depicted in Fig. 24, which also displays the computed chemisorption energy as the function of coverage for the graphene containing different defects. The chemisorption energy of a single hydrogen atom to the defect-free graphene was given as 1.5 eV, while those of 0.30 eV for SW defects, 0.93 eV for bivacancies, and 0.36 eV for substitution impurities of nitrogen in graphene were reported. This indicates the significant influence of defects on single hydrogen atom chemisorption energy in graphene. The calculated chemisorption energy for different nonequivalent carbon atoms of the graphene containing SW defect reveals that the chemisorption energy for the entire area surrounding the SW defect is lower compared to the perfect graphene.

Further, the defects also decrease the chemisorption energy of two hydrogen atoms at adjacent positions compared to the defect-free graphene. It was reported that, for the complete coverage, the binding energy is smaller for the hydrogen chemisorption of graphene with defects than the perfect graphene. Thus, completely hydrogenated graphene is less stable with defects than without them (Boukhvalov and Katsnelson 2008).

Graphenes with various kinds of defects may have different types of properties and applications. Therefore, obtaining knowledge on graphenes with defects is important. DFT calculations showed only physisorption of water molecule with perfect graphene, while the vacancy defect greatly assists the dissociative chemisorption of water molecules in the graphene (Cabrera-Sanfeliix and Darling 2007; Kostov et al. 2005). There can be many possible reaction pathways for the dissociation of water molecule over defective sites in the graphene (Kostov et al. 2005). Computational studies provide evidence that defects such as Stone–Wales and vacancy strongly influence the chemisorption of functional groups in the graphene (Boukhvalov and Katsnelson 2008; Boukhvalov et al. 2008; Cabrera-Sanfeliix and Darling 2007; Kostov et al. 2005).

Fullerenes

Computational Studies of Fullerene Isomers

Computational methods were employed to systematically search and study the low-lying isomeric structures of fullerenes, and such thorough investigations have been useful to predict the best candidates for the lowest-energy structures of higher fullerenes because of the growing experimental interest (Shao et al. 2006, 2007; Slanina et al. 2000a, b; Sun 2003; Sun and Kertesz 2002; Zhao et al. 2004a, b). Fullerene C_{86} has 19 possible isomers obeying IPR, and all of these isomers were studied using B3LYP functional with different basis sets (Sun and Kertesz 2002). Among 19 isomers, the isomer **17** with C_2 symmetry is the most stable followed by isomer **16** with C_s symmetry, and these two isomers are shown in Fig. 25. It should be noted that these two isomers were experimentally observed. At the B3LYP/6-31G level, isomer **16** was predicted to be about 6 kcal/mol less stable than isomer **17**, albeit the former one has slightly larger HOMO–LUMO gap than the latter. The variation of relative stability at different theoretical levels for all 19 IPR-satisfying isomers of C_{86} is depicted in Fig. 26. The relative stabilities were calculated with respect to the lowest-energy isomer (**17**). The HF/3-21G level and the semiempirical AM1 Hamiltonian overestimate the relative energies compared to the density functional theory (DFT) levels (Sun and Kertesz 2002).

Experimental study identified two isomers of fullerene C_{86} and characterized them using ^{13}C NMR spectroscopy (Miyake et al. 2000). Based on the experimental NMR spectra, C_2 and C_s point groups were assigned for the two isomers, but there are more than single C_2 and C_s isomers. Theoretical calculations play a crucial role in identifying the correct structure by comparing theoretical and experimental ^{13}C NMR spectra. The ^{13}C NMR chemical shifts were calculated for all of the 19

Fig. 25 Two experimentally observed IPR isomers of fullerene C_{86} . Their point groups are given in parentheses (Reprinted with permission from reference Sun and Kertesz (2002). Copyright 2002 Elsevier)

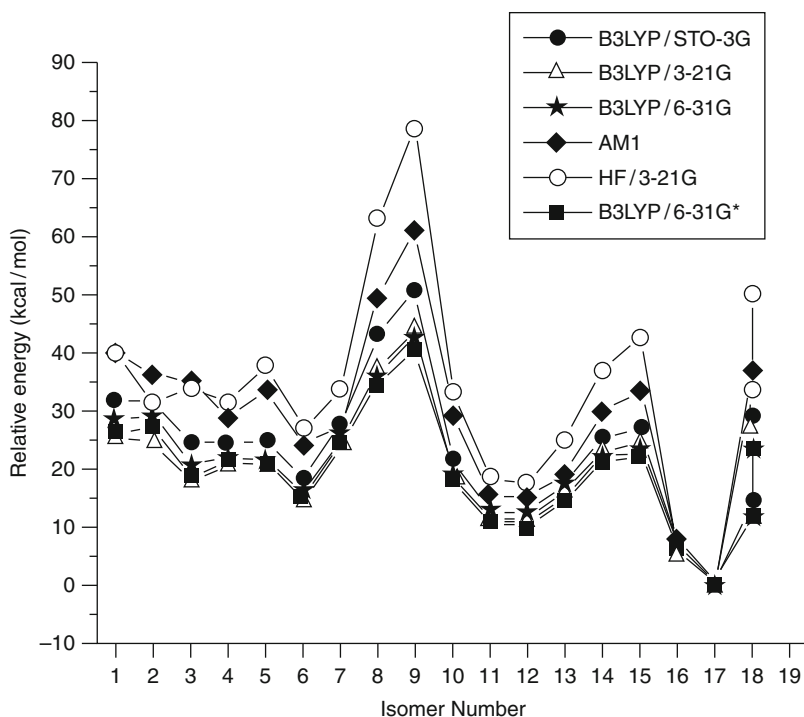
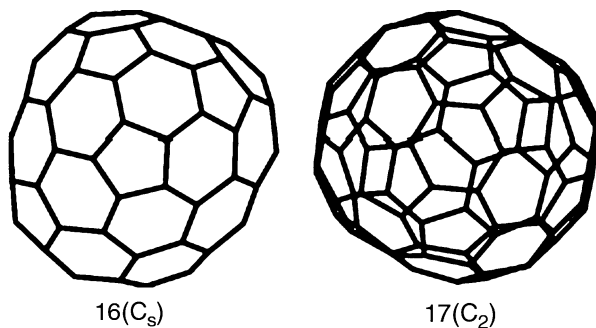


Fig. 26 The relative energy of IPR-satisfying isomers of C_{84} at various levels of theory. Isomers **1**, **5**, **7**, **11**, **12**, and **13** have C_1 symmetry. Isomers **2**, **3**, **4**, **6**, **14**, and **17** possess C_2 point group. Isomers **9** and **10** have C_{2v} point group. Isomers **8**, **15**, and **16** have C_s point group. Isomers **18** and **19** possess C_3 and D_3 point groups, respectively (The data was taken from reference Okada and Saito (1996))

IPR isomers of C_{86} , except isomer **8**. Theoretical ^{13}C NMR spectra complement the experimental spectra as evidenced from Fig. 27. Computational study revealed that isomer **17** has high thermodynamic and kinetic stability among the six IPR isomers of C_{84} possessing C_2 point group. The computed NMR spectrum of isomer **17** supports the results of experimental spectrum. Among the C_s isomers, the second

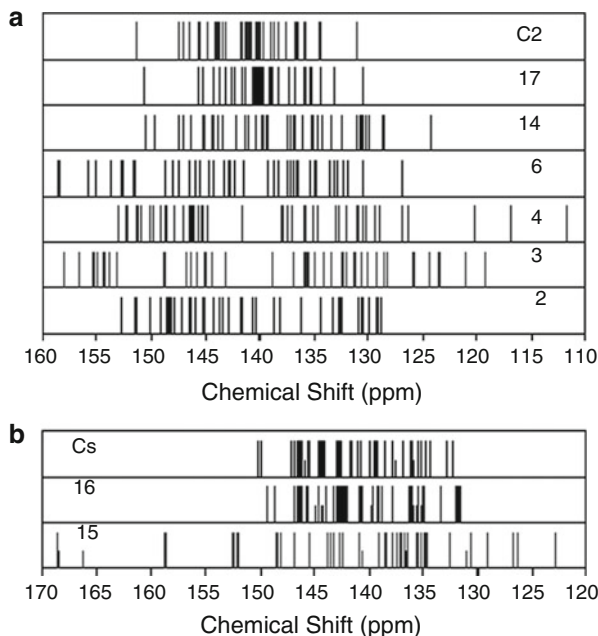


Fig. 27 Experimental and theoretical ^{13}C NMR spectra of (a) C_2 isomers of fullerene C_{86} . (b) C_s isomers of fullerene C_{86} . Theoretical spectra are labeled by isomer number and experimental spectrum labeled by symmetry (Reprinted with permission from reference Sun and Kertesz (2002). Copyright 2002 Elsevier)

most stable isomer **16** has large HOMO–LUMO gap (Sun and Kertesz 2002). Isomers **6**, **10**, **11**, **12**, **13**, and **18** were predicted to have relative energies less than 20 kcal/mol and moderate HOMO–LUMO gap, thus indicating the possibility of experimental realization (Sun and Kertesz 2002).

Okada and Saito proposed the number of extractable fullerenes among the IPR-satisfying isomers of fullerenes from C_{60} to C_{90} . They found that C_{84} is unique since the number of preferable isomers is more than for other fullerenes and this was attributed to the abundant production of C_{84} after C_{60} and C_{70} . All 24 IPR-satisfying isomers of C_{84} were studied computationally (Okada and Saito 1996). A complete set of 187 isomers that obey IPR of fullerene C_{96} was systematically investigated using various theoretical methods including molecular mechanics (MM3), semiempirical (AM1, MNDO, and PM3), and quantum mechanical (HF/4-31G and B3LYP/6-31G) methods. All of the theoretical levels unequivocally predicted that isomer **183** with D_2 point group is the lowest-energy one. The relative energies for some of the isomers were reported to be quite method sensitive and varied dramatically with different methods. The computational study highlighted the importance of the entropy effect in examining the relative stability of IPR-obeying isomers of fullerene C_{96} (Zhao et al. 2004a). A large set of 450 IPR

isomers of C_{100} has been explored using the abovementioned semiempirical and molecular mechanics (MM3) methods. Systematic theoretical calculations predicted the isomer with D_2 point group (isomer **449**) as the lowest energy by all of the methods employed (Zhao et al. 2004a).

Shao et al. searched the lowest-energy isomer of the fullerenes C_{38} to C_{80} and C_{112} to C_{120} . For the first set (C_{38} to C_{80}), all IPR and all non-IPR isomers were considered, and only IPR-satisfying isomers were considered for the second set of fullerenes (C_{112} to C_{120}). Thus, a large set of molecules was taken for optimizations at the semiempirical density functional-based tight-binding (DFTB) method and the single-point energy calculations at the DFT (Shao et al. 2007). It is known that the fullerene with large HOMO–LUMO gap and high symmetry is not necessarily the lowest-energy structure. The decreasing trend of HOMO–LUMO gap was reported with increasing the fullerene size (Shao et al. 2006, 2007). An unexpected manner of pentagonal adjacency was observed in the low-lying isomers in the series of fullerenes C_{38} to C_{80} (Shao et al. 2007).

In a comprehensive computational study, Shao et al. identified 20 isomers as the best candidates for the lowest-energy structures. Among the 20 isomers, 10 isomers with relative energies less than 1 kcal/mol are depicted in Fig. 28 (Shao et al. 2007). Thus, these 10 isomers can be observed experimentally. The ^{13}C NMR chemical shifts for these 10 isomers were calculated. Theoretical study predicted that C_{116} :**6061** can be more easily isolated and characterized in the laboratory than other higher fullerenes from C_{112} to C_{120} (Shao et al. 2007). In a different study, Shao et al. proposed the seven best candidates of the lowest-energy isomers for the fullerenes C_{98} to C_{110} based on the systematic study using DFTB and

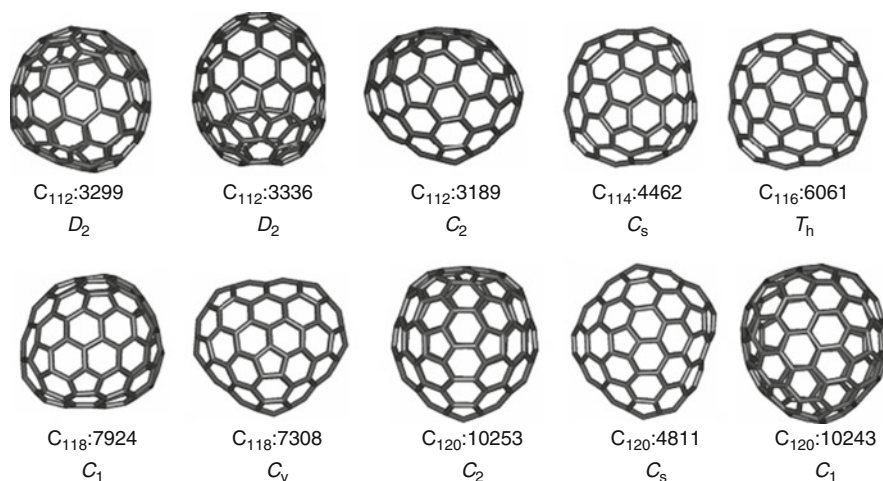


Fig. 28 Best candidates for the lowest-energy structure of higher fullerenes (C_{112} to C_{120}). The isomer number and point group are given (Reprinted with permission from reference Shao et al. (2007). Copyright 2007 American Chemical Society)

DFT methods. They pointed out that C_{102} (C_1 : **603**) and C_{108} (D_2 : **1771**) isomers can be easily synthesized (Shao et al. 2006). The concepts of cage connectivity and frontier π -orbitals play important roles to understand the relative stability of charged fullerene isomers without performing extensive quantum chemical calculations. This theoretical study correctly predicted the structures observed experimentally and explained why the isolated pentagon rule is often violated for fullerene anions, but the opposite is found for fullerene cations (Wang et al. 2015b).

Fullerenes C_{50} and $C_{50}Cl_{10}$ were computationally studied using B3LYP/6-31G(d) level due to the experimental report of the latter compound (Lu et al. 2004). The computational study thoroughly explored the structures, relative energies, HOMO–LUMO energies, and HOMO–LUMO gap for low-lying isomers of C_{50} and its anions. The computed IR, Raman, ^{13}C NMR, and UV–Vis spectra of the $C_{50}Cl_{10}$ with D_{5h} symmetry showed very good agreement with the reported experimental data. The pentagon–pentagon fusions were found to be the active sites of addition reactions in both D_3 and D_{5h} symmetric isomers of fullerene C_{50} . It was observed that HOMO and LUMO coefficients of C_{50} (D_{5h}) are distributed around the equatorial pentagon–pentagon fusion sites. This was given as a reason for the binding of Cl atoms around the equatorial pentagon–pentagon fusion sites of C_{50} yielded $C_{50}Cl_{10}$ (Lu et al. 2004).

Giant Fullerenes

Giant fullerenes have been the subject of theoretical interests (Calaminici et al. 2009; Dulap and Zope 2006; Dunlap et al. 1991; Gueorguiev et al. 2004; Lopez-Urias et al. 2003; Zope et al. 2008). The structures and stabilities of the giant fullerenes C_{180} , C_{240} , C_{320} , and C_{540} were investigated using high-level density functional theory calculations (Calaminici et al. 2009). The results of the uncorrected binding energy (in eV) per carbon atom for the giant fullerenes obtained using the VWN functional are depicted in Fig. 29. The inclusion of the basis set superposition error (BSSE) decreases the calculated binding energies but does not alter the trend. The increasing trend of binding energy indicates that the large fullerenes become more and more stable with increasing size. Fullerene C_{540} has a similar binding energy to diamond, giving the hope that such giant fullerenes could be prepared. However, the binding energy per carbon atom of the fullerene, C_{540} , is considerably lower than that of graphene (Calaminici et al. 2009). Gueorguiev et al. performed the calculations for giant fullerenes using semiclassical approximation LR-LCAO (linear response model in the framework of the linear combination of atomic orbitals). They reported the decreasing trend of HOMO–LUMO gaps (except C_{20}) and the considerably large increase of the static polarizability as increasing the size of the fullerene cage (Fig. 30; Gueorguiev et al. 2004). The static dipole polarizability per atom in C_{2160} is three times larger than that in C_{60} (Dunlap et al. 1991).

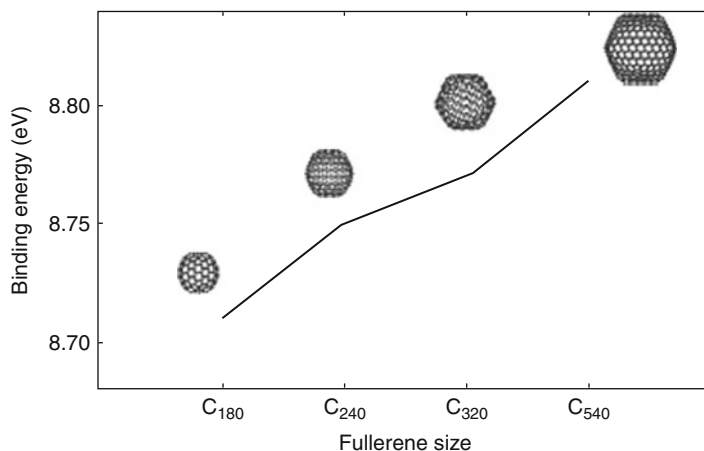


Fig. 29 Binding energy (in eV) for C₁₈₀, C₂₄₀, C₃₂₀, and C₅₄₀ fullerenes. The calculations have been performed with the VWN functional in combination with DZVP basis sets (Reprinted with permission from reference Calaminici et al. (2009). Copyright 2009 American Chemical Society)

Local Strain in Curved Polycyclic Systems: POAV and Pyramidalization Angle

Fullerenes experience large strain energy because of their spherical shape. The curvature-induced pyramidalization of the carbon atoms of fullerenes weakens the π -conjugation. The curved π -conjugation in carbon networks of fullerenes has not only π -character but also substantial s-character. The π -orbital axis vector (POAV) analysis developed by Haddon is useful in measuring the local curvature of the nonplanar conjugated organic molecules, fullerenes, and SWCNTs (Haddon 1993; Haddon and Scott 1986). In general, the sp^2 -hybridized carbon atom prefers to be in the planar arrangement, but it is pyramidalized in fullerenes. The local strain of carbon framework in fullerenes and SWCNTs is reflected in the pyramidalization angle θ_P at the carbon atoms (Niyogi et al. 2002). The pyramidalization angle (θ_P) equals to the difference between the π -orbital axis vector (POAV) and the normal right angle 90° : thus, $\theta_P = (\theta_{\sigma\pi} - 90^\circ)$, where the $\theta_{\sigma\pi}$ is the angle between the π -orbital of the conjugated atom and the σ -orbital of the surrounding atoms. As shown in Fig. 31, the pyramidalization angle is 0° and 19.47° for a planar sp^2 -hybridized carbon and a tetrahedral sp^3 -hybridized carbon, respectively. All carbon atoms in the icosahedral C₆₀ have the same $\theta_{\sigma\pi}$ of 11.6° .

Pyramidalization angle of a carbon atom in fullerenes and SWCNTs is helpful in predicting the chemical reactivity (Akdin et al. 2007; Bettinger 2005; Dinadayalane and Leszczynski 2007a, b; Lu and Chen 2005; Lu et al. 2005). The larger pyramidalization angle of carbon atom indicates the higher reactivity toward addition reactions in the curved systems of fullerenes and SWCNTs. Curvature-induced pyramidalization and the π -orbital misalignment cause local strain in SWCNTs

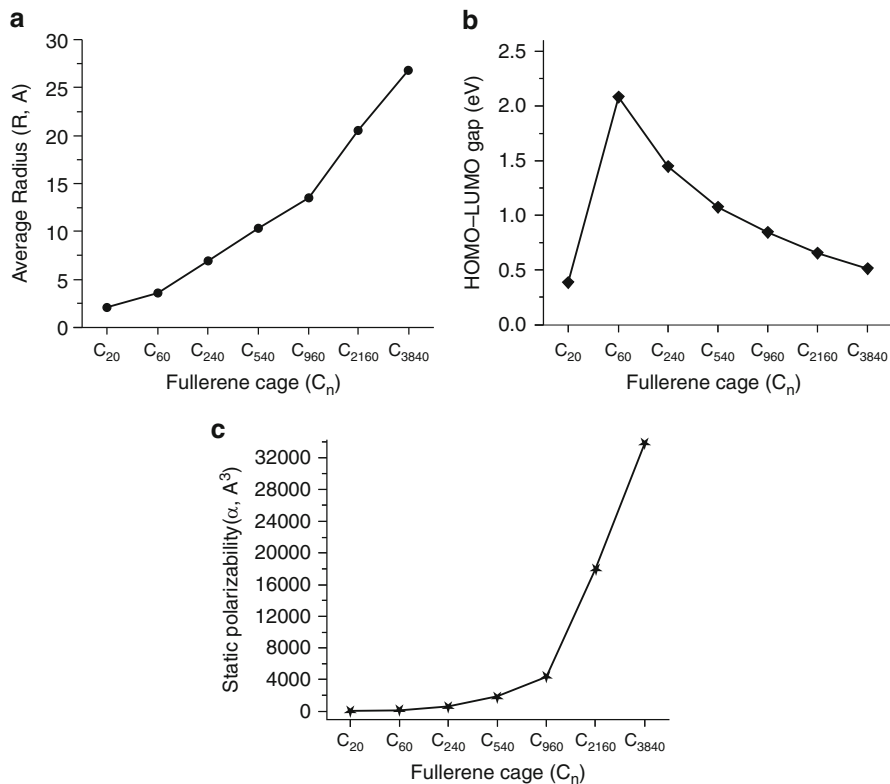


Fig. 30 The variation of (a) radius (R , Å), (b) HOMO–LUMO gap (eV), and (c) static polarizability as increasing the size of the fullerene cage (C_{20} , C_{60} , C_{240} , C_{540} , C_{960} , C_{2160} , and C_{3840}) (The data for the plots was taken from reference Gueorguiev et al. 2004)

(Fig. 32). Hence, carbon atoms of SWCNTs are more reactive than that of a perfect graphene sheet (Niyogi et al. 2002; Park et al. 2003). Cyranski et al. studied the structures and energetics of the 12 lowest-energy isomers of neutral, closed-shell IPR fullerenes C_{60} – C_{96} using B3LYP/6-31G(d) level. They obtained the decreasing values of pyramidalization angles, while no regular trend was obtained for HOMO–LUMO gaps with increasing size of fullerenes (Fig. 32; Cyranski et al. 2004).

Decachloro-derivative of C_{50} fullerene has been synthesized and experimental characterization confirmed the existence of C_{50} cage. Two C_{20} caps and five C_2 units around the equator are present in the C_{50} core of $C_{50}Cl_{10}$ (Xie et al. 2004). The calculated pyramidalization angle for the carbon atoms of the C_{20} caps of fullerene C_{50} ranges from 10.7 to 12.88°, which are comparable to that of C_{60} (11.68°). However, a large pyramidalization angle (15.58°) is obtained for the equatorial C atoms (pentagon–pentagon fusion). Such large value was attributed to high reactivity of those carbon atoms in addition reactions to form exohedral adducts (Chen 2004; Lu et al. 2004). Such structural features were reasoned for the instability of a bare C_{50} cage and the stability of $C_{50}Cl_{10}$ (Chen 2004).

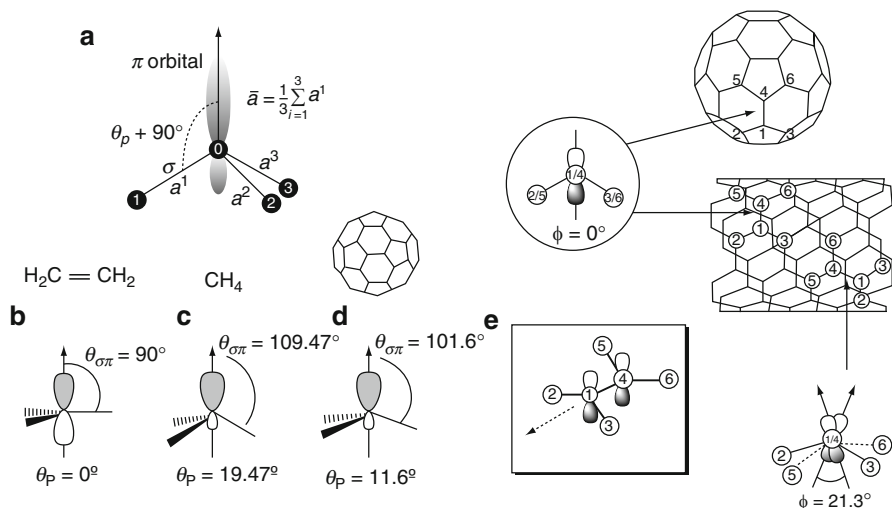


Fig. 31 (a) Pyramidalization angle (θ_p) is defined by the angle between the π -orbital and σ -bond minus 90° so that $\theta_p = 0^\circ$ for a graphene sheet and $\theta_p = 19.47^\circ$ for sp^3 -hybridized carbon. For practical reasons, we take the average of three θ_p values: (b) θ_p for a perfect planar sp^2 -hybridized carbon atom (e.g., in C_2H_4), (c) θ_p for a perfect tetrahedral sp^3 -hybridized carbon atom (e.g., CH_4), (d) θ_p for a nonplanar sp^2 -hybridized carbon atom (e.g., C atom in C_{60} or SWCNT). (e) The π -orbital misalignment angle (ϕ) along the C1–C4 bond in the (5,5) SWCNT and the fullerene C_{60} (Pictures were reprinted with permission from references Lu and Chen (2005), Niyogi et al. (2002), and Park et al. (2003). Copyright 2002, 2003 and 2005 American Chemical Society)

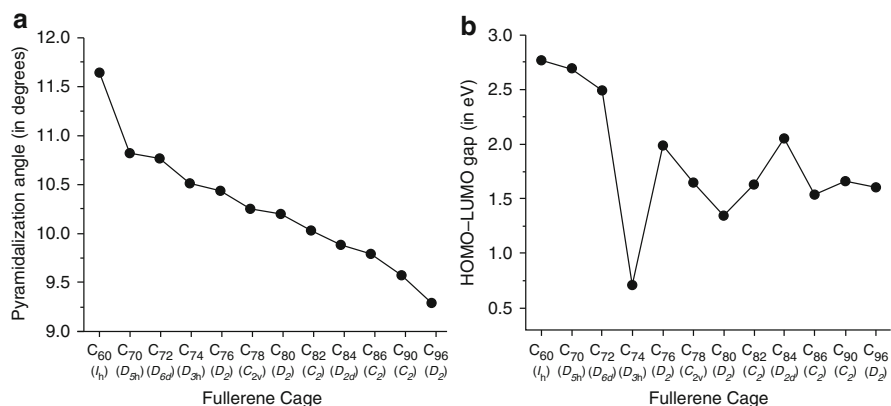


Fig. 32 (a) Variation of pyramidalization angle for the carbon atom of the most stable IPR fullerene as increasing the size of fullerene. (b) Variation of HOMO–LUMO gap for the most stable IPR fullerene as increasing the size of fullerene. The point groups are given in the parentheses. The values for these points were taken from reference Cyranski et al. (2004)

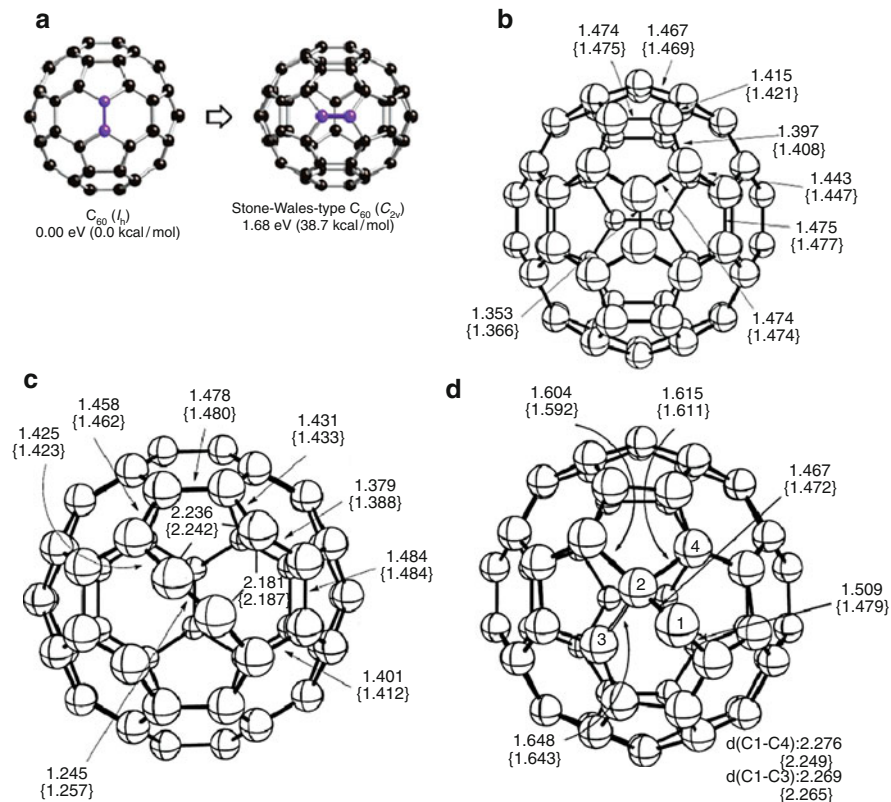


Fig. 33 (a) Buckminsterfullerene to $C_{60}-C_{2v}$ with Stone–Wales defect generated by the 90° rotation of the C–C bond in *blue color* of $C_{60} (I_h)$. (b) Optimized structure of the C_{2v} symmetry isomer. (c) Structure of the C_2 symmetry transition state for the concerted Stone–Wales transformation pathway. (d) Structure of the asymmetric transition state between carbene intermediate and $C_{60}-C_{2v}$ isomer. Bond lengths were obtained at the B3LYP/6-31G* and PBE/6-31G* (in curly brackets) levels of theory and are given in Å (Pictures were reprinted with permission from reference Bettinger et al. (2003) and Yumura et al. (2007). Copyright 2003 and 2007 American Chemical Society)

Stone–Wales Defect in C_{60}

Fullerene isomers are likely to interconvert through Stone–Wales transformation (Stone and Wales 1986; Troyanov and Tamm 2009). Very recently, experimental study has reported that the chlorine-functionalized D_2-C_{76} IPR isomer transformed to non-IPR isomer, and this transformation was proposed to include seven single Stone–Wales rearrangements (Ioffe et al. 2009). Computational chemists strived to understand the energy barriers for the Stone–Wales transformation and the possible mechanisms involved in this rearrangement, particularly considering the C_{60} fullerene (Bettinger et al. 2003; Eggen et al. 1996; Yumura et al. 2007).

Stone–Wales transformation is a thermally forbidden rearrangement by following the orbital symmetry considerations of Woodward and Hoffmann (1969). The icosahedral C_{60} fullerene (buckminsterfullerene) gives an isomer of C_{60} with C_{2v} point group that violates the isolated pentagon rule. Two different pathways, namely, concerted and stepwise pathways, and two different (symmetric and asymmetric) transition states were identified theoretically for the Stone–Wales transformation in C_{60} fullerene. The C_{60} - C_{2v} isomer, which is a Stone–Wales-type defect structure with two adjacent pentagons, was reported to be less stable by 33.9–38.7 kcal/mol (1.47–1.68 eV) than the buckminsterfullerene using various density functional theory levels (Yumura et al. 2007).

Bettinger et al. listed the C–C bond lengths of C_{60} (I_h) and the activation barrier for the Stone–Wales defect transformation through different transition states at various levels of theory. Computed geometries of buckminsterfullerene at different levels showed shorter [6,6] C–C bond length than the [5,6] C–C bond length, in consistent with experimental results (Bettinger et al. 2003). Figure 33 depicts the concerted C_2 symmetric transition state and asymmetric transition state involved in SW transformation of C_{60} - I_h to C_{60} - C_{2v} . The intrinsic reaction coordinate calculations by Bettinger et al. support the concerted pathway rather than stepwise pathway for the SW transformation in the C_{60} fullerene. Based on the computed activation energies, both concerted and stepwise pathways are highly competitive. The rigorous computational study of SW transformation in buckminsterfullerene revealed that the empirical schemes such as Tersoff–Brenner potentials and density functional-based tight binding (DFTB) underestimate the barrier heights, and semiempirical AM1 appears to be promising for such investigations (Bettinger et al. 2003).

Computational Studies on Vacancy Defects in Fullerene C_{60}

Vacancy defects in fullerene C_{60} were studied using quantum chemical methods (Hu and Ruckenstein 2003, 2004; Lee and Han 2004). They were generated by removal of 1–4 carbon atoms in C_{60} as shown in Figs. 34 and 35. Different modes are possible to remove carbon atoms from C_{60} to generate vacancy defects; hence, different sizes of rings (four-, seven-, eight-, and nine-membered rings) were produced by removing carbon atoms in C_{60} . Removing one, two, three, and four adjacent carbon atoms from the C_{60} cluster generates two, three, three, and six different isomers for the C_{59} , C_{58} , C_{57} , and C_{56} clusters, respectively (Hu and Ruckenstein 2004; Lee and Han 2004). The odd-numbered clusters have unsaturated carbon, which favors being located in a six-membered ring rather than a five-membered ring. Two-atom vacancies give structure with seven- and eight-membered rings, whereas one-atom vacancy gives the structure with nine-membered ring. Four-atom vacancies provide the most stable structure with only five- and six-membered rings. Thus, increasing the number of vacancies need not increase the size of the hole (Hu and Ruckenstein 2003, 2004).

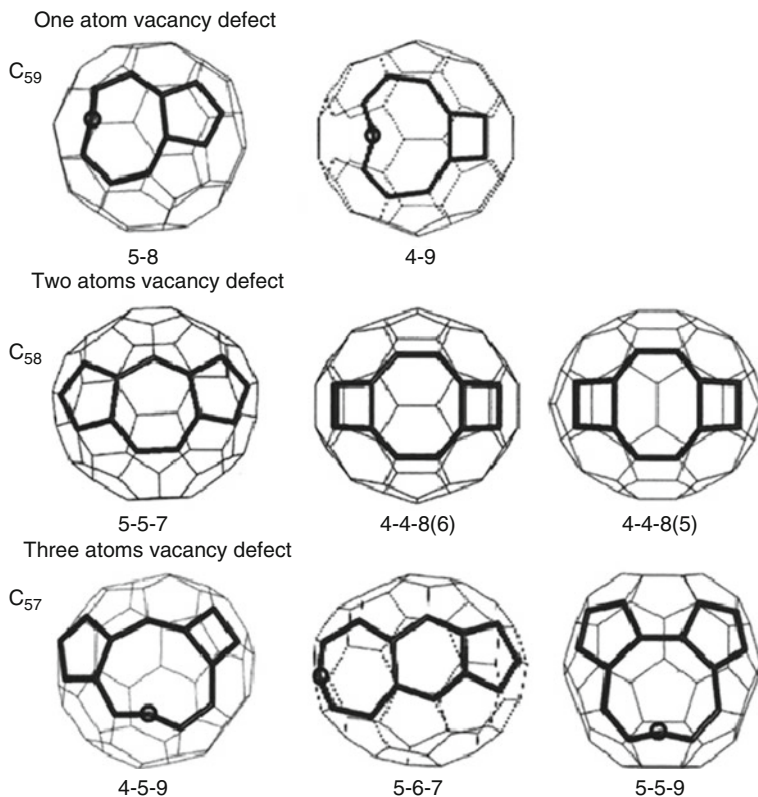


Fig. 34 B3LYP/6-31G(d) optimized structures of C_{59} , C_{58} , and C_{57} clusters. Description indicates highlighted rings. “A-B” denotes A- and B-membered ring. Circle denotes an unsaturated atom (Reprinted with permission from reference Lee and Han (2004). Copyright 2004, American Institute of Physics)

The singlet structures are more stable than the triplet ones for C_{58} cluster, while the reverse is true in the case of C_{57} clusters. The reported stabilization energy per atom at the B3LYP/6-311G(d)//B3LYP/6-31G(d) level is 2.18, 1.49, and 3.10 kcal/mol for the C_{59} , C_{58} , and C_{57} , respectively. Quantum chemical calculations provide relationship between structure and stability of the defect fullerene clusters (Hu and Ruckenstein 2003, 2004; Lee and Han 2004). In case of removal of four adjacent carbon atoms in C_{60} , additional five-membered rings are formed in geometry optimizations (e.g., isomer **1** in Fig. 35). The isomer **4** has only five- and six-membered rings (12 five-membered rings and 18 six-membered rings) and was predicted to be the most stable among the isomers depicted in Fig. 35. The stability energy for the isomers generated by removing four carbon atoms has the following sequence: isomer **4** > isomer **3** > isomer **2** > isomer **5** > isomer **6** > isomer **1**. All defect clusters have lower stability energy per atom than C_{60} . The removal of carbon atoms from C_{60} increases the HOMO and decreases the

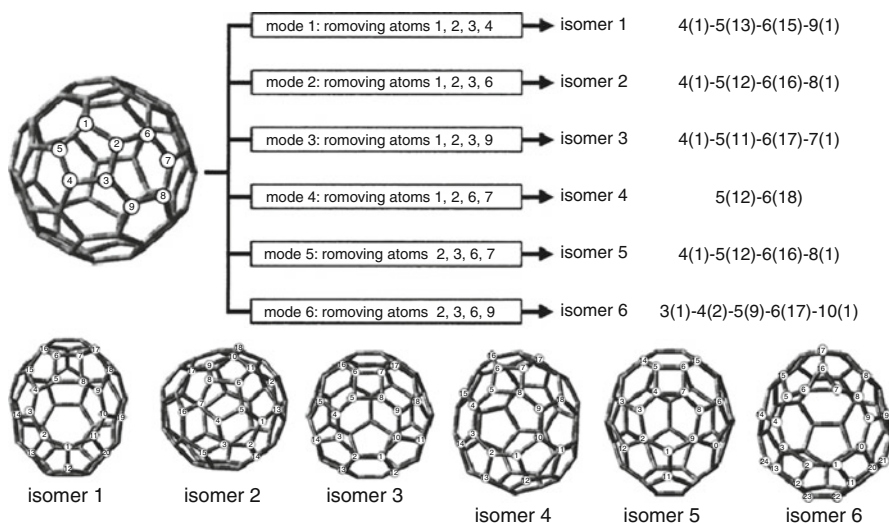


Fig. 35 Different modes to generate isomers of C_{60} with four vacancies by removing four adjacent atoms from the perfect C_{60} structure. Structures of isomers 1–6 of defect C_{60} with four vacancies. The ring size and the number of rings (in parentheses) for each isomer are given; for example, isomer 1-4(1)-5(13)-6(15)-9(1) means 1 four-membered, 13 five-membered, 15 six-membered, and 1 nine-membered rings (Reprinted with permission from reference Hu and Ruckenstein (2004). Copyright 2004, American Institute of Physics)

LUMO energy. Consequently, the defect structures exhibit lower HOMO–LUMO gap compared to C_{60} . No relationship was obtained between the stability energy per carbon atom and the HOMO–LUMO gap for the defective carbon clusters of C_{60} (Hu and Ruckenstein 2003).

Computational Studies of Single-Walled Carbon Nanotubes

Computational chemists explored the structures, electronic properties, and reactivities of SWCNTs of varying lengths and diameters (Bettinger 2004; Dinadayalane and Leszczynski 2009; Dinadayalane et al. 2007b; Galano 2006; Kaczmarek et al. 2007; Matsuo et al. 2003; Niyogi et al. 2002; Yang et al. 2006c). They also tried to understand the influence of different defects on these properties at reliable theoretical methods within the limitations of hardware and software (Akdım et al. 2007; Andzelm et al. 2006; Bettinger 2005; Dinadayalane and Leszczynski 2007a, b; Govind et al. 2008; Lu et al. 2005; Nishidate and Hasegawa 2005; Wanbaylor and Ruangpornvisuti 2008; Wang et al. 2006; Yang et al. 2006a, b). A series of finite-length hydrogen-terminated armchair SWCNTs have been computationally studied to obtain knowledge on the influence of diameter and length on the structural and electronic properties (Galano 2006). The optimized armchair (n,n) SWCNTs possess D_{nh} and D_{nd} point groups for $\vartheta/2$ even and odd, respectively. The different

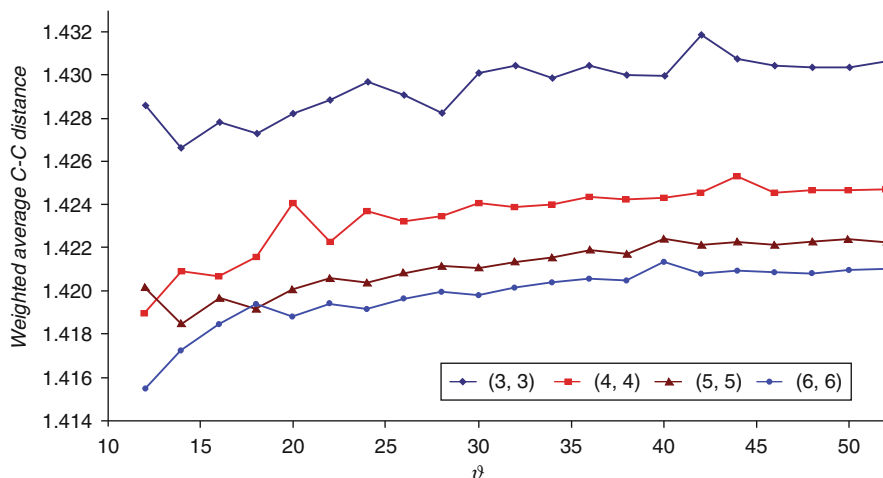


Fig. 36 Calculated weighted average values of the C–C distance as a function of the tube length for armchair (n,n) SWCNTs (Reprinted with permission from reference Galano (2006). Copyright 2006 Elsevier)

lengths of the armchair SWCNTs have the general formula $C_{(2n)k}H_{4n}$ with $k = \vartheta/2$. Galano considered (3,3), (4,4), (5,5), and (6,6) armchair SWCNTs with k of 6 to 26 (i.e., from 6 to 26 carbon atom layers). There are two types of bonds in the perfect (n,n) armchair SWCNTs; one is perpendicular to the tube axis (r_I) and another one is nearly parallel to the tube axis (r_{II}). The maximum difference between r_I and r_{II} was obtained in the case of the narrow diameter (3,3) tube. The influence of diameter on the weighted average values of the C–C distances is larger than the influence of the tube length (Fig. 36; Galano 2006).

The frontier molecular orbitals (HOMO and LUMO) play an important role in SWCNTs since they are helpful in predicting a number of ground-state properties of molecules. According to Huckel theory, the (n,n) armchair SWCNTs should be metallic (Saito et al. 1998), but the finite-length armchair SWCNTs are semi-conducting with a finite size of the HOMO–LUMO gap (Cioslowski et al. 2002). The computed HOMO–LUMO gaps for (3,3) to (6,6) SWCNTs were reported to be lower than the corresponding value for fullerene C_{60} . The HOMO–LUMO gap oscillates as the tube length increases for all of these armchair tubes (Fig. 37). The behavior of narrow diameter (3,3) tube is different from other armchair SWCNTs (Galano 2006).

Matsuo et al. classified the structures of finite-length armchair (5,5) and (6,6) SWCNTs as Kekule, incomplete Clar, and complete Clar networks depending on the exact length of the tubes. The (5,5) and (6,6) SWCNTs were elongated layer by layer of 10 and 12 carbon atoms, respectively (Matsuo et al. 2003). The local aromaticity of different lengths of the tubes was evaluated using the NICS (nucleus-independent chemical shift) calculations (GIAO-SCF/6-31G*//HF/6-31G* level).

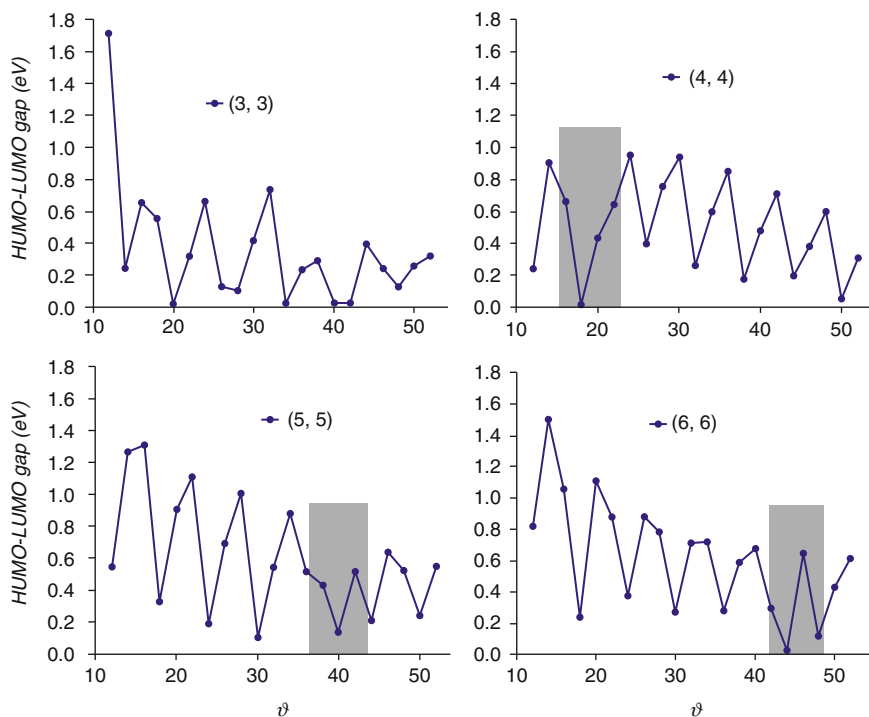


Fig. 37 Variations of HOMO–LUMO gaps as increasing the tube length for the armchair SWCNTs. There is no periodicity in (3,3) tube and the shaded region indicated the broken periodicity in other tubes (Reprinted with permission from reference Galano (2006). Copyright 2006 Elsevier)

Matsuo et al. pointed out that the geometry of $C_{50}H_{10}$ is similar to the equatorial belt of the fullerene C_{70} . Bond lengths of optimized structures exhibit oscillation with increase in tube length for both (5,5) and (6,6) armchair SWCNTs. The schematic structures of Kekule, incomplete Clar, and complete Clar networks for (5,5) and (6,6) SWCNTs are depicted in Fig. 38 along with the NICS values of dissimilar benzenoid rings. The energy of frontier molecular orbitals and HOMO–LUMO gap also oscillate as the length of the nanotube increases (Fig. 39). The Kekule structure shows larger HOMO–LUMO gap than the other two. It was reported that the band gap will eventually disappear at a certain tube length (Matsuo et al. 2003).

The pyramidalization angle (θ_P) and π -orbital misalignment angles are useful to gauge the reactivity of the carbon atom sites of SWCNTs. The end caps of SWCNTs resemble fullerene hemisphere; thus, the end caps are expected to be more reactive than sidewalls irrespective of the diameter of the SWCNTs. Carbon atoms in fullerene are more distorted than those in the corresponding SWCNTs. For example, the carbon atom of (10,10) armchair SWCNT has the pyramidalization angle (θ_P) of about 3.0° , while the carbon atom of the fullerene with corresponding

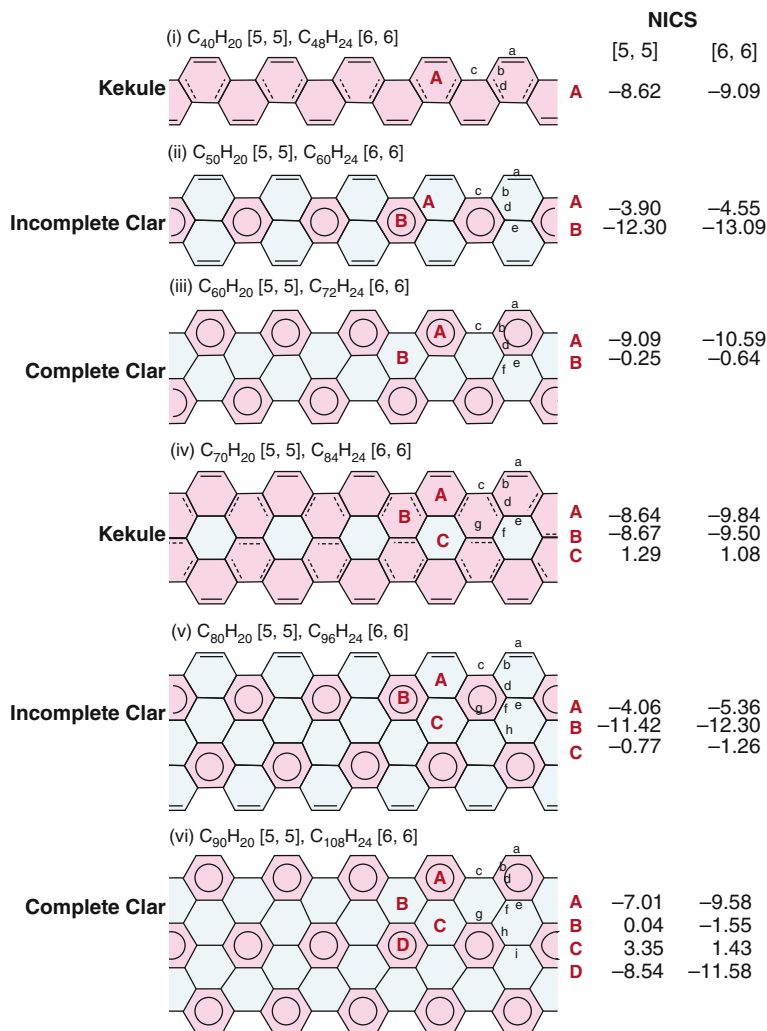


Fig. 38 Schematic structures and color-coded NICS maps of finite-length (5,5) and (6,6) SWCNTs. Hydrogen atoms are omitted for clarity. Chemical bonds are schematically represented by using single bond (*solid single line*; bond length > 1.43 Å), double bond (*solid double line*; bond length < 1.38 Å), single bond halfway to double bond (*solid dashed line*; 1.43 Å > bond length > 1.38 Å), and Clar structure (i.e., ideal benzene). NICS coding: *red*, aromatic < -4.5; *blue*, nonaromatic > -4.5 (Reprinted with permission from reference Matsuo et al. 2003. Copyright 2003 American Chemical Society)

radius (fullerene C_{240}) has the θ_P of about 9.7° (the hemisphere of fullerene C_{240} can be capped to (10,10) SWCNT) (Niyogi et al. 2002). Chen et al. mentioned that the pyramidalization angle of the C atoms of the sidewalls of SWCNTs is smaller compared to that of the fullerenes of same radius. As a consequence, the

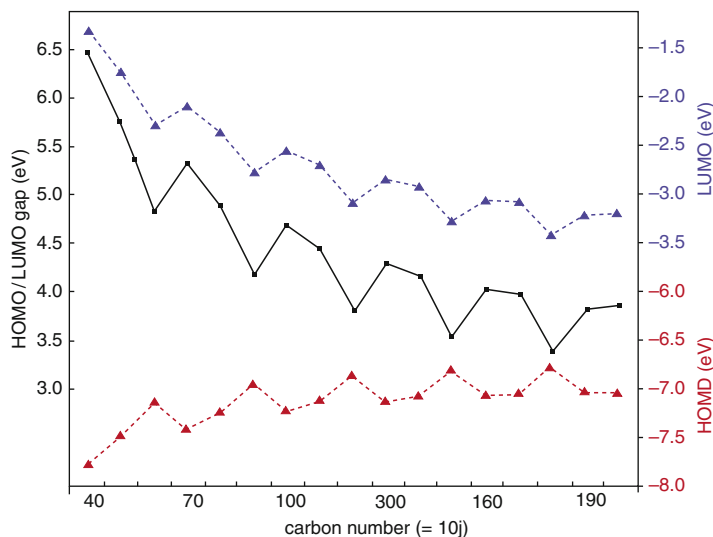


Fig. 39 Variation of HOMO–LUMO energies and HOMO–LUMO energy gap with increase in tube length of (5,5) armchair SWCNT ($C_{10j}H_{20}$) (Reprinted with permission from reference Matsuo et al. (2003). Copyright 2003 American Chemical Society)

covalent functionalization to SWCNTs is less favorable compared to fullerenes of same radius (Chen et al. 2003). The π -orbital misalignment is likely to be a main source of strain in the SWCNTs. For both armchair and zigzag SWCNTs, the pyramidalization angle and the π -orbital misalignment angle decrease with increase in diameter of the tube (Fig. 40; Niyogi et al. 2002).

Covalent Functionalization of SWCNTs: H and F Atom Chemisorptions

The covalent functionalization of SWCNTs, which modifies the properties of the tubes, has become a challenging field of research for the past few years (Bettinger 2006; Hirsch 2002; Niyogi et al. 2002; Vostrowsky and Hirsch 2004). Functionalization of tubes is considered to be promising to produce carbon nanotube-based materials for selective applications (Bettinger 2006; Cho et al. 2008; Denis et al. 2009). The binding of hydrogen with SWCNTs has generated a lot of experimental and theoretical interests due to their potential application in hydrogen storage (Dillon et al. 1997; Dinadayalane and Leszczynski 2009; Dinadayalane et al. 2007b; Kaczmarek et al. 2007; Nikitin et al. 2005; Ormsby and King 2007; Yang et al. 2006c; Zhang et al. 2006). Scientists have tried to obtain knowledge on the mechanism of hydrogen adsorption in SWCNTs. They attempted to design the viable nanotube-based hydrogen storage material to meet the Department of Energy (DOE) target of 6.5 wt% at ambient temperature (Dillon et al. 1997; Dinadayalane

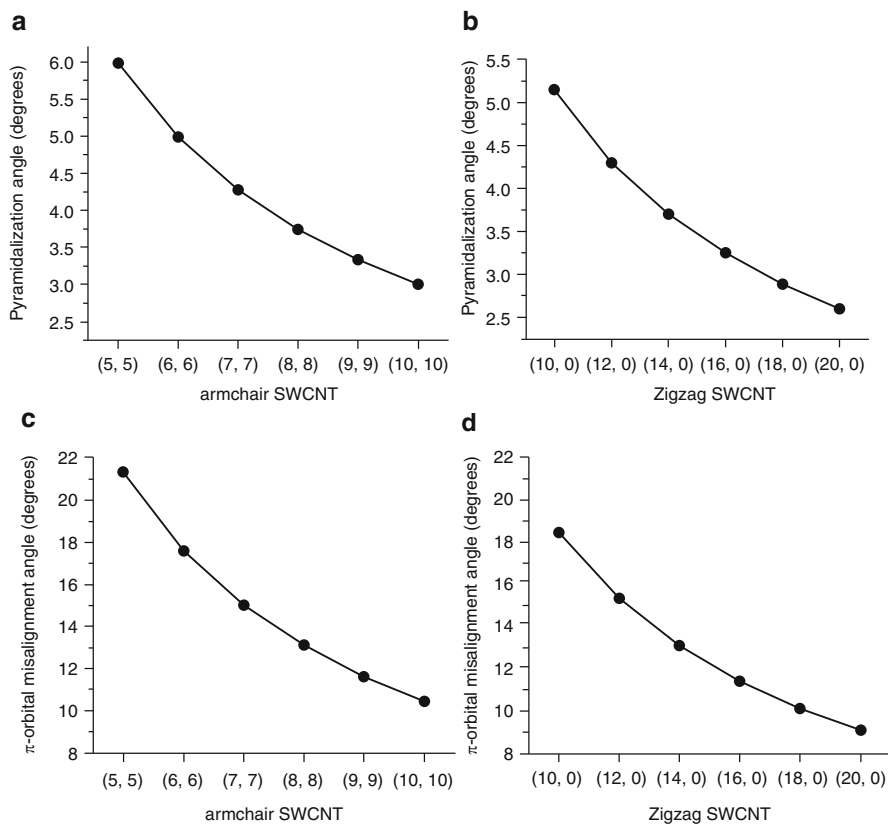


Fig. 40 The change of pyramidalization angle (**a, b**) at the carbon atom and the π -orbital misalignment angle (**c, d**) between two adjacent carbon atoms of armchair (**a, c**) and zigzag (**b, d**) SWCNTs. The π -orbital misalignment angle is zero for the carbon atoms of the circumferential bond in armchair tube and axial bond in zigzag tube (The data was taken from reference Niyogi et al. (2002))

and Leszczynski 2009; Dinadayalane et al. 2007b; Kaczmarek et al. 2007; Nikitin et al. 2005; Ormsby and King 2007; Yang et al. 2006c; Zhang et al. 2006). The experimental investigations reported the chemisorption of H atoms on the surface of SWCNTs as promising approach to meet DOE's target of hydrogen storage (Nikitin et al. 2005; Zhang et al. 2006). Chemisorptions of hydrogen atoms on the surface of SWCNTs were investigated (Dinadayalane and Leszczynski 2009; Dinadayalane et al. 2007b; Kaczmarek et al. 2007; Ormsby and King 2007; Yang et al. 2006c). The covalent functionalization of SWCNTs by H atoms is a hot topic. Thus, we discuss the quantum chemical studies of the hydrogen chemisorption on different types of SWCNTs.

Yang et al. studied, using DFT and ONIOM calculations, the chemisorption of atomic hydrogen(s) on the open-ended finite-sized (5,0), (7,0), and (9,0) zigzag and

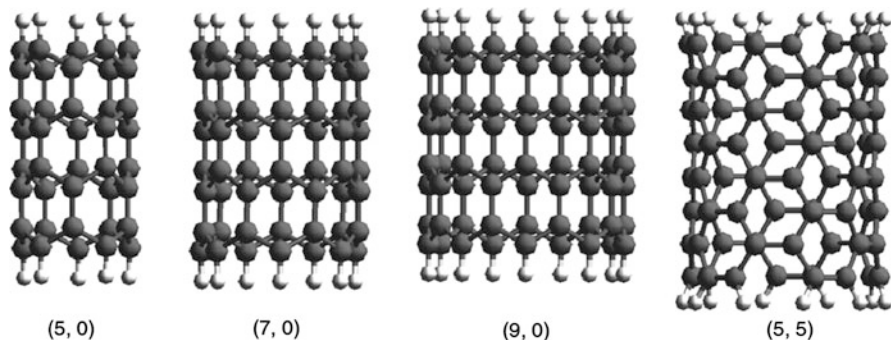


Fig. 41 Finite-sized small carbon nanotube models of (5,0), (7,0), and (9,0) zigzag and (5,5) armchair SWCNTs considered for low occupancy of H chemisorptions (Reprinted with permission from reference Yang et al. (2006c). Copyright 2006 American Chemical Society)

(5,5) armchair SWCNTs (Fig. 41; Yang et al. 2006c). They compared the binding energies obtained for nanotubes with results of the model graphene sheet in order to examine the effect of curvature. It was reported that the chemisorptions of H atoms to the exterior wall of the SWCNTs are more favorable than the interior walls. The H chemisorption has strong dependence of tube diameter and helicity or chirality in both interior (endohedral) and exterior (exohedral) addition. In case of single H atom addition, the binding energy (chemisorption energy), which is the reaction energy for H chemisorption with SWCNT, decreases with increase in tube diameter. In the chemisorption of two hydrogen atoms in the interior and exterior walls of (5,0) and (7,0) SWCNTs, two hydrogen atoms prefer to bind at alternate positions rather than adjacent positions. This was attributed to the crowding effect when two hydrogen atoms occupy in the adjacent positions. In the case of (5,0) SWCNT, chemisorption of ten hydrogen atoms (33 % coverage) decreases the magnitude of chemisorption energy, which is further decreased by an increase in the coverage to 50 %. Similar to the situation in zigzag SWCNTs, two hydrogen atoms prefer to attach at alternate carbon sites rather than adjacent sites in the graphene sheet. Significantly large deviation of chemisorption energy between the graphene sheet and zigzag SWCNTs (H atoms chemisorbed on the exterior wall) was reported. It was found that the chemisorptions of H atoms with small diameter SWCNTs are much more favorable than with the graphene sheet (Yang et al. 2006c).

We investigated a single-hydrogen chemisorption, the preference of the positions (i.e., 1-2, 1-2', 1-3, or 1-4 positions) for the chemisorption of two hydrogen atoms considering (3,3), (4,4), (5,5), and (6,6) armchair SWCNTs of 9 and 15 carbon atom layers. The SWCNTs of 15 carbon layers considered in our study are shown in Fig. 42. The addition of H atoms on the outer wall of SWCNT (exohedral addition) has only been considered in our study. We performed DFT calculations using B3LYP/6-31G(d) level for full geometry optimizations. The finite-length SWCNTs were capped with hydrogen atoms to avoid dangling bonds. The reaction energies for hydrogen chemisorption (E_r) on the external surface of

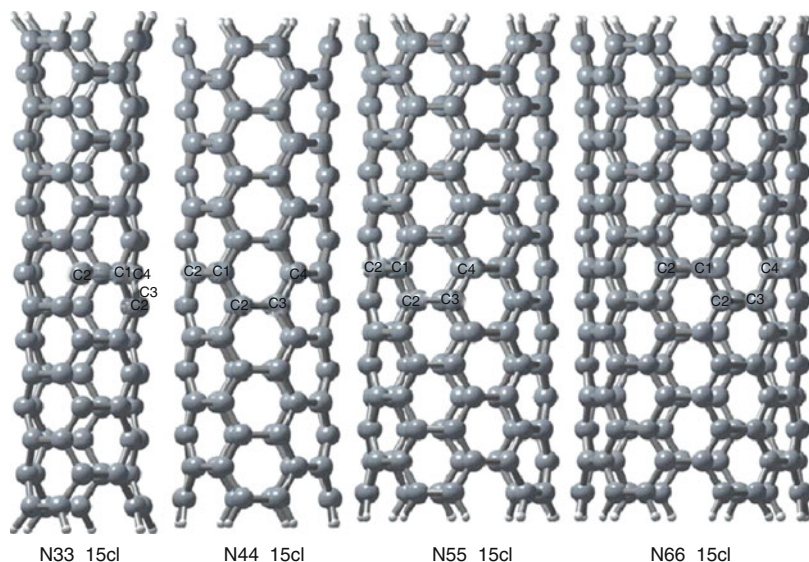


Fig. 42 Structures of (3,3), (4,4), (5,5), and (6,6) armchair SWCNTs of 15 carbon layers (15 cl) considered for the chemisorption of one and two H atoms. The carbon atom sites for attachment of H atoms are shown (Reprinted with permission from reference Dinadayalane et al. (2007b). Copyright 2007 American Chemical Society)

SWCNTs have been calculated using the formula $E_r = E_{\text{SWCNT}+n\text{H}} - E_{\text{SWCNT}} - nE_{\text{H}}$, where $E_{\text{SWCNT}+n\text{H}}$ denotes the total energy of hydrogen-chemisorbed nanotube, n represents the number of hydrogen atoms chemisorbed, and E_{SWCNT} and E_{H} correspond to the energies of pristine nanotube and the hydrogen atom, respectively. The reaction energy E_r can also be considered as hydrogen chemisorption energy. The chemisorption of hydrogen is an exothermic process if the value of E_r is negative (Dinadayalane et al. 2007b).

We observed the rupture of circumferential C1–C2 bond when two hydrogen atoms were chemisorbed in the case of (3,3) SWCNT of 15 carbon layers. As shown in Fig. 43, the reactions of single as well as two hydrogen chemisorptions on the surface of armchair SWCNTs are highly exothermic. The reaction energy for the addition of two H atoms is more than two times that of one H chemisorption except for H(1,3) addition. Our computational study revealed a competition between H(1,2) and H(1,4) addition in the case of (5,5) and (6,6) SWCNTs, but such competition was not seen in the case of narrow diameter (3,3) and (4,4) SWCNTs. Increasing the length of the tube has pronounced effect on the reaction energy of hydrogen chemisorption. In case of armchair SWCNTs, the chemisorption of two hydrogen atoms at alternate positions is thermodynamically less favored compared to H(1,2) and H(1,2') additions regardless of the length and diameter of the tubes (Dinadayalane et al. 2007b). The least positional preference of H(1,3) for armchair SWCNTs is different compared to the results of zigzag-type nanotubes (Yang et al.

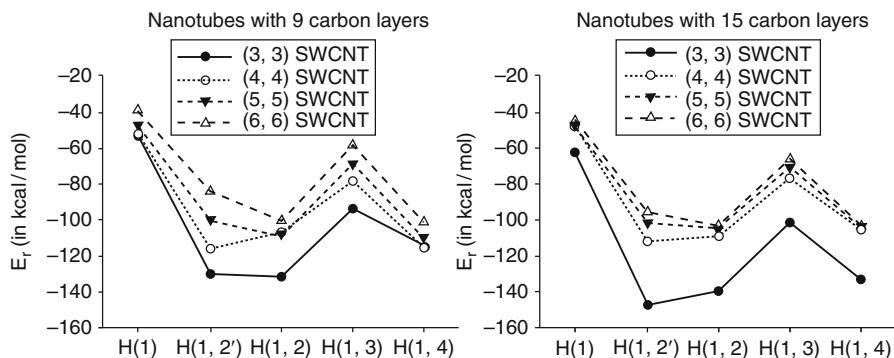


Fig. 43 The variation of reaction energies at the B3LYP/6-31G(d) level for the chemisorption of one and two hydrogen atoms on the external surface of (3,3), (4,4), (5,5), and (6,6) armchair single-walled carbon nanotubes (SWCNTs) (Reprinted with permission from reference Dinadayalane et al. (2007b). Copyright 2007 American Chemical Society)

2006c; Dinadayalane et al. 2007b). We found that the H chemisorption on nanotubes of different diameters and the positions of two hydrogen atoms chemisorbed on the surface of armchair SWCNTs can be characterized by C–H stretching frequencies of chemisorbed hydrogen atoms (Dinadayalane et al. 2007b). In the investigation of chemisorption of H atoms with (3,3) and (4,4) SWCNTs of different lengths, we found that changing the length of the nanotube has significant effect on the reaction energy of hydrogen chemisorption and HOMO–LUMO gap of pristine and hydrogen-chemisorbed SWCNTs (Kaczmarek et al. 2007).

The reactivity pattern was predicted for the hydrogenation in chiral SWCNTs (Ormsby and King 2007). Investigations involving chiral SWCNTs are more challenging than for zigzag and armchair SWCNTs because single unit cell contains many atoms; consequently, more computational resources are required. Computational study demonstrated that hydrogenation of the fully benzenoid (12,9) SWCNT was significantly less energetic (by 8 kcal/mol per mol H₂) than the hydrogenation of (12,7) and (12,8) SWCNTs (Fig. 44). Furthermore, the hydrogenation at an internal Clar double bond or bonds was reported to be more exothermic than at randomly selected internal bonds. Like other polycyclic aromatic hydrocarbons, hydrogenation of double bonds is energetically preferred over hydrogenation of aromatic sextets. The frontier molecular orbitals (HOMO and LUMO) of chiral SWCNTs have maximum amplitude at the double bonds suggesting that Clar's model also predicts the kinetic reactivity. Thus, the frontier molecular orbitals are useful in predicting the favorable sites for hydrogenation in chiral SWCNTs (Ormsby and King 2007).

In the early 2000s, experimentalists found that partial fluorination of the SWCNTs could be used as a technique for cutting the nanotubes of varying lengths. However, the mechanism of cutting the nanotube is not clear (Gu et al. 2002). In a computational study, Bettinger observed a strong oscillation of reaction energy for

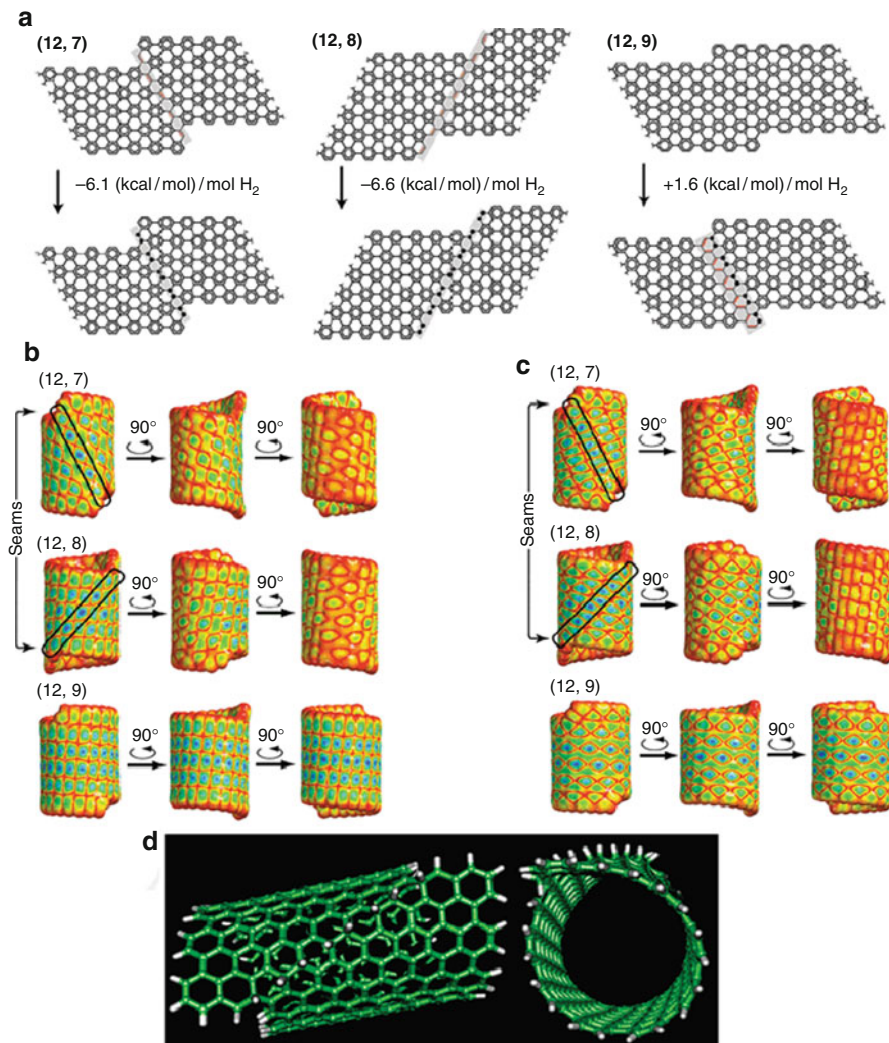


Fig. 44 (a) Hydrogenation of equivalent vectors of (12,7), (12,8), and (12,9) chiral SWCNT segments (planar representation). (b) HOMO plotted on the isodensity surface for (12,7), (12,8), and (12,9) chiral SWCNTs. (c) LUMO plotted on the isodensity surface for (12,7), (12,8), and (12,9) chiral SWCNTs. (d) The hydrogenated model chiral SWCNT (*side* and *top* views). The locations of double bond are indicated in (b) and (c). HOMO and LUMO isodensity surface structures generated at AM1 method (Reprinted with permission from reference Ormsby and King (2007). Copyright 2007 American Chemical Society)

the addition of F atom on the external surface of (5,5) SWCNT of varying lengths (Fig. 45). The computed reaction energy oscillation ranges from 43 to 68 kcal/mol at the UB3LYP/6-31G(d) level using UPBE/3-21G optimized geometries. The shortest tube exhibited the highest exothermicity. The energy oscillation was reported to be

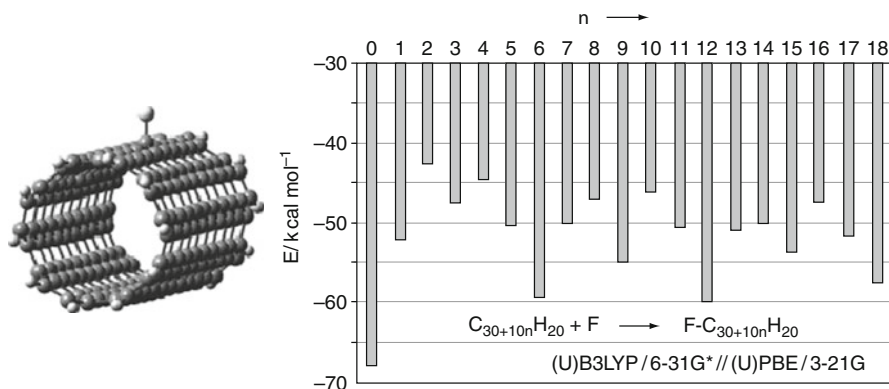


Fig. 45 Variation of reaction energy for the addition of F atom to (5,5) armchair SWCNT of different lengths (n – increasing number of carbon layers) (Reprinted with permission from reference Bettinger (2004). Copyright 2004 American Chemical Society)

periodic with large exothermicity for the fully benzenoid frameworks, in agreement with their smaller band gaps compared to Kekule and incomplete Clar structures. Computational study demonstrated that the addition of F atom to the sidewall of SWCNT strongly depends on the length of the nanotube (Bettinger 2004). As observed in H atom addition (Dinadayalane et al. 2007b; Kaczmarek et al. 2007; Yang et al. 2006c), the F atom addition to the sidewalls of SWCNTs transforms the carbon atom hybridization from sp^2 to sp^3 (Bettinger 2004).

The chemical reactivity of carbon nanotubes is governed by the local atomic structure. As mentioned earlier, the pyramidalization angle is an important parameter in predicting the chemical reactivity of SWCNTs. Park et al. predicted the hydrogenation and fluorination energies of each carbon from its pyramidalization angle for zigzag SWCNTs (Fig. 46). They formulated the E_{total} for chemisorption of H and F atom on the external surface of zigzag tubes of different diameters as a function of pyramidalization angle of the binding site of tubes. They revealed that the metallic zigzag SWCNTs are slightly more reactive than the semiconducting SWCNTs. Furthermore, the fluorination is more viable than the hydrogenation (Park et al. 2003).

Theoretical Studies on Common Defects in SWCNTs

Investigating the atomic defects is important in tailoring the electronic properties of SWCNTs. Recent experimental study reported a method to selectively modify the electronic properties of semiconductor SWCNTs by the creation and annihilation of point defects on their surface with the tip of a scanning tunneling microscope (STM) (Berthe et al. 2007). Such experimental study motivates theoreticians to explore the structures, energetics, reactivities, and electronic properties of SWCNTs containing different types of defects.

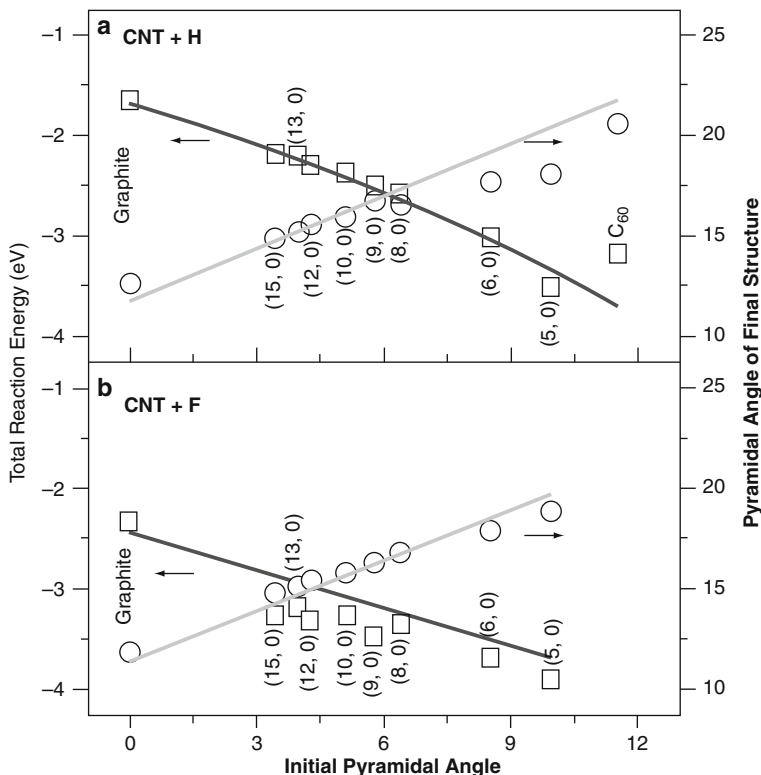


Fig. 46 DFT-computed E_{total} (□) and pyramidalization angle (θ_p) values (○) for fully relaxed configurations and their estimated values (solid curves) for (a) hydrogenation and (b) fluorination (Reprinted with permission from reference Park et al. (2003). Copyright 2003 American Chemical Society)

Stone–Wales Defect

The Stone–Wales defect can be created by 90° rotation of one of the C–C bonds in the hexagonal network of SWCNTs. Two types of C–C bonds exist in each of armchair and zigzag SWCNTs. Therefore, one can generate Stone–Wales defect in two different orientations in both armchair and zigzag SWCNTs (Fig. 47). DFT calculations revealed that the formation energies of (5,5) SWD_II and (10,0) SWD_II are lower than those of (5,5) SWD_I and (10,0) SWD_I. The computed formation energies of (5,5) SWD_I, (5,5) SWD_II, (10,0) SWD_I, and (10,0) SWD_II are 66.4 (2.88 eV), 57.0 (2.47 eV), 68.9 (2.99 eV), and 63.0 (2.73 eV) kcal/mol, respectively. The formation energy was calculated as the relative energy of the Stone–Wales defective tube with respect to the corresponding defect-free SWCNT. It was reported that the formation of Stone–Wales defect causes no change in the HOMO–LUMO gaps (Yang et al. 2006a, b).

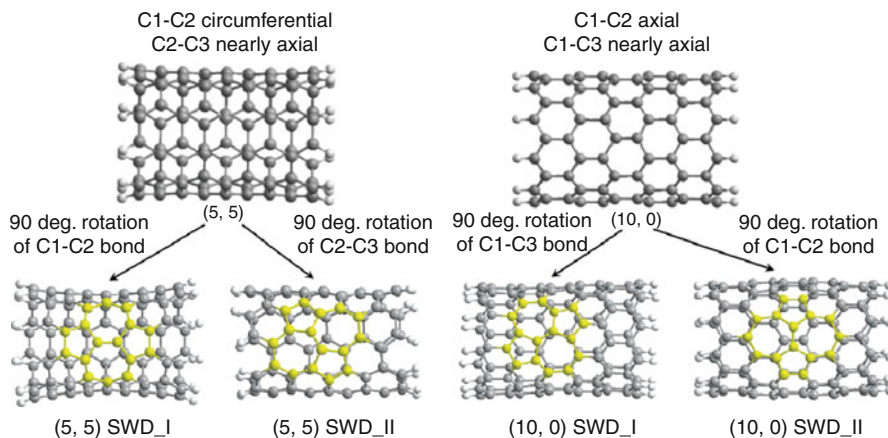


Fig. 47 Generation of Stone–Wales defect with different orientations in (5,5) armchair and (10,0) zigzag SWCNTs. The atoms in the Stone–Wales defect region are highlighted in *yellow color* (Reprinted with permission from reference Yang et al. (2006a). Copyright 2006 American Chemical Society. Reprinted with permission from Yang et al. (2006b). Copyright 2006, American Institute of Physics)

We investigated the structures, formation energies, and reactivities of Stone–Wales defect with two different orientations and different locations from the end of tube in armchair (5,5) SWCNTs of $C_{80}H_{20}$ (I) and $C_{100}H_{20}$ (II) (Fig. 48a; Dinadayalane and Leszczynski 2007b). We employed HF/4-31G, HF/6-31G(d), B3LYP/3-21G, and B3LYP/6-31G(d) levels of theory. HF/4-31G level overestimates the Stone–Wales defect formation energy compared to B3LYP/6-31G(d) level. Our study revealed that B3LYP/3-21G level, which provides reasonable energy estimation, may be employed for large nanotube systems to compute the defect formation energies when the calculations at the B3LYP functionals with large basis sets are prohibitive. Our study showed that the Stone–Wales defective (5,5) armchair SWCNTs generated by rotation of nearly axial bond (ABR) are more stable than those created by circumferential bond rotation (CBR) as shown in Fig. 48b. The SW defect structures generated by ABR show lower HOMO–LUMO gap than those created by CBR and the defect-free SWCNTs (Dinadayalane and Leszczynski 2007b).

Bettinger demonstrated in a comprehensive computational study that some of the bonds of SW defect show higher reactivity than pristine tube; others are less reactive (Bettinger 2005). Computational studies explained the reactivity of carbon atom sites based on the pyramidalization angles (Akdim et al. 2007; Bettinger 2005; Lu et al. 2005). Lu et al. investigated addition of O, CH_2 , and O_3 across C–C bonds of SW defective and defect-free armchair SWCNTs. They showed that the central C–C bond of the SW defect in armchair SWCNT (SW defect generated by CBR) is chemically less reactive than that in the perfect tube, and it was attributed to small local curvature in the carbon atoms of central C–C bond of the SW defect (Lu et al. 2005). We found that the values of pyramidalization angles do not completely

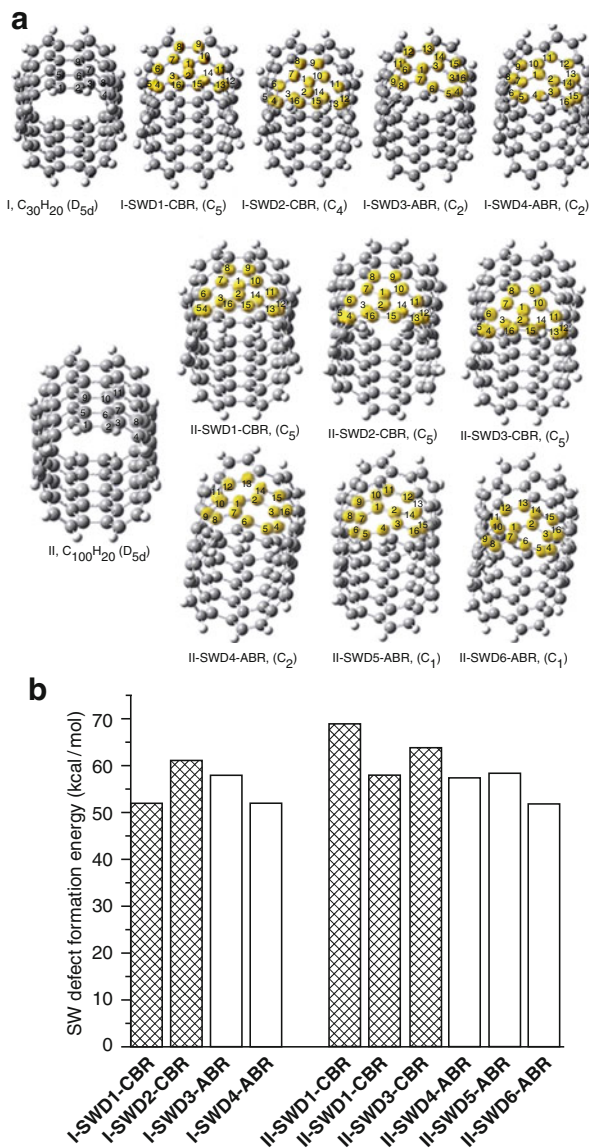


Fig. 48 (a) B3LYP/6-31G(d) level optimized structures of defect-free (5,5) armchair SWCNTs and the Stone–Wales defect tubes generated by the 90° rotation of circumferential and nearly axial C–C bonds. (b) The Stone–Wales defect formation energy obtained at the B3LYP/6-31G(d) level (Reprinted with permission from reference Dinadayalane and Leszczynski (2007b). Copyright 2007 Elsevier)

explain the reactivity of different bonds of SW defect region for cycloaddition reactions and the reactivity may arise from various other reasons, in addition to topology. We concluded that the cycloaddition reactions across the C–C bond

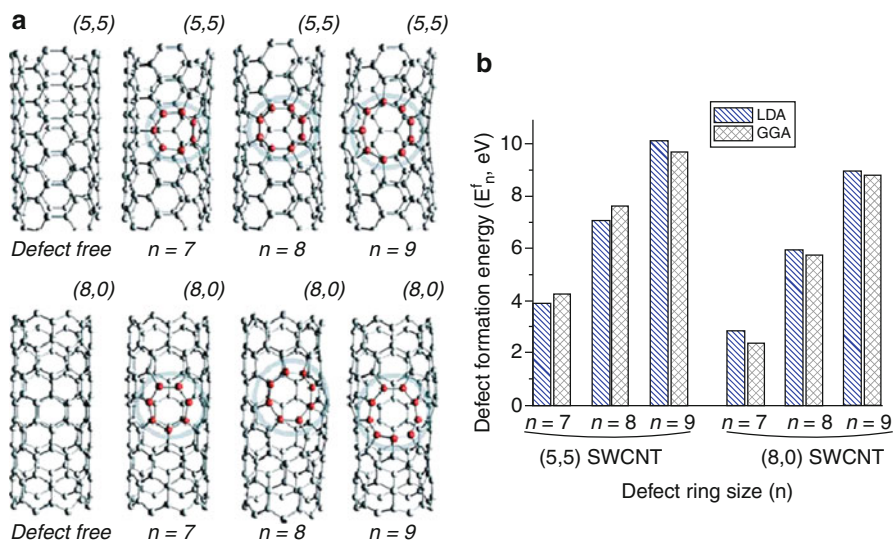


Fig. 49 (a) Fully relaxed structures of the defect-free and the defective (5,5) (upper panel) and (8,0) (lower panel) SWCNTs obtained by the GGA calculations. Gray balls and rods are the carbon atoms and bonds shorter than 1.5 Å, respectively. Carbon ring defects (given in red color) are indicated by the thick circles. (b) Defect formation energy (E_f^n) of n -membered carbon rings for (5,5) and (8,0) calculated using LDA and GGA (The pictures in (a) and the data for (b) were taken with permission from Nishidate and Hasegawa (2005). Copyright 2005 by the American Physical Society)

shared by two heptagons (7-7 ring fusion) need not always be less reactive than the corresponding bond in the pristine structure and the reactivity of that bond depends on the orientation of the SWD in the SWCNTs (Dinadayalane and Leszczynski 2007b).

Topological Ring Defects

Nishidate and Hasegawa calculated the formation energies of n -membered topological ring defects with $n = 7$ (heptagon), $n = 8$ (octagon), and $n = 9$ (enneagon) in (5,5) armchair and (8,0) zigzag SWCNTs (Fig. 49a). They used both local density approximation (LDA) and the generalized gradient approximation (GGA: PW91). The spin-polarized projector augmented-wave (PAW) implemented in VASP code was employed for calculations, and periodic boundary condition was used (Nishidate and Hasegawa 2005).

The defect formation energy (E_f^n) of n -membered rings was calculated as the energy difference between the total energy of defective SWCNT (E_{tot}^n) and that of the pristine SWCNT (E_{tot}); i.e., $E_f^n = E_{tot}^n - E_{tot}$. The number of atoms of each defective SWCNT is the same as that of the corresponding SWCNT. Computational

study showed that the defect formation energy increases with increase in defect ring size (n) (Fig. 49b). In general, LDA method yielded higher defect formation energy than generalized gradient approximation. The SWCNTs were reported to be more fragile than the graphene sheet for defect formation. Distortion of the SWCNTs became larger as the defect ring size increases (Nishidate and Hasegawa 2005).

Single- and Di-vacancy Defects

An ideal single vacancy (SV) with three dangling bonds (DBs) was generated by removing one carbon atom from the perfect (5,5) and (10,0) SWCNTs. Upon geometry optimization, an ideal SV with three dangling bonds rearranged into a pentagonal ring and one DB (Yang et al. 2006a). Hence, this defect is called a 5-1DB defect (Lu and Pan 2004; Yang et al. 2006a). In case of (5,5) SWCNT, the optimization of the ideal SV resulted in two different 5-1DB defects. The structures were named as (5,5) SV_I and (5,5) SV_II as shown in Fig. 50. The latter structure

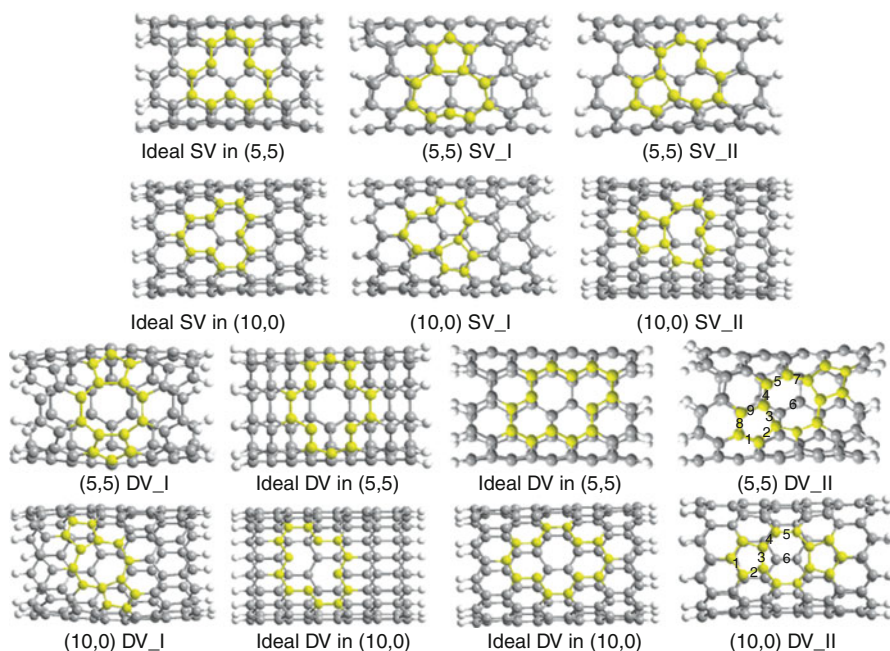


Fig. 50 Configurations of the single (SV) and double (DV) vacancies. Ideal SV and DV mean one and two carbon atoms removed from pristine SWCNT and the structures were not relaxed. Single and double vacancies in two different orientations (I and II) given for (5,5) and (10,0) SWCNTs. Carbon atoms in defect region are given in *yellow color* (Reprinted with permission from reference Yang et al. (2006a). Copyright 2006 American Chemical Society)

was reported to be energetically more favorable (by 1.20 eV) than the former one. A similar behavior was observed for (10,0) zigzag SWCNT. The bond length of the new C–C bond forming five-membered ring is 1.64 Å in (5,5) SV_I, while that of 1.55 Å was obtained for (5,5) SV_II (Yang et al. 2006a). Lu and Pan found using tight-binding calculations that the single-vacancy defect formation energy for (n,n) armchair SWCNTs increases monotonically with increasing tube radii. The formation energy curve of single-vacancy defects in the zigzag (n,0) SWCNTs is periodic, which is mainly characterized by metallic (n,0) tubes (such as (6,0), (9,0), (12,0), (15,0), etc.) (Lu and Pan 2004).

An ideal di-vacancy can be generated by removing two carbon atoms. Two different orientations are possible because of the presence of two different types of bonds in both (5,5) armchair and (10,0) zigzag SWCNTs. Upon geometry optimizations, SWCNTs with ideal di-vacancies yielded 5-8-5 (five-eight-five-membered rings) defects in two different orientations in both armchair- and zigzag-type tubes. Computational study revealed that (5,5) DV_II and (10,0) DV_II configurations are energetically more favorable than (5,5) DV_I and (10,0) DV_I by 0.97 and 0.63 eV, respectively. (5,5) DV_II and (10,0) DV_II can be obtained by removing the carbon atoms with the dangling bond from (5,5) SV_II and (10,0) SV_II, which are the most stable configurations among the possible types of SVs in each type of tubes (Yang et al. 2006a).

The di-vacancy in graphene as well as SWCNTs generates structures possessing two pentagons side by side with an octagon (585 structure) as a result of geometry optimization. The 585 configuration can reconstruct further into a complex structure composed of three pentagons and three heptagons, called 555777 defect structure. In fact, 555777 configuration is more stable than 585 configuration in graphene. Amorim et al. investigated the stability of these types of configurations, derived by di-vacancies, in armchair and zigzag SWCNTs considering different tube diameters (Amorim et al. 2007). The 585 defect in SWCNTs has two possible orientations with respect to the tube axis: perpendicular and tilted in armchair and parallel and tilted in zigzag SWCNTs. For the (5,5) SWCNT, the perpendicular orientation is less stable than the tilted one by 3.5 eV. In case of (8,0) SWCNT, the tilted orientation is less stable by 2.7 eV than the parallel one. Only the tilted and the parallel orientations of defects were considered for armchair and zigzag SWCNTs, respectively (Fig. 51; Amorim et al. 2007).

In contrast to graphene, the 585 defect was predicted to be more stable than the 555777 defect in both armchair and zigzag SWCNTs. Both 585 and 555777 defects in nanotubes (both armchair and zigzag) are more stable than in graphene. The defect formation energy increases monotonically as the diameter of the armchair SWCNT increases (Fig. 52a). The energy difference between the 585 defect and the 555777 defect was computed to be 1.6 eV in case of zigzag tubes, while the difference was reported to be 0.7 eV for armchair SWCNTs. Zigzag SWCNTs exhibited oscillations in the formation energies and the oscillations were related to the alternation between semiconductor and metallic character of the (n,0) zigzag SWCNTs (Fig. 52b; Amorim et al. 2007).

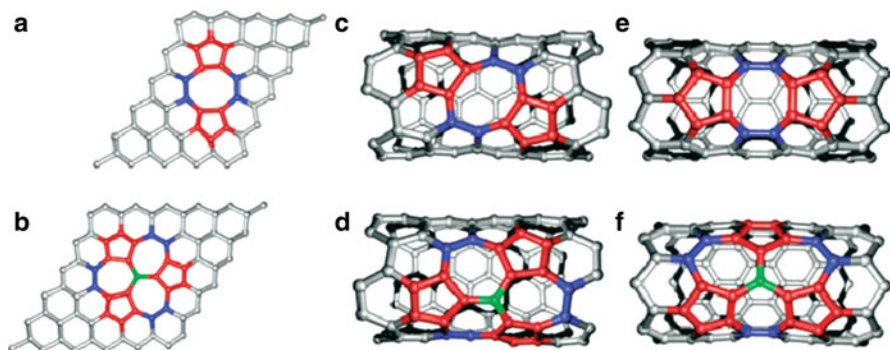


Fig. 51 Ball and stick models for the final geometries of the defects: (a) 585 and (b) 555777 in graphene; (c) 585 tilted and (d) 555777 in the (5,5) armchair SWCNT; (e) 585 parallel and (f) 555777 in the (8,0) zigzag SWCNT. The carbon atoms in pentagons are marked in *red*, the ones that complete either the octagons or the heptagons are colored in *blue*, and the ones at the center of the C_3 symmetry operation in the 555777 defects are colored in *green*. All the others are colored *gray* (Reprinted with permission from reference Amorim et al. (2007). Copyright 2007 American Chemical Society)

Outlook of Potential Applications of Carbon Nanostructures

Graphene is used as a base material for nanoelectromechanical systems (NEMS) due to its lightweight and stiffness properties (Bunch et al. 2007; Robinson et al. 2008). Functionalized graphene can be exploited for water splitting and hydrogen production. Scientists produce graphene-based materials with high structural and electronic quality for the preparation of transparent conducting electrodes for displays and touch screens. Solution processing and chemical vapor deposition are the ideal means to produce thin films that can be used as electrodes in energy devices such as solar panels, batteries, fuel cells, or in hydrogen storage (Bonaccorso et al. 2015). Graphene-based resonators have notable advantages in comparison with nanotubes. Reduced graphene oxide films are used to make drum resonators. The high Young's modulus, extremely low mass, and large surface area make the graphene-based resonators ideally suited for use as mass, force, and charge sensors (Ekinci et al. 2004; Knobel and Cleland 2003; Lavrik and Datskos 2003). Graphene can be used for metallic transistor applications and ballistic transport. One of the potential applications of graphene sheet is its use as membrane for separation (Jiang et al. 2009). Graphene may be useful for electro- and magneto-optics (Geim 2009). Graphane (fully hydrogenated graphene) nanoribbons have quite promising applications in optics and opto-electronics due to the wide band gap. Graphene may also be used for transistor applications (Novoselov et al. 2004). Research into applications for carbon graphene nanosheets has focused on their uses as platforms for next-wave microchips, active materials in field emitter arrays for flat panel screen displays; in gas sensors (Wang et al. 2015a), biological sensors, and medical imaging devices; in solar energy cells; and in high-surface area electrodes for use

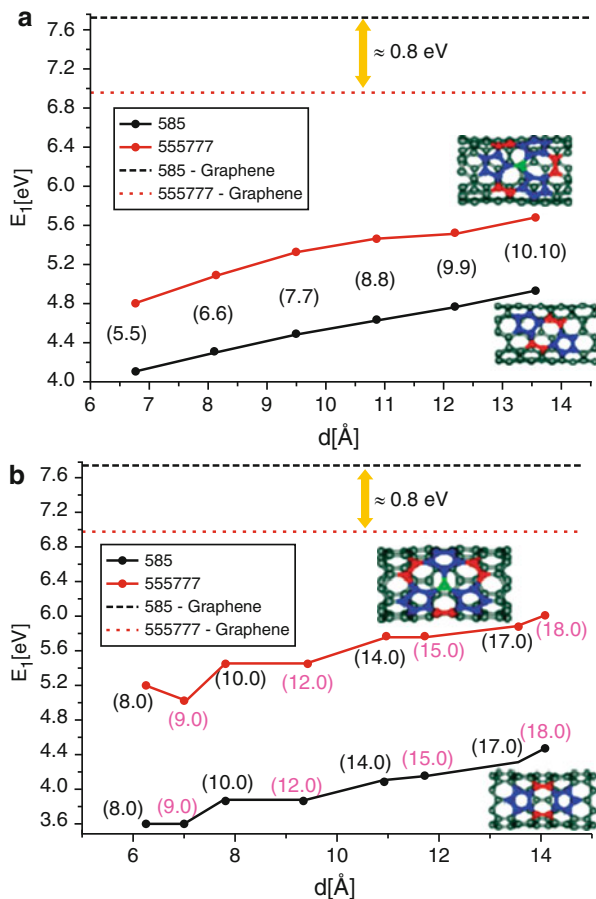


Fig. 52 Formation energy vs. diameter for the 585 and 555777 defects in (a) armchair and (b) zigzag SWCNTs. The limits for graphene were given in both graphs (Reprinted with permission from reference Amorim et al. (2007). Copyright 2007 American Chemical Society)

in bioscience. Graphene-based materials are known for energy and environmental applications (Bo et al. 2015; Chen et al. 2015; Yuan and He 2015). Graphene is a possible replacement material where carbon nanotubes are presently used (Xia et al. 2009). Graphene-based liquid crystal devices (LCD) show excellent performance with high contrast ratio. Thus, LCDs might be graphene's first realistic commercial application (Blake et al. 2008).

Fullerenes hold possibilities of application in many areas including antiviral activity, enzyme inhibition, DNA cleavage, photodynamic therapy, electron transfer, ball bearings, lightweight batteries, new lubricants, nanoscale electrical switches, new plastics, antitumor therapy for cancer patients, and combustion science and astrophysics (Dresselhaus et al. 1996; Lebedeva et al. 2015). The fullerene deriva-

tives obtained by attachment of electron donor moieties are used as photovoltaic devices. The supramolecular design of molecular assemblies involving fullerenes holds the possibility to reach new efficient photovoltaic devices (Hudhomme and Cousseau 2007). Fullerenes show promising biomedical applications (Bakry et al. 2007; Bosi et al. 2003; Mashino et al. 2003; Stoilova et al. 2007; Thrash et al. 1999). The fullerene derivatives showed antibacterial and antiproliferative activities; they inhibited bacteria and cancer cell growth effectively (Mashino et al. 2003). Cationic fullerenes were identified to work as antimicrobial photosensitizers. Bis-functionalized C_{60} derivatives have shown the activity against HIV-1 and HIV-2 strains (Bosi et al. 2003). The antiviral activity of fullerene derivatives is based on several biological properties including their unique molecular architecture and antioxidant activity (Bakry et al. 2007). Fullerenes derivatized by hydrophilic moieties are capable of carrying drugs and genes for the cellular delivery (Thrash et al. 1999). The localization of the metallofullerol in bone might be a useful chemotherapeutic agent for treatment of leukemia and bone cancer (Thrash et al. 1999).

Several potential applications have been proposed for carbon nanotubes, for example, conductive and high-strength composites, energy storage and energy conversion devices, sensors, field emission display and radiation sources, and nanotube-based semiconductor devices (Baughman et al. 2002; Sinha and Yeow 2005). Supercapacitors with carbon nanotube electrodes can be used for devices that require higher power capabilities than batteries. Nanotubes have potential application as hydrogen storage (Dinadayalane and Leszczynski 2009). CNTs can be added to aircraft to offer EMI (electromagnetic interference) shielding and lightning strike protection. They will also make the aircraft stronger and lighter, allowing for larger payloads and greater fuel efficiency. They may be used in commercial aircraft and in notebook computers to efficiently draw away generated heat without adding additional weight (Sinha and Yeow 2005). Nanotube films may be used by the automobile industry to make cars and trucks stronger yet lighter and, therefore, more fuel efficient. Three-dimensional graphene-CNT hollow fibers with radially aligned CNTs could be useful for efficient energy conversion and storage (Xue et al. 2015; Yu et al. 2014; Zhu et al. 2012).

MWCNTs show great potential for use in nanofluidic devices because of their high mechanical strength and fluid transport ability at near-molecular length scales (Sinha and Yeow 2005). Due to the advantages of miniature size of the nanotube and the small amount of material required, the carbon nanotubes are being explored for chemical sensing applications. SWCNTs are promising materials for building high-performance nano-sensors and devices (Close et al. 2008). Defects in SWCNTs play an important role in chemical sensing applications (Robinson et al. 2006). CNTs can be used as implanted sensors to monitor pulse, temperature, blood glucose, and heart's activity level and can also be used for repairing damaged cells or killing them by targeting tumors by chemical reactions (Sinha and Yeow 2005). Some of the selected applications of carbon nanotubes are shown in Fig. 53.

Potential biological and biomedical applications of CNTs are under investigation (Dhar et al. 2008; Karousis et al. 2009; Liu et al. 2007; Sinha and Yeow 2005).

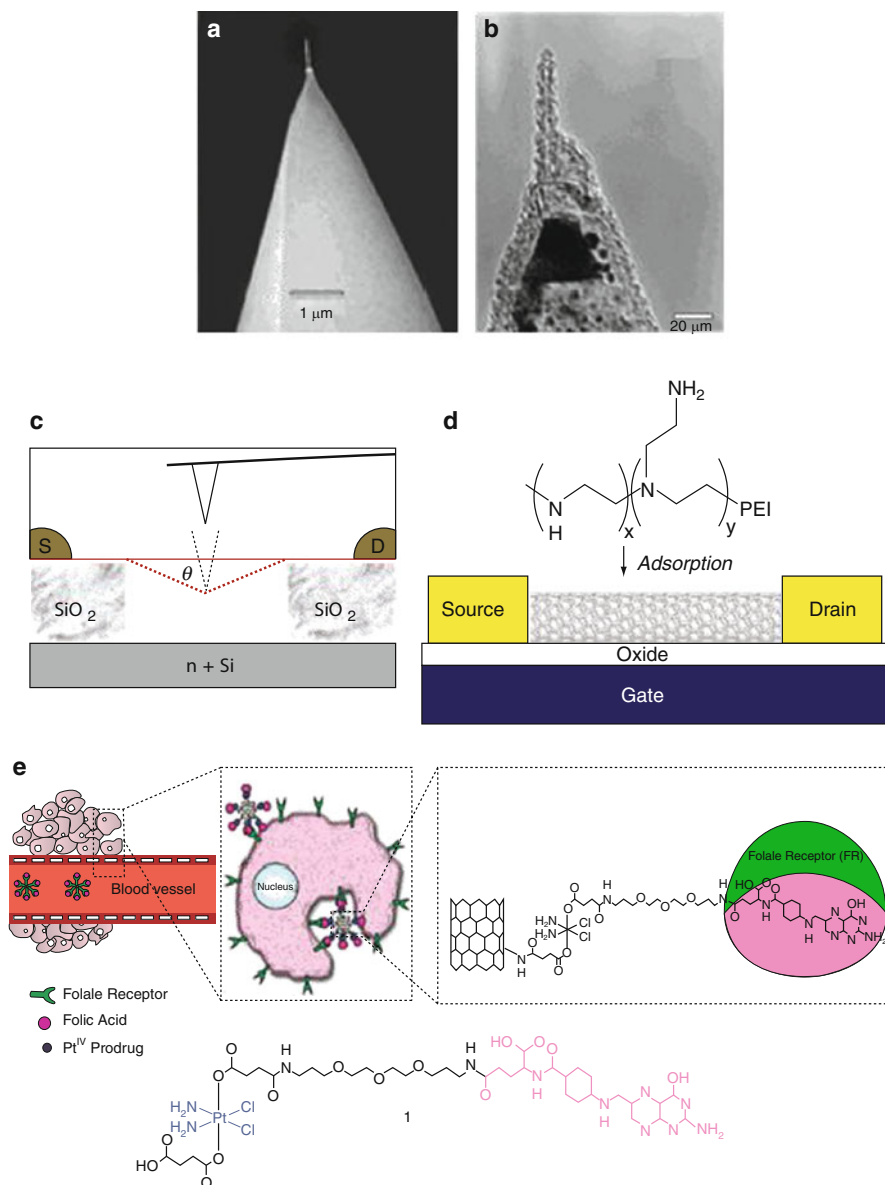


Fig. 53 (a) and (b) Electron micrographs of two different AFM cantilever tips, each with a nanotube attached; (a) an SEM (scanning electron microscope) micrograph of a nanotube. (b) A TEM (transmission electron microscope) micrograph of a nanotube (Reprinted with permission from reference Stevens et al. (2000). Copyright 2000 Institute of Physics). (c) Nanoelectromechanics of suspended nanotubes – experimental scheme for measuring the electromechanical property of the nanotube. (d) The SWCNT evolves into an n-type FET (field-effect transistor) after adsorption of PEI (polyethyleneimine). Pictures (c) and (d) were reprinted with permission from reference Dai (2002). Copyright 2002 American Chemical Society). (e) Folate receptor (FR)-mediated targeting and SWCNT-mediated delivery of Pt-containing complex (Reprinted with permission from reference Dhar et al. (2008). Copyright 2008 American Chemical Society)

Carbon nanotubes have potential to make miniature biological electronic devices, including probes and sensors (Sotiropoulou and Chaniotakis 2003; Stevens et al. 2000, 2004). Water-soluble peptidomimetic-functionalized carbon nanotubes have been reported to have antitrypsin activity (Dhar et al. 2008). Functionalized and water-soluble SWCNTs have been explored to find biological applications in the area of drug delivery (Dhar et al. 2008; Karousis et al. 2009; Liu et al. 2007). CNTs could be used as potential delivery tools for peptide-based synthetic vaccines. CNTs are currently being considered as suitable substrate for neuronal growth, as ion channel blockers, and as vectors for gene transfection (Sinha and Yeow 2005). Carbon nanotubes have provided possibilities for applications in nanotechnology. Continuous and optimistic research efforts in the area of carbon nanotubes are required to realize a lot of breakthrough commercial applications.

Summary and Outlook

In this chapter, we provided vital information and up-to-date research on carbon nanostructures, particularly graphene, fullerenes, and carbon nanotubes, which are critical in the nanotechnology revolution. This chapter also covered the modeling aspects, especially the current trends of computational chemistry applications in understanding the structures, reactivity, and other properties of abovementioned carbon nanostructures, and their importance in supporting the experimental results. Many aspects of basic research and practical application requirements have been motivating both theoreticians and experimentalists to gain better understanding about the carbon nanostructures. Obtaining knowledge on a specific class of chemical reactions with graphene, fullerenes, and SWCNTs is required for making novel materials as well as producing carbon-based nanomaterials for specific applications. Computational investigations provide opportunity to understand the structures, binding of atoms/molecules with the carbon nanotubes. A systematic and careful computational chemistry approaches could have important implications for the rational design of novel CNT composite materials, novel nanotube-based sensors, as well as for the development of new chemical strategies for SWCNT functionalization. Strong interactions between experimentalists and theoreticians working in the area of carbon nanostructures will enhance the real-time applications rapidly. Future efforts should not only provide high-tech nano-devices but also address fundamental scientific questions. Further exciting developments in nanoscience and nanotechnology are expected.

Acknowledgments This work was supported by the High Performance Computational Design of Novel Materials (HPCDNM) Project funded by the Department of Defense (DoD) through the US Army Engineer Research and Development Center (Vicksburg, MS) Contract # W912HZ-06-C-0057 and by the Office of Naval Research (ONR) grant 08PRO2615-00/N00014-08-1-0324. JL acknowledges the support from the National Science Foundation (NSF) for the Interdisciplinary Center for Nanotoxicity (ICN) through CREST grant HRD-0833178. TCD acknowledges the start up support provided by the Clark Atlanta University.

Bibliography

- Abanin, D. A., Lee, P. A., & Levitov, L. S. (2006). Spin-filtered edge states and quantum hall effect in graphene. *Physical Review Letters*, *96*, 176803-1–176803-4.
- Achiba, Y., Kikuchi, K., Aihara, Y., Wakabayashi, Y., Miyake, Y., & Kainosho, M. (1995). In P. Bernier, D. S. Bethune, L. Y. Chiang, T. W. Ebbesen, R. M. Metzger, & J. W. Mintmire, J. W. (Eds.), *Higher fullerenes: Structure and properties* (Materials Research Society symposium proceedings, Vol. 359, p. 3). Pittsburgh, PA: Materials Research Society.
- Achiba, Y., Kikuchi, K., Aihara, Y., Wakabayashi, T., Miyake, Y., & Kainosho, M. (1996). Fullerenes and Endofullerenes: Model Substances?. In W. Andreoni (Ed.), *The chemical physics of fullerenes 10 (and 5) years later: the far-reaching impact of the discovery of C₆₀* (p. 139). Dordrecht: Kluwer.
- Akdim, B., Kar, T., Duan, X., & Pachter, R. (2007). Density functional theory calculations of ozone adsorption on sidewall single-wall carbon nanotubes with Stone-Wales defects. *Chemical Physics Letters*, *445*, 281–287.
- Amorim, R. G., Fazzio, A., Antonelli, A., Novaes, F. D., & da Silva, A. J. R. (2007). Divacancies in graphene and carbon nanotubes. *Nano Letters*, *7*, 2459–2462.
- Amsharov, K. Y., & Jensen, M. (2008). A C₇₈ fullerene precursor: Toward the direct synthesis of higher fullerenes. *Journal of Organic Chemistry*, *73*, 2931–2934.
- An, W., Gao, Y., Bulusu, S., & Zeng, X. C. (2005). Ab initio calculation of bowl, cage, and ring isomers of C₂₀ and C₂₀⁻. *Journal of Chemical Physics*, *122*, 204109-1–204109-8.
- Andzelm, J., Govind, N., & Maiti, A. (2006). Nanotube-based gas sensors – Role of structural defects. *Chemical Physics Letters*, *421*, 58–62.
- Arnold, M. S., Green, A. A., Hulvat, J. F., Stupp, S. I., & Hersam, M. C. (2006). Sorting carbon nanotubes by electronic structure using density differentiation. *Nature Nanotechnology*, *1*, 60–65.
- Austin, S. J., Fowler, P. W., Manolopoulos, D. E., & Zerbetto, F. (1995). The Stone-Wales map for C₆₀. *Chemical Physics Letters*, *235*, 146–151.
- Avila, A. F., & Lacerda, G. S. R. (2008). Molecular mechanics applied to single-walled carbon nanotubes. *Materials Research*, *11*, 325–333.
- Avouris, P., Chen, Z. H., & Perebeinos, V. (2007). Carbon-based electronics. *Nature Nanotechnology*, *2*, 605–615.
- Bakry, R., Vallant, R. M., Najam-ul-Haq, M., Rainer, M., Szabo, Z., Huck, C. W., & Bonn, G. K. (2007). Medicinal applications of fullerenes. *International Journal of Nanomedicine*, *2*, 639–649.
- Balandin, A. A., Ghosh, S., Bao, W., Calizo, I., Teweldebrhan, D., Miao, F., & Lau, C. N. (2008). Superior thermal conductivity of single-layer graphene. *Nano Letters*, *8*, 902–907.
- Barth, W. E., & Lawton, R. G. (1966). Dibenz[ghi, mno]fluoranthene. *Journal of the American Chemical Society*, *88*, 380–381.
- Baughman, R. H., Zakhidov, A. A., & de Heer, W. A. (2002). Carbon nanotubes-the route toward applications. *Science*, *297*, 787–792.
- Beavers, C. M., Zuo, T., Duchamp, J. C., Harich, K., Dorn, H. C., Olmstead, M. M., & Balch, A. L. (2006). Tb₃N@C₈₄: An improbable, egg-shaped endohedral fullerene that violates the isolated pentagon rule. *Journal of the American Chemical Society*, *128*, 11352–11353.
- Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *Journal of Chemistry Physics*, *98*, 5648–5652.
- Becker, L., Bada, J. L., Winans, R. E., Hunt, J. E., Bunch, T. E., & French, B. M. (1994). Fullerenes in the 1.85-billion-year-old Sudbury impact structure. *Science*, *265*, 642–645.
- Berthe, M., Yoshida, S., Ebine, Y., Kanazawa, K., Okada, A., Taninaka, A., Takeuchi, O., Fukui, N., Shinohara, H., Suzuki, S., Sumitomo, K., Kobayashi, Y., Grandidier, B., Stievenard, D., & Shigekawa, H. (2007). Reversible defect engineering of single-walled carbon nanotubes using scanning tunneling microscopy. *Nano Letters*, *7*, 3623–3627.

- Bethune, D. S., Kiang, C.-H., de Vries, M. S., Gorman, G., Savoy, R., Vazquez, J., & Beyers, R. (1993). Cobalt-catalysed growth of carbon nanotubes with single-atomic-layer walls. *Nature*, *363*, 605–607.
- Bettinger, H. F. (2004). Effects of finite carbon nanotube length on sidewall addition of fluorine atom and methylene. *Organic Letters*, *6*, 731–734.
- Bettinger, H. F. (2005). The reactivity of defects at the sidewalls of single-walled carbon nanotubes: The Stone–Wales defect. *Journal of Physical Chemistry B*, *109*, 6922–6924.
- Bettinger, H. F. (2006). Addition of carbenes to the sidewalls of single-walled carbon nanotubes. *Chemistry – A European Journal*, *12*, 4372–4379.
- Bettinger, H. F., Yakobson, B. I., & Scuseria, G. E. (2003). Scratching the surface of buckminsterfullerene: The barriers for Stone–Wales transformation through symmetric and asymmetric transition states. *Journal of the American Chemical Society*, *125*, 5572–5580.
- Blake, P., Brimicombe, P. D., Nair, R. R., Booth, T. J., Jiang, D., Schedin, F., Ponomarenko, L. A., Morozov, S. V., Gleeson, H. F., Hill, E. W., Geim, A. K., & Novoselov, K. S. (2008). Graphene-based liquid crystal device. *Nano Letters*, *8*, 1704–1708.
- Bo, Z., Mao, S., Han, Z. J., Cen, K., Chen, J., & Ostrikov, K. (2015). Emerging energy and environmental applications of vertically-oriented graphenes. *Chemical Society Reviews*, *44*, 2108–2121.
- Bonaccorso, F., Colombo, L., Yu, G., Stoller, M., Tozzini, V., Ferrari, A. C., Ruoff, R. S., & Vittorio Pellegrini, V. (2015). Graphene, related two-dimensional crystals, and hybrid systems for energy conversion and storage. *Science*, *347*, 1246501.
- Bosi, S., Ros, T. D., Spalluto, G., Balzarini, J., & Prato, M. (2003). Synthesis and anti-HIV properties of new water-soluble bis-functionalized[60]fullerene derivatives. *Bioorganic and Medicinal Chemistry Letters*, *13*, 4437–4440.
- Boukhvalov, D. W., & Katsnelson, M. I. (2008). Chemical functionalization of graphene with defects. *Nano Letters*, *8*, 4373–4379.
- Boukhvalov, D. W., Katsnelson, M. I., & Lichtenstein, A. I. (2008). Hydrogen on graphene: Electronic structure, total energy, structural distortions and magnetism from first-principles calculations. *Physical Review B*, *77*, 035427-1–035427-7.
- Bunch, J. S., van der Zande, A. M., Verbridge, S. S., Frank, I. W., Tanenbaum, D. M., Parpia, J. M., Craighead, H. G., & McEuen, P. L. (2007). Electromechanical resonators from graphene sheets. *Science*, *315*, 490–493.
- Burda, C., Samia, A. C. S., Hathcock, D. J., Huang, H., & Yang, S. (2002). Experimental evidence for the photoisomerization of higher fullerenes. *Journal of the American Chemical Society*, *124*, 12400–12401.
- Buseck, P. R., Tshipursky, S. J., & Hettich, R. (1992). Fullerenes from the geological environment. *Science*, *257*, 215–217.
- Cabrera-Sanfeliu, P., & Darling, G. R. (2007). Dissociative adsorption of water at vacancy defects in graphite. *The Journal of Physical Chemistry C*, *111*, 18258–18263.
- Calaminici, P., Geudtner, G., & Koster, A. M. (2009). First-principle calculations of large fullerenes. *Journal of Chemical Theory and Computation*, *5*, 29–32.
- Carlson, J. M., & Scheffler, M. (2006). Structural, electronic, and chemical properties of nanoporous carbon. *Physical Review Letters*, *96*, 046806-1–046806-4.
- Carpio, A., Bonilla, L. L., de Juan, F., & Vozmediano, M. A. H. (2008). Dislocations in graphene. *New Journal of Physics*, *10*, 053021-1–053021-13.
- Chakraborty, A. K., Woolley, R. A. J., Butenko, Y. V., Dhanak, V. R., Šiller, L., & Hunt, M. R. C. (2007). A photoelectron spectroscopy study of ion-irradiation induced defects in single-wall carbon nanotubes. *Carbon*, *45*, 2744–2750.
- Chandra, N., Namila, S., & Shet, C. (2004). Local elastic properties of carbon nanotubes in the presence of Stone–Wales defects. *Physical Review B*, *69*, 094101-1–094101-12.
- Charlier, J.-C. (2002). Defects in carbon nanotubes. *Accounts of Chemical Research*, *35*, 1063–1069.
- Charlier, J.-C., Ebbsen, T. W., & Lambin, P. (1996). Structural and electronic properties of pentagon–heptagon pair defects in carbon nanotubes. *Physical Review B*, *53*, 11108–11113.

- Chaur, M. N., Valencia, R., Rodríguez-Forteza, A., Poblet, J. M., & Echegoyen, L. (2009a). Trimetallic nitride endohedral fullerenes: Experimental and theoretical evidence for the $M_3N^{6+}@C_{2n}^{6-}$ model. *Angewandte Chemie, International Edition*, 48, 1425–1428.
- Chaur, M. N., Melin, F., Ortiz, A. L., & Echegoyen, L. (2009b). Chemical, electrochemical, and structural properties of endohedral metallofullerenes. *Angewandte Chemie, International Edition*, 48, 7514–7538.
- Chen, Z. (2004). The smaller fullerene C_{50} , isolated as $C_{50}Cl_{10}$. *Angewandte Chemie, International Edition*, 43, 4690–4691.
- Chen, Z., Thiel, W., & Hirsch, A. (2003). Reactivity of the convex and concave surfaces of single-walled carbon nanotubes (SWCNTs) towards addition reactions: Dependence on the carbon-atom pyramidalization. *Chemical Physics and Physical Chemistry*, 4, 93–97.
- Chen, K., Song, S., Liu, F., & Xue, D. (2015). Structural design of graphene for use in electrochemical energy storage devices. *Chemical Society Reviews*, 44, 6230–6257.
- Chico, L., Crespi, V. H., Benedict, L. X., Louie, S. G., & Cohen, M. L. (1996). Pure carbon nanoscale devices: Nanotube heterojunctions. *Physical Review Letters*, 76, 971–974.
- Cho, E., Shin, S., & Yoon, Y.-G. (2008). First-principles studies on carbon nanotubes functionalized with azomethine ylides. *The Journal of Physical Chemistry C*, 112, 11667–11672.
- Christian, J. F., Wan, Z., & Anderson, S. L. (1992). $O^+ + C_{60} \rightarrow C_{60}O^+$ production and decomposition, charge transfer, and formation of $C_{59}O^+$. Dopeyball or $[CO@C_{58}^+]$. *Chemical Physics Letters*, 199, 373–378.
- Chun, H., Hahm, M. G., Homma, Y., Meritz, R., Kuramochi, K., Menon, L., Ci, L., Ajayan, P. M., & Jung, Y. J. (2009). Engineering low-aspect ratio carbon nanostructures: Nanocups, nanorings, and nanocontainers. *ACS Nano*, 3, 1274–1278.
- Chuvilin, A., Kaiser, U., Bichoutskaia, E., Besley, N. A., & Khlobystov, A. N. (2010). Direct transformation of graphene to fullerene. *Nature Chemistry*, 2, 450–453.
- Cioslowski, J., Rao, N., & Moncrieff, D. (2002). Electronic structures and energetics of [5,5] and [9,0] single-walled carbon nanotubes. *Journal of the American Chemical Society*, 124, 8485–8489.
- Close, G. F., Yasuda, S., Paul, B., Fujita, S., & Wong, H. S. P. (2008). A 1 GHz integrated circuit with carbon nanotube interconnects and silicon transistors. *Nano Letters*, 8, 706–709.
- Cohen, M. L. (1993). Predicting useful materials. *Science*, 261, 307–308.
- Crassous, J., Rivera, J., Fender, N. S., Shu, L., Echegoyen, L., Thilgen, C., Herrmann, A., & Diederich, F. (1999). Chemistry of C_{84} : Separation of three constitutional isomers and optical resolution of D_2-C_{84} by using the “Bingel-Retro-Bingel” strategy. *Angewandte Chemie, International Edition*, 38, 1613–1617.
- Cuesta, I. G., Pedersen, T. B., Koch, H., & de Meras, A. S. (2006). Carbon nanorings: A challenge to theoretical chemistry. *Chemical Physics and Physical Chemistry*, 7, 2503–2507.
- Cyranski, M. K., Howard, S. T., & Chodkiewicz, M. L. (2004). Bond energy, aromatic stabilization energy and strain in IPR fullerenes. *Chemical Communications*, 2458–2459
- Dai, H. (2002). Carbon nanotubes: Synthesis, integration, and properties. *Accounts of Chemical Research*, 35, 1035–1044.
- David, W. I. F., Ibberson, R. M., Matthewman, J. C., Prassides, K., Dennis, T. J. S., Hare, J. P., Kroto, H. W., Taylor, R., & Walton, D. R. M. (1991). Crystal structure and bonding of ordered C_{60} . *Nature*, 353, 147–149.
- David, V. P., Lin, X., Zhang, H., Liu, S., & Kappes, M. M. (1992). Transmission electron microscopy of C_{70} single crystals at room temperature. *Journal of Materials Research*, 7, 2440–2446.
- Deng, J.-P., Ju, D.-D., Her, G.-R., Mou, C.-Y., Chen, C.-J., Lin, Y.-Y., & Han, C.-C. (1993). Odd-numbered fullerene fragment ions from C_{60} oxides. *Journal of Physical Chemistry*, 97, 11575–11577.
- Denis, P. A., Iribarne, F., & Faccio, R. (2009). Hydrogenated double wall carbon nanotubes. *Journal of Chemical Physics*, 130, 194704-1–194704-10.
- Dennis, T. J. S., & Shinohara, H. (1998). Isolation and characterisation of the two major isomers of [84]fullerene (C_{84}). *Chemical Communications*, 619–620.

- Dereli, G., & Sungu, B. (2007). Temperature dependence of the tensile properties of single-walled carbon nanotubes: O(N) tight-binding molecular-dynamics simulations. *Physical Review B*, *75*, 184104-1–184104-6.
- Dewar, M. J. S., & Thiel, W. (1977). Ground states of molecules. 38. The MNDO method. Approximations and parameters. *Journal of the American Chemical Society*, *99*, 4899–4907.
- Dewar, M. J. S., Zoebisch, E. G., Healy, E. F., & Stewart, J. J. P. (1985). Development and use of quantum mechanical molecular models. 76. AM1: A new general purpose quantum mechanical molecular model. *Journal of the American Chemical Society*, *107*, 3902–3909.
- Dewar, M. J. S., Jie, C., & Yu, J. (1993). SAM1; The first of a new series of general purpose quantum mechanical molecular models. *Tetrahedron*, *49*, 5003–5038.
- Dhar, S., Liu, Z., Thomale, J., Dai, H., & Lippard, S. J. (2008). Targeted single-wall carbon nanotube-mediated Pt(IV) prodrug delivery using folate as a homing device. *Journal of the American Chemical Society*, *130*, 11467–11476.
- Diederich, F., Ettl, R., Rubin, Y., Whetten, R. L., Beck, R., Alvarez, M., Anz, S., Sensharma, D., Wudl, F., Khemani, K. C., & Koch, A. (1991a). The higher fullerenes: Isolation and characterization of C₇₆, C₈₄, C₉₀, C₉₄, and C₇₀O, an oxide of D_{5h}-C₇₀. *Science*, *252*, 548–551.
- Diederich, F., Whetten, R. L., Thilgen, C., Ettl, R., Chao, I., & Alvarez, M. M. (1991b). Fullerene isomerism: Isolation of C_{2v}-C₇₈ and D₃-C₇₈. *Science*, *254*, 1768–1770.
- Dillon, A. C., Jones, K. M., Bekkedahl, T. A., Kiang, C. H., Bethune, D. S., & Heben, M. J. (1997). Storage of hydrogen in single-walled carbon nanotubes. *Nature*, *386*, 377–379.
- Dinadayalane, T. C., & Leszczynski, J. (2007a). Toward nanomaterials: Structural, energetic and reactivity aspects of single-walled carbon nanotubes. In P. B. Balbuena & J. M. Seminario (Eds.), *Nanomaterials: Design and simulation* (Theoretical and computational chemistry, Vol. 18, pp. 167–199). Amsterdam: Elsevier.
- Dinadayalane, T. C., & Leszczynski, J. (2007b). StoneWales defects with two different orientations in (5, 5) single-walled carbon nanotubes: A theoretical study. *Chemical Physics Letters*, *434*, 86–91.
- Dinadayalane, T. C., & Leszczynski, J. (2009). Toward understanding of hydrogen storage in single-walled carbon nanotubes by chemisorption mechanism. In J. Leszczynski & M. K. Shukla (Eds.), *Practical aspects of computational chemistry: Methods, concepts and applications* (pp. 297–313). Dordrecht: Springer.
- Dinadayalane, T. C., & Sastry, G. N. (2001). Synthetic strategies toward buckybowls and C₆₀: Benzannulation is remarkably facile compared to cyclopentannulation. *Tetrahedron Letters*, *42*, 6421–6423.
- Dinadayalane, T. C., & Sastry, G. N. (2002a). Structure-energy relationships in curved polycyclic aromatic hydrocarbons: Study of benzocorannulenes. *Journal of Organic Chemistry*, *67*, 4605–4607.
- Dinadayalane, T. C., & Sastry, G. N. (2002b). An assessment of semiempirical (MNDO, AM1 and PM3) methods to model buckybowls. *Journal of Molecular Structure (THEOCHEM)*, *579*, 63–72.
- Dinadayalane, T. C., & Sastry, G. N. (2003). Isolated pentagon rule in buckybowls: A computational study on thermodynamic stabilities and bowl-to-bowl inversion barriers. *Tetrahedron*, *59*, 8347–8351.
- Dinadayalane, T. C., Priyakumar, U. D., & Sastry, G. N. (2001). Theoretical studies on the effect of sequential benzannulation to corannulene. *Journal of Molecular Structure (THEOCHEM)*, *543*, 1–10.
- Dinadayalane, T. C., Priyakumar, U. D., & Sastry, G. N. (2002). Ring closure synthetic strategies toward buckybowls: Benzannulation versus cyclopentannulation. *Journal of the Chemical Society, Perkin Transactions*, *2*, 94–101.
- Dinadayalane, T. C., Deepa, S., & Sastry, G. N. (2003). Is peri hydrogen repulsion responsible for flattening buckybowls? The effect of ring annelation to the rim of corannulene. *Tetrahedron Letters*, *44*, 4527–4529.

- Dinadayalane, T. C., Deepa, S., Reddy, A. S., & Sastry, G. N. (2004). Density functional theory study on the effect of substitution and ring annelation to the rim of corannulene. *Journal of Organic Chemistry*, 69, 8111–8114.
- Dinadayalane, T. C., Gorb, L., Simeon, T., & Dodziuk, H. (2007a). Cumulative p-p interaction triggers unusually high stabilization of linear hydrocarbons inside the single-walled carbon nanotube. *International Journal of Quantum Chemistry*, 107, 2204–2210.
- Dinadayalane, T. C., Kaczmarek, A., Łukaszewicz, J., & Leszczynski, J. (2007b). Chemisorption of hydrogen atoms on the sidewalls of armchair single-walled carbon nanotubes. *Journal of Physical Chemistry C*, 111, 7376–7383.
- Ding, F. (2005). Theoretical study of the stability of defects in single-walled carbon nanotubes as a function of their distance from the nanotube end. *Physical Review B*, 72, 245409-1–245409-7.
- Dresselhaus, M. S., Dresselhaus, G., & Eklund, P. C. (1996). *Science of fullerenes and carbon nanotubes: Their properties and applications*. San Diego: Academic.
- Dresselhaus, M. S., Dresselhaus, G., & Avouris, P. (Eds.). (2001). *Carbon nanotubes: Synthesis, structure, properties, and applications*. Berlin: Springer.
- Dresselhaus, M. S., Dresselhaus, G., Jorio, A., Filho, A. G. S., Pimenta, M. A., & Saito, R. (2002). Single nanotube Raman spectroscopy. *Accounts of Chemical Research*, 35, 1070–1078.
- Dresselhaus, M. S., Dresselhaus, G., Saito, R., & Jorio, A. (2005). Raman spectroscopy of carbon nanotubes. *Physics Reports*, 409, 47–99.
- Dresselhaus, M. S., Dresselhaus, G., & Jorio, A. (2007). Raman spectroscopy of carbon nanotubes in 1997 and 2007. *The Journal of Physical Chemistry C*, 111, 17887–17893.
- Dulap, B. I., & Zope, R. R. (2006). Efficient quantum-chemical geometry optimization and the structure of large icosahedral fullerenes. *Chemical Physics Letters*, 422, 451–454.
- Dunlap, B. I., Brenner, D. W., Mintmire, J. W., Mowrey, R. C., & White, C. T. (1991). Local density functional electronic structures of three stable icosahedral fullerenes. *Journal of Physical Chemistry*, 95, 8737–8741.
- Duplock, E. J., Scheffler, M., & Lindan, P. J. D. (2004). Hallmark of perfect graphene. *Physical Review Letters*, 92, 225502-1–225502-4.
- EGgen, B. R., Heggie, M. I., Jungnickel, G., Latham, C. D., Jones, R., & Briddon, P. R. (1996). Autocatalysis during fullerene growth. *Science*, 272, 87–89.
- Ekinici, K. L., Huang, X. M. H., & Roukes, M. L. (2004). Ultrasensitive nanoelectromechanical mass detection. *Applied Physics Letters*, 84, 4469–4471.
- Elias, D. C., Nair, R. R., Mohiuddin, T. M. G., Morozov, S. V., Blake, P., Halsall, M. P., Ferrari, A. C., Boukhalov, D. W., Katsnelson, M. I., Geim, A. K., & Novoselov, K. S. (2009). Control of graphene's properties by reversible hydrogenation: Evidence for graphene. *Science*, 323, 610–613.
- Ertekin, E., Chrzan, D. C., & Daw, M. S. (2009). Topological description of the Stone-Wales defect formation energy in carbon nanotubes and graphene. *Physical Review B*, 79, 155421-1–155421-17.
- Esquivel, E. V., & Murr, L. E. (2004). A TEM analysis of nanoparticulates in a polar ice core. *Materials Characterization*, 52, 15–25.
- Feng, X., Irlle, S., Witek, H., Morokuma, K., Vidic, R., & Borguet, E. (2005). Sensitivity of ammonia interaction with single-walled carbon nanotube bundles to the presence of defect sites and functionalities. *Journal of the American Chemical Society*, 127, 10533–10538.
- Fischer, J. E., Heiney, P. A., McGhie, A. R., Romanow, W. J., Denenstien, A. M., McCauley, J. P., Jr., & Smith, A. B., III. (1991). Compressibility of solid C₆₀. *Science*, 252, 1288–1290.
- Fowler, P. W., & Heine, T. (2001). Stabilisation of pentagon adjacencies in the lower fullerenes by functionalisation. *Journal of the Chemical Society, Perkin Transactions*, 2, 487–490.
- Fowler, P. W., & Manolopoulos, D. E. (1995). *An atlas of fullerenes*. New York: Oxford University Press.
- Frisch, M. J. et al. (2003). *Gaussian 03, revision E.1*. Pittsburg, PA: Gaussian.
- Fu, W., Xu, L., Azurmendi, H., Ge, J., Fuhrer, T., Zuo, T., Reid, J., Shu, C., Harich, K., & Dorn, H. C. (2009). ⁸⁹Y and ¹³C NMR cluster and carbon cage studies of an yttrium

- metallofullerene family, $Y_3N@C_{2n}$ ($n = 40 - 43$). *Journal of the American Chemical Society*, *131*, 11762–11769.
- Galano, A. (2006). On the influence of diameter and length on the properties of armchair single-walled carbon nanotubes: A theoretical chemistry approach. *Chemical Physics*, *327*, 159–170.
- Geim, A. K. (2009). Graphene: Status and prospects. *Science*, *324*, 1530–1534.
- Geim, A. K., & Novoselov, K. S. (2007). The rise of graphene. *Nature Materials*, *6*, 183–191.
- Govind, N., Andzelm, J., & Maiti, A. (2008). Dissociation chemistry of gas molecules on carbon nanotubes applications to chemical sensing. *IEEE Sensors Journal*, *8*, 837–841.
- Gu, Z., Peng, H., Hauge, R. H., Smalley, R. E., & Margrave, J. L. (2002). Cutting single-wall carbon nanotubes through fluorination. *Nano Letters*, *2*, 1009–1013.
- Gueorguiev, G. K., Pacheco, J. M., & Tomanek, D. (2004). Quantum size effects in the polarizability of carbon fullerenes. *Physical Review Letters*, *92*, 215501-1–215501-4.
- Guo, T., Diener, M. D., Chai, Y., Alford, M. J., Hauffer, R. E., McClure, S. M., Ohno, T., Weaver, J. H., Scuseria, G. E., & Smalley, R. E. (1992). Uranium stabilization of C_{28} : A tetravalent fullerene. *Science*, *257*, 1661–1663.
- Haddon, R. C. (1993). Chemistry of the fullerenes: The manifestation of strain in a class of continuous aromatic molecules. *Science*, *261*, 1545–1550.
- Haddon, R. C., & Scott, L. T. (1986). π -Orbital conjugation and rehybridization in bridged annulenes and deformed molecules in general: π -orbital axis vector analysis. *Pure and Applied Chemistry*, *58*, 137–142.
- Hamada, N., Sawada, S., & Oshiyama, A. (1992). New one-dimensional conductors: Graphitic microtubules. *Physical Review Letters*, *68*, 1579–1581.
- Harutyunyan, A. R., Chen, G., Paronyan, T. M., Pigos, E. M., Kuznetsov, O. A., Hewaparakrama, K., Kim, S. M., Zakharov, D., Stach, E. A., & Sumanasekera, G. U. (2009). Preferential growth of single-walled carbon nanotubes with metallic conductivity. *Science*, *326*, 116–120.
- He, H. Y., & Pan, B. C. (2009). Electronic structures and Raman features of a carbon nanobud. *The Journal of Physical Chemistry C*, *113*, 20822–20826.
- Heath, J. R. (1991). Synthesis of C_{60} from small carbon clusters: a model based on experiment and theory. *ACS Symposium Series*, *481*, 1–23.
- Helden, G. v., Gotts, N. G., & Bowers, M. T. (1993). Experimental evidence for the formation of fullerenes by collisional heating of carbon rings in the gas phase. *Nature*, *363*, 60–63.
- Hernández, E., Ordejón, P., & Terrones, H. (2001). Fullerene growth and the role of nonclassical isomers. *Physical Review B*, *63*, 193403-1–193403-4.
- Heymann, D., Chibante, L. P. F., Brooks, R. R., Wolbach, W. S., & Smalley, R. E. (1994). Fullerenes in the cretaceous-tertiary boundary layer. *Science*, *265*, 645–647.
- Heymann, D., Jenneskens, L. W., Jehlička, J., Koper, C., & Vlietstra, E. (2003). Terrestrial and extraterrestrial fullerenes. *Fullerenes, Nanotubes, and Carbon Nanostructures*, *11*, 333–370.
- Hirahara, K., Suenaga, K., Bandow, S., Kato, H., Okazaki, T., Shinohara, H., & Iijima, S. (2000). One-dimensional metallofullerene crystal generated inside single-walled carbon nanotubes. *Physical Review Letters*, *85*, 5384–5387.
- Hirsch, A. (2002). Functionalization of single-walled carbon nanotubes. *Angewandte Chemie, International Edition*, *41*, 1853–1859.
- Howard, J. B., Mckinnon, J. T., Makarovskiy, Y., Lafleur, A. L., & Johnson, M. E. (1991). Fullerenes C_{60} and C_{70} in flames. *Nature*, *352*, 139–141.
- <http://www.nndc.bnl.gov/content/elements.html>
- Hu, Y. H., & Ruckenstein, E. (2003). Ab initio quantum chemical calculations for fullerene cages with large holes. *Journal of Chemical Physics*, *119*, 10073–10080.
- Hu, Y. H., & Ruckenstein, E. (2004). Quantum chemical density-functional theory calculations of the structures of defect C_{60} with four vacancies. *Journal of Chemical Physics*, *120*, 7971–7975.
- Hudhomme, P., & Cousseau, J. (2007). Plastic solar cells using fullerene derivatives in the photoactive layer. In F. Langa & J.-F. Nierengarten (Eds.), *Fullerenes: Principles and applications*. London: Royal Society of Chemistry.
- Hutter, J., et al. Computer code CPMD, version 3.11 (The CPMD program is © 2000–2016 jointly by IBM Corp. and by Max Planck Institute, Stuttgart.), <http://www.cpmd.org/>.

- Iijima, S. (1991). Helical microtubules of graphitic carbon. *Nature*, 354, 56–58.
- Iijima, S. (2007). A career in carbon. *Nature Nanotechnology*, 2, 590–591.
- Iijima, S., & Ichihashi, T. (1993). Single-shell carbon nanotubes of 1-nm diameter. *Nature*, 363, 603–605.
- Iijima, S., Yudasaka, M., Yamada, R., Bandow, S., Suenaga, K., Kokai, F., & Takahashi, K. (1999). Nano-aggregates of single-walled graphitic carbon nano-horns. *Chemical Physics Letters*, 309, 165–170.
- Ioffe, I. N., Goryunkov, A. A., Tamm, N. B., Sidorov, L. N., Kemnitz, E., & Troyanov, S. I. (2009). Fusing pentagons in a fullerene cage by chlorination: IPR D_2-C_{76} Rearranges into non-IPR $C_{76}Cl_{24}$. *Angewandte Chemie, International Edition*, 48, 5904–5907.
- Jia, T.-T., Zheng, M.-M., Fan, X.-Y., Su, Y., Li, S.-J., Liu, H.-Y., Chen, G., & Kawazoe, Y. (2016). Dirac cone move and bandgap on/off switching of graphene superlattice. *Scientific Reports*, 6, 18869.
- Jiang, H., Nasibulin, A. G., Brown, D. P., & Kauppinen, E. I. (2007). Unambiguous atomic structural determination of single-walled carbon nanotubes by electron diffraction. *Carbon*, 45, 662–667.
- Jiang, D., Cooper, V. R., & Dai, S. (2009). Porous graphene as the ultimate membrane for gas separation. *Nano Letters*, 9, 4019–4024.
- Kaczmarek, A., Dinadayalane, T. C., Łukaszewicz, J., & Leszczynski, J. (2007). Effect of tube length on the chemisorptions of one and two hydrogen atoms on the sidewalls of (3,3) and (4,4) single-walled carbon nanotubes: A theoretical study. *International Journal of Quantum Chemistry*, 107, 2211–2219.
- Kadish, K. M., & Ruoff, R. S. (Eds.). (2002). *Fullerene: Chemistry physics and technology*. New York: Wiley.
- Kar, T., Bettinger, H. F., Scheiner, S., & Roy, A. K. (2008). Noncovalent $\pi - \pi$ stacking and CH- π interactions of aromatics on the surface of single-wall carbon nanotubes: An MP2 study. *The Journal of Physical Chemistry C*, 112, 20070–20075.
- Karousis, N., Papi, R. M., Siskos, A., Vakalopoulou, P., Glezakos, P., Sarigiannis, Y., Stavropoulos, G., Kyriakidis, D. A., & Tagmatarchis, N. (2009). Peptidomimetic-functionalized carbon nanotubes with antitrypsin activity. *Carbon*, 47, 3550–3558.
- Kessler, B., Bringer, A., Cramm, S., Schlebusch, C., Eberhardt, W., Suzuki, S., Achiba, Y., Esch, F., Barnaba, M., & Cocco, D. (1997). Evidence for incomplete charge transfer and La-derived states in the valence bands of endohedrally doped $La@C_{82}$. *Physical Review Letters*, 79, 2289–2292.
- Kikuchi, K., Nakahara, N., Wakabayashi, T., Suzuki, S., Shiromaru, H., Miyake, Y., Saito, K., Ikemoto, I., Kainosho, M., & Achiba, Y. (1992a). NMR characterization of isomers of C_{78} , C_{82} and C_{84} fullerenes. *Nature*, 357, 142–145.
- Kikuchi, K., Nakahara, N., Wakabayashi, T., Honda, M., Matsumiya, H., Moriwaki, T., Suzuki, S., Shiromaru, H., Saito, K., Yamauchi, K., Ikemoto, I., & Achiba, Y. (1992b). Isolation and identification of fullerene family: C_{76} , C_{78} , C_{82} , C_{84} , C_{90} and C_{96} . *Chemical Physics Letters*, 188, 177–180.
- Kimura, T., Sugai, T., Shinohara, H., Goto, T., Tohji, K., & Matsuoka, I. (1995). Preferential arc-discharge production of higher fullerenes. *Chemical Physics Letters*, 246, 571–576.
- Klein, D. J., & Schmalz, T. G. (1990). In I. Hargittai (Ed.), *Quasicrystals, networks, and molecules of fivefold symmetry* (p. 239). New York: VCH.
- Knobel, R. G., & Cleland, A. N. (2003). Nanometre-scale displacement sensing using a single electron Transistor. *Nature*, 424, 291–293.
- Kostov, M. K., Santiso, E. E., George, A. M., Gubbins, K. E., & Nardelli, M. B. (2005). Dissociation of water on defective carbon substrates. *Physical Review Letters*, 95, 136105-1–136105-4.
- Krätschmer, W., Lamb, L. D., Fostiropoulos, K., & Huffman, D. R. (1990). Solid C_{60} : A new form of carbon. *Nature*, 347, 354–358.
- Kresse, G., & Furthmüller, J. (1996a). Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54, 11169–11186.

- Kresse, G., & Furthmuller, J. (1996b). Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, *6*, 15–50.
- Kroto, H. W. (1987). The stability of the fullerenes C_n , with $n = 24, 28, 32, 36, 50, 60$ and 70 . *Nature*, *329*, 529–531.
- Kroto, H. W., Heath, J. R., O'Brien, S. C., Curl, R. F., & Smalley, R. E. (1985). Buckminsterfullerene. *Nature*, *318*, 162–163.
- Kubozono, Y., Maeda, H., Takabayashi, Y., Hiraoka, K., Nakai, T., Kashino, S., Emura, S., Ukita, S., & Sogabe, T. (1996). Extractions of $Y@C_{60}$, $Ba@C_{60}$, $La@C_{60}$, $Ce@C_{60}$, $Pr@C_{60}$, $Nd@C_{60}$, and $Gd@C_{60}$ with aniline. *Journal of the American Chemical Society*, *118*, 6998–6999.
- Launio, P., Chorro, M., Verberck, B., Albouy, P.-A., Rouziere, S., Colson, D., Foget, A., Noe, L., Kataura, H., Monthieux, M., & Cambedouzou, J. (2010). Transformation of C_{70} peapods into double walled carbon nanotubes. *Carbon*, *48*, 89–98.
- Lavrik, N. V., & Datskos, P. G. (2003). Femtogram mass detection using photothermally actuated nanomechanical resonators. *Applied Physics Letters*, *82*, 2697–2699.
- Lebedeva, M. A., Chamberlain, T. W., & Khlobystov, A. N. (2015). Harnessing the synergistic and complementary properties of fullerene and transition-metal compounds for nanomaterial applications. *Chemical Reviews*, *115*, 11301–11351.
- Lee, S. U., & Han, Y.-K. (2004). Structure and stability of the defect fullerene clusters of C_{60} : C_{59} , C_{58} , and C_{57} . *Journal of Chemical Physics*, *121*, 3941–3942.
- Lee, C., Yang, W., & Parr, R. G. (1988). Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, *37*, 785–789.
- Lee, C., Kim, D., Jurecka, P., Tarakeswar, P., Hobza, P., & Kim, K. S. (2007). Understanding of assembly phenomena by aromatic – aromatic interactions: Benzene dimer and the substituted systems. *The Journal of Physical Chemistry. A*, *111*, 3446–3457.
- Lee, C., Wei, X., Kysar, J. W., & Hone, J. (2008). Measurement of the elastic properties and intrinsic strength of monolayer grapheme. *Science*, *321*, 385–388.
- Lherbier, A., Blase, X., Niquet, Y. M., Triozon, N., & Roche, S. (2008). Charge transport in chemically doped 2D graphene. *Physical Review Letters*, *101*, 036808–1–036808–4.
- Li, L., Reich, S., & Robertson, J. (2005). Defect energies of graphite: Density-functional calculations. *Physical Review B*, *72*, 184109–1–184109–10.
- Li, J., Wu, C., & Guan, L. (2009a). Lithium insertion/extraction properties of nanocarbon materials. *The Journal of Physical Chemistry C*, *113*, 18431–18435.
- Li, Y., Zhou, Z., Shen, P., & Chen, Z. (2009b). Structural and electronic properties of graphane nanoribbons. *The Journal of Physical Chemistry*, *113*, 15043–15045.
- Li, Y., Liu, S., Datta, D., & Li, Z. (2015). Surface hydrogenation regulated wrinkling and torque capability of hydrogenated graphene annulus under circular shearing. *Scientific Reports*, *5*, 16556.
- Liu, J., Dai, H., Hafner, J. H., Colbert, D. T., Smalley, R. E., Tans, S. J., & Dekker, C. (1997). Fullerene ‘crop circles’. *Nature*, *385*, 780–781.
- Liu, Z., Sun, X., Nakayama-Ratchford, N., & Dai, H. (2007). Supramolecular chemistry on water-soluble carbon nanotubes for drug loading and delivery. *ACS Nano*, *1*, 50–56.
- Lopez-Urias, F., Terrones, M., & Terrones, H. (2003). Electronic properties of giant fullerenes and complex graphitic nanostructures with novel morphologies. *Chemical Physics Letters*, *381*, 683–690.
- Lu, X., & Chen, Z. (2005). Curved pi-conjugation, aromaticity, and the related chemistry of small fullerenes ($< C_{60}$) and single-walled carbon nanotubes. *Chemical Reviews*, *105*, 3643–3696.
- Lu, A. J., & Pan, B. C. (2004). Nature of single vacancy in achiral carbon nanotubes. *Physical Review Letters*, *92*, 105504–1–105504–4.
- Lu, J., Zhang, X., & Zhao, X. (2000). Metal-cage hybridization in endohedral $La@C_{60}$, $Y@C_{60}$ and $Sc@C_{60}$. *Chemical Physics Letters*, *332*, 51–57.
- Lu, X., Chen, Z., Thiel, W., von Rague Schleyer, P., Huang, R., & Zheng, L. (2004). Properties of fullerene[50] and D_{5h} decachlorofullerene[50]: A computational study. *Journal of the American Chemical Society*, *126*, 14871–14878.

- Lu, X., Chen, Z., & von Rague Schleyer, P. (2005). Are Stone – Wales defect sites always more reactive than perfect sites in the sidewalls of single-wall carbon nanotubes? *Journal of the American Chemical Society*, *127*, 20–21.
- Lu, J., Yuan, D., Liu, J., Leng, W., & Kopley, T. E. (2008). Three dimensional single-walled carbon nanotubes. *Nano Letters*, *8*, 3325–3329.
- Ma, J., Alfe, D., Michaelides, A., & Wang, E. (2009). Stone-Wales defects in graphene and other planar sp^2 -bonded materials. *Physical Review B*, *80*, 033407-1–033407-4.
- MacKenzie, K. J., See, C. H., Dunens, O. M., & Harris, A. T. (2008). Do single-walled carbon nanotubes occur naturally? *Nature Nanotechnology*, *3*, 310.
- Malyi, O. I., Sopiha, K., Kulish, V. V., Tan, T. L., Manzhos, S., & Persson, C. (2015). A computational study of Na behavior on graphene. *Applied Surface Science*, *333*, 235–243.
- Manolopoulos, D. E., & Fowler, P. W. (1991). Structural proposals for endohedral metal-fullerene Complexes. *Chemical Physics Letters*, *187*, 1–7.
- Manolopoulos, D. E., & Fowler, P. W. (1992). Molecular graphs, point groups, and fullerenes. *Journal of Chemical Physics*, *96*, 7603–7614.
- Maruyama, S., & Yamaguchi, Y. (1998). A molecular dynamics demonstration of annealing to a perfect C_{60} Structure. *Chemical Physics Letters*, *286*, 343–349.
- Maseras, F., & Morokuma, K. (1995). IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *Journal of Computational Chemistry*, *16*, 1170–1179.
- Mashino, T., Nishikawa, D., Takahashi, K., Usui, N., Yamori, T., Seki, M., Endo, T., & Mochizuki, M. (2003). Antibacterial and antiproliferative activity of cationic fullerene derivatives. *Bioorganic and Medicinal Chemistry Letters*, *13*, 4395–4397.
- Matsuo, Y., Tahara, K., & Nakamura, E. (2003). Theoretical studies on structures and aromaticity of finite-length armchair carbon nanotubes. *Organic Letters*, *5*, 3181–3184.
- McKenzie, D. R., Davis, C. A., Cockayne, D. J. H., Muller, D. A., & Vassallo, A. M. (1992). The structure of the C_{70} molecule. *Nature*, *355*, 622–624.
- Mehta, G., & Rao, H. S. P. (1998). Synthetic studies directed towards bucky-balls and bucky-bowls. *Tetrahedron*, *54*, 13325–13370.
- Mehta, G., Panda, G., Yadav, R. D., & Kumar, K. R. (1997). A synthetic approach towards Pinakene, a $C_{28}H_{14}$ fragment of [70]-fullerene. *Indian Journal of Chemistry Section B*, *36*, 301–302.
- Melin, F., Chaur, M. N., Engmann, S., Elliott, B., Kumbhar, A., Athans, A. J., & Echegoyen, L. (2007). The large $Nd_3N@C_{2n}$ ($40 \leq n \leq 49$) cluster fullerene family: Preferential templating of a C_{88} cage by a trimetallic nitride cluster. *Angewandte Chemie, International Edition*, *46*, 9032–9035.
- Menon, M., & Srivastava, D. (1997). Carbon nanotube T junctions: Nanoscale metal-semiconductor-metal contact devices. *Physical Review Letters*, *79*, 4453–4456.
- Meyer, J. C., Kisielowski, C., Erni, R., Rossell, M. D., Crommie, M. F., & Zettl, A. (2008). Direct imaging of lattice atoms and topological defects in graphene membranes. *Nano Letters*, *8*, 3582–3586.
- Mielke, S. L., Troya, D., Zhang, S., Li, J.-L., Xiao, S., Car, R., Ruoff, R. S., Schatz, G. C., & Belytschko, T. (2004). The role of vacancy defects and holes in the fracture of carbon nanotubes. *Chemical Physics Letters*, *390*, 413–420.
- Mintmire, J. W., Dunlap, B. I., & White, C. T. (1992). Are fullerene tubules metallic? *Physical Review Letters*, *68*, 631–634.
- Miwa, R. H., Martins, T. B., & Fazzio, A. (2008). Hydrogen adsorption on boron doped graphene: An ab initio study. *Nanotechnology*, *19*, 155708-1–155708-7.
- Miyake, Y., Minami, T., Kikuchi, K., Kainosho, M., & Achiba, Y. (2000). Trends in structure and growth of higher fullerenes isomer structure of C_{86} and C_{88}^- . *Molecular Crystals and Liquid Crystals*, *340*, 553–558.
- Miyamoto, Y., Rubio, A., Berber, S., Yoon, M., & Tomanek, D. (2004). Spectroscopic characterization of Stone-Wales defects in nanotubes. *Physical Review B*, *69*, 121413-1–121413-4.

- Mizorogi, N., & Aihara, J. (2003). PM3 localization energies for the isolated-pentagon isomers of the C₈₄ Fullerene. *Physical Chemistry Chemical Physics*, 5, 3368–3371.
- Monthieux, M., & Kuznetsov, V. L. (2006). Who should be given the credit for the discovery of carbon nanotubes? *Carbon*, 44, 1621–1623.
- Moothi, K., Simate, G. S., Falcon, R., Iyuke, S. E., & Meyyappan, M. (2015). Carbon nanotube synthesis using coal pyrolysis. *Langmuir*, 31, 9464–9472.
- Moro, L., Ruoff, R. S., Becker, C. H., Lorents, D. C., & Malhotra, R. (1993). Studies of metallofullerene primary soots by laser and thermal desorption mass spectrometry. *Journal of Physical Chemistry*, 97, 6801–6805.
- Morokuma, K., Wang, Q., & Vreven, T. (2006). Performance evaluation of the three-layer ONIOM method: Case study for a zwitterionic peptide. *Journal of Chemical Theory and Computation*, 2, 1317–1324.
- Murry, R. L., Strout, D. L., Odom, G. K., & Scuseria, G. E. (1993). Role of sp³ carbon and 7-membered rings in fullerene annealing and fragmentation. *Nature*, 366, 665–667.
- Nasibulin, A. G., Pikhitsa, P. V., Jiang, H., Brown, D. P., Krashenninnikov, A. V., Anisimov, A. S., Queipo, P., Moysala, A., Gonzalez, D., Lientschnig, G., Hassanien, A., Shandakov, S. D., Lolli, G., Resasco, D. E., Choi, M., Tomanek, D., & Kauppinen, E. I. (2007a). A novel hybrid carbon material. *Nature Nanotechnology*, 2, 156–161.
- Nasibulin, A. G., Anisimov, A. S., Pikhitsa, P. V., Jiang, H., Brown, D. P., Choi, M., & Kauppinen, E. I. (2007b). Investigations of nanobud formation. *Chemical Physics Letters*, 446, 109–114.
- Neto, A. H. C. (2010). The carbon new age. *Materials Today*, 13(3), 12–17.
- Neto, A. H. C., Guinea, F., Peres, N. M. R., Novoselov, K. S., & Geim, A. K. (2009). The electronic properties of graphene. *Reviews of Modern Physics*, 81, 109–162.
- Nikitin, A., Ogasawara, H., Mann, D., Denecke, R., Zhang, Z., Dai, H., Cho, K., & Nilsson, A. (2005). Hydrogenation of single-walled carbon nanotubes. *Physical Review Letters*, 95, 225507-1–225507-1.
- Nishidate, K., & Hasegawa, M. (2005). Energetics of lithium ion adsorption on defective carbon nanotubes. *Physical Review B*, 71, 245418-1–245418-6.
- Niyogi, S., Hamon, M. A., Hu, H., Zhao, B., Bhowmik, P., Sen, R., Itkis, M. E., & Haddon, R. C. (2002). Chemistry of single-walled carbon nanotubes. *Accounts of Chemical Research*, 35, 1105–1113.
- Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Zhang, Y., Dubonos, S. V., Grigorieva, I. V., & Firsov, A. A. (2004). Electric field effect in atomically thin carbon films. *Science*, 306, 666–669.
- Novoselov, K. S., Jiang, D., Schedin, F., Booth, T. J., Khotkevich, V. V., Morozov, S. V., & Geim, A. K. (2005a). Two-dimensional atomic crystals. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 10451–10453.
- Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Katsnelson, M. I., Grigorieva, I. V., Dubonos, S. V., & Firsov, A. A. (2005b). Two-dimensional gas of massless Dirac fermions in graphene. *Nature*, 438, 197–200.
- Oberlin, A., Endo, M., & Koyama, T. (1976). Filamentous growth of carbon through benzene Decomposition. *Journal of Crystal Growth*, 32, 335–349.
- O'Brien, S. C., Heath, J. R., Curl, R. F., & Smalley, R. E. (1988). Photophysics of buckminsterfullerene and other carbon cluster ions. *Journal of Chemical Physics*, 88, 220–230.
- Okada, S. (2007). Radial-breathing mode frequencies for nanotubes encapsulating fullerenes. *Chemical Physics Letters*, 438, 59–62.
- Okada, S., & Saito, S. (1996). Number of extractable fullerene isomers and speciality of C₈₄. *Chemical Physical Letters*, 252, 94–100.
- Ormsby, J. L., & King, B. T. (2007). The regioselectivity of addition to carbon nanotube segments. *Journal of Organic Chemistry*, 72, 4035–4038.
- Osawa, E. (1970). Superaromaticity. *Kagaku (Kyoto)*, 25, 854–863.
- Osuna, S., Morera, J., Cases, M., Morokuma, K., & Sola, M. (2009). Diels – Alder reaction between cyclopentadiene and C₆₀: An analysis of the performance of the ONIOM method for

- the study of chemical reactivity in fullerenes and nanotubes. *The Journal of Physical Chemistry. A*, *113*, 9721–9726.
- Ouyang, M., Huang, J.-L., & Lieber, C. M. (2002). Fundamental electronic properties and applications of single-walled carbon nanotubes. *Accounts of Chemical Research*, *35*, 1018–1025.
- Palkar, A., Kumbhar, A., Athans, A. J., & Echegoyen, L. (2008). Pyridyl-functionalized and water-soluble carbon nano onions: First supramolecular complexes of carbon nano onions. *Chemistry of Materials*, *20*, 1685–1687.
- Park, S., Srivastava, D., & Cho, K. (2003). Generalized chemical reactivity of curved surfaces: Carbon Nanotubes. *Nano Letters*, *3*, 1273–1277.
- Park, S. S., Liu, D., & Hagelberg, F. (2005). Comparative investigation on non-IPR C_{68} and IPR C_{78} fullerenes encaging Sc_3N molecules. *The Journal of Physical Chemistry. A*, *109*, 8865–8873.
- Peng, X., Komatsu, N., Bhattacharya, S., Shimawaki, T., Aonuma, S., Kimura, T., & Osuka, A. (2007). Optically active single-walled carbon nanotubes. *Nature Nanotechnology*, *2*, 361–365.
- Perdew, J. P., Burke, K., & Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Physical Review Letters*, *77*, 3865–3868.
- Pereira, V. M., Neto, A. H. C., & Peres, N. M. R. (2009). Tight-binding approach to uniaxial strain in graphene. *Physical Review B*, *80*, 045401-1–045401-8.
- Pierson, H. O. (1993). *Handbook of carbon, graphite, diamonds and fullerenes: Processing, properties and applications*. New Jersey: Noyes.
- Piskoti, C., Yarger, J., & Zettl, A. (1998). C_{36} , a new carbon solid. *Nature*, *393*, 771–774.
- Ponomarenko, L. A., Schedin, F., Katsnelson, M. I., Yang, R., Hill, E. W., Novoselov, K. S., & Geim, A. K. (2008). Chaotic Dirac billiard in graphene quantum dots. *Science*, *320*, 356–358.
- Poonjarernsilp, C., Sano, N., Tamon, H., & Charinpanitkul, T. (2009). A model of reaction field in gas-injected arc-in-water method to synthesize single-walled carbon nanohorns: Influence of water temperature. *Journal of Applied Physics*, *106*, 104315-1–104315-7.
- Prinzbach, H., Weiler, A., Landenberger, P., Wahl, F., Worth, J., Scott, L. T., Gelmont, M., Olevano, D., & Issendorff, B. v. (2000). Gas-phase production and photoelectron spectroscopy of the smallest fullerene, C_{20} . *Nature*, *407*, 60–63.
- Priyakumar, U. D., & Sastry, G. N. (2001a). Heterobuckybowls: A theoretical study on the structure, bowl-to-bowl inversion barrier, bond length alternation, structure-inversion barrier relationship, stability, and synthetic feasibility. *Journal of Organic Chemistry*, *66*, 6523–6530.
- Priyakumar, U. D., & Sastry, G. N. (2001b). Tailoring the curvature, bowl rigidity and stability of heterobuckybowls: Theoretical design of synthetic strategies towards heterosumanenes. *Journal of Molecular Graphics and Modelling*, *19*, 266–269.
- Priyakumar, U. D., & Sastry, G. N. (2001c). Theory provides a clue to accomplish the synthesis of sumanene, $C_{21}H_{12}$, the prototypical C_{3v} -buckybowl. *Tetrahedron Letters*, *42*, 1379–1381.
- Qin, L.-C. (2007). Determination of the chiral indices (n, m) of carbon nanotubes by electron diffraction. *Physical Chemistry Chemical Physics*, *9*, 31–48.
- Radushkevich, L. V., & Lukyanovich, V. M. (1952). O strukture ugleroda, obrazujucesja pri termiceskom razlozenii okisi ugleroda na zeleznom kontakte. *Zurn Fisis Chim*, *26*, 88–95.
- Rao, C. N. R., Voggu, R., & Govindaraj, A. (2009a). Selective generation of single-walled carbon nanotubes with metallic, semiconducting and other unique electronic properties. *Nanoscale*, *1*, 96–105.
- Rao, F., Li, T., & Wang, Y. (2009b). Growth of all-carbon single-walled carbon nanotubes from diamonds and Fullerenes. *Carbon*, *47*, 3580–3584.
- Robertson, D. H., Brenner, D. W., & Mintmire, J. W. (1992). Energetics of nanoscale graphitic tubules. *Physical Review B*, *45*, 12592–12595.
- Robinson, J. A., Snow, E. S., Badescu, S. C., Reinecke, T. L., & Perkins, F. K. (2006). Role of defects in single-walled carbon nanotube chemical sensors. *Nano Letters*, *6*, 1747–1751.
- Robinson, J. T., Perkins, F. K., Snow, E. S., Wei, Z., & Sheehan, P. E. (2008). Reduced graphene oxide molecular sensors. *Nano Letters*, *8*, 3137–3140.
- Rohlfing, E. A., Cox, D. M., & Kaldor, A. (1994). Production and characterization of supersonic carbon cluster Beams. *Journal of Chemical Physics*, *81*, 3322–3330.

- Rojas, A., Martínez, M., Amador, P., & Torres, L. A. (2007). Increasing stability of the fullerenes with the number of carbon atoms: The experimental evidence. *The Journal of Physical Chemistry B*, *111*, 9031–9035.
- Saito, M., & Miyamoto, Y. (2001). Theoretical identification of the smallest fullerene, C₂₀. *Physical Review Letters*, *87*, 035503-1–035503-4.
- Saito, R., Fujita, M., Dresselhaus, G., & Dresselhaus, M. S. (1992). Electronic structure of chiral graphene tubules. *Applied Physics Letters*, *60*, 2204–2206.
- Saito, R., Dresselhaus, G., & Dresselhaus, M. S. (1998). *Physical properties of carbon nanotubes*. London: Imperial College Press.
- Sakurai, H., Daiko, T., & Hirao, T. (2003). A synthesis of sumanene, a fullerene fragment. *Science*, *301*, 1878.
- Sano, M., Kamino, A., Okamura, J., & Shinkai, S. (2001). Ring closure of carbon nanotubes. *Science*, *293*, 1299–1301.
- Sastry, G. N., & Priyakumar, U. D. (2001). The role of heteroatom substitution in the rigidity and curvature of buckybowls. A theoretical study. *Journal of the Chemical Society, Perkin Transactions*, *2*, 30–40.
- Sastry, G. N., Jemmis, E. D., Mehta, G., & Shah, S. R. (1993). Synthetic strategies towards C₆₀. Molecular mechanics and MNDO study on sumanene and related structures. *Journal of the Chemical Society, Perkin Transactions*, *2*, 1867–1871.
- Sastry, G. N., Rao, H. S. P., Bednarek, P., & Priyakumar, U. D. (2000). Effect of substitution on the curvature and bowl-to-bowl inversion barrier of bucky-bowls. Study of mono-substituted corannulenes (C₁₉XH₁₀, X = B⁻, N⁺, P⁺ and Si). *Chemical Communications*, 843–844.
- Saunders, M., Jiménez-Vázquez, H. A., Cross, R. J., & Poreda, R. J. (1993). Stable compounds of helium and neon: He@C₆₀ and Ne@C₆₀. *Science*, *259*, 1428–1430.
- Scheina, S., & Friedrich, T. (2008). A geometric constraint, the head-to-tail exclusion rule, may be the basis for the isolated-pentagon rule in fullerenes with more than 60 vertices. *Proceeding of the National Academy of Sciences of the United States of America*, *105*, 19142–19147.
- Scott, L. T., Boorum, M. M., McMahon, B. J., Hagen, S., Mack, J., Blank, J., Wegner, H., & de Meijere, A. (2002). A rational chemical synthesis of C₆₀. *Science*, *295*, 1500–1503.
- Scuseria, G. E. (1996). Ab initio calculations of fullerenes. *Science*, *271*, 942–945.
- Seiders, T. J., Elliot, E. L., Grube, G. H., & Siegel, J. S. (1999). Synthesis of corannulene and alkyl derivatives of corannulene. *Journal of the American Chemical Society*, *121*, 7804–7813.
- Seiders, T. J., Baldrige, K. K., Grube, G. H., & Siegel, J. S. (2001). Structure/energy correlation of bowl depth and inversion barrier in corannulene derivatives: Combined experimental and quantum mechanical analysis. *Journal of the American Chemical Society*, *123*, 517–525.
- Serra, S., Cavazzoni, C., Chiarotti, G. L., Scandolo, S., & Tosatti, E. (1999). Pressure-induced solid carbonates from molecular CO₂ by computer simulation. *Science*, *284*, 788–790.
- Shao, N., Gao, Y., Yoo, S., An, W., & Zeng, X. C. (2006). Search for lowest-energy fullerenes: C₉₈ to C₁₁₀. *The Journal of Physical Chemistry A*, *110*, 7672–7676.
- Shao, N., Gao, Y., & Zeng, X. C. (2007). Search for lowest-energy fullerenes 2: C₃₈ to C₈₀ and C₁₁₂ to C₁₂₀. *The Journal of Physical Chemistry C*, *111*, 17671–17677.
- Shukla, M. K., & Leszczynski, J. (2009). Fullerene (C₆₀) forms stable complex with nucleic acid base guanine. *Chemical Physics Letters*, *469*, 207–209.
- Shustova, N. B., Kuvychko, I. V., Bolskar, R. D., Seppelt, K., Strauss, S. H., Popov, A. A., & Boltalina, O. V. (2006). Trifluoromethyl derivatives of insoluble small-HOMO – LUMO-Gap hollow higher fullerenes. NMR and DFT structure elucidation of C₂-(C₇₄-D_{3h})(CF₃)₁₂, C₅-(C₇₆-T_d(2))(CF₃)₁₂, C₂-(C₇₈-D_{3h}(5))(CF₃)₁₂, C₅-(C₈₀-C_{2v}(5))(CF₃)₁₂, and C₂-(C₈₂-C₂(5))(CF₃)₁₂. *Journal of the American Chemical Society*, *128*, 15793–15798.
- Shustova, N. B., Newell, B. S., Miller, S. M., Anderson, O. P., Bolskar, R. D., Seppelt, K., Popov, A. A., Boltalina, O. V., & Strauss, S. H. (2007). Discovering and verifying elusive fullerene cage isomers: Structures of C₂-p¹¹-(C₇₄-D_{3h})(CF₃)₁₂ and C₂-p¹¹-(C₇₈-D_{3h}(5))(CF₃)₁₂. *Angewandte Chemie, International Edition*, *46*, 4111–4114.
- Simeon, T. M., Yanov, I., & Leszczynski, J. (2005). Ab initio quantum chemical studies of fullerene molecules with substitutes C₅₉X [X = Si, Ge, Sn], C₅₉X⁻ [X = B, Al, Ga, In], and C₅₉X [X = N, P, As, Sb]. *International Journal of Quantum Chemistry*, *105*, 429–436.

- Sinha, N., & Yeow, J. T.-W. (2005). Carbon nanotubes for biomedical applications. *IEEE Transactions on Nano Bioscience*, 4, 180–195.
- Sinnokrot, M. O., & Sherrill, C. D. (2004). Highly accurate coupled cluster potential energy curves for the benzene dimer: Sandwich, T-shaped, and parallel-displaced configurations. *The Journal of Physical Chemistry A*, 108, 10200–10207.
- Slanina, Z., Zhao, X., Lee, S.-L., & Osawa, E. (1997). C₉₀ temperature effects on relative stabilities of the IPR Isomers. *Chemical Physics*, 219, 193–200.
- Slanina, Z., Uhlik, F., Yoshida, M., & Osawa, E. (2000a). A computational treatment of 35 IPR isomers of C₈₈. *Fullerene Science and Technology*, 8, 417–432.
- Slanina, Z., Zhao, X., Deota, P., & Osawa, E. (2000b). Relative stabilities of C₉₂ IPR fullerenes. *Journal of Molecular Modeling*, 6, 312–317.
- Smalley, R. E. (1992). Self-assembly of the fullerenes. *Accounts of Chemical Research*, 25, 98–105.
- Smith, B. W., Monthieux, M., & Luzzi, D. E. (1998). Encapsulated C₆₀ in carbon nanotubes. *Nature*, 396, 323–324.
- Smith, B. W., Monthieux, M., & Luzzi, D. E. (1999). Carbon nanotube encapsulated fullerenes: A unique class of hybrid materials. *Chemical Physics Letters*, 315, 31–36.
- Sofo, J. O., Chaudhari, A. S., & Barber, G. D. (2007). Graphane: A two-dimensional hydrocarbon. *Physical Review B*, 75, 153401-1–153401-4.
- Sotiropoulou, S., & Chaniotakis, N. A. (2003). Carbon nanotube array-based biosensor. *Analytical and Bioanalytical Chemistry*, 375, 103–105.
- Stevens, R. M. D., Frederick, N. A., Smith, B. L., Morse, D. E., Stucky, G. D., & Hansma, P. K. (2000). Carbon nanotubes as probes for atomic force microscopy. *Nanotechnology*, 11, 1–5.
- Stevens, R. M. D., Nguyen, C. V., & Meyyappan, M. (2004). Carbon nanotube scanning probe for imaging in aqueous environment. *IEEE Transactions on Nano Bioscience*, 3, 56–60.
- Stewart, J. J. P. (1989). Optimization of parameters for semiempirical methods I Method. *Journal of Computational Chemistry*, 10, 209–220.
- Stoilova, O., Jérôme, C., Detrembleur, C., Mouithys-Mickalad, A., Manolova, N., Rashkova, I., & Jérôme, R. (2007). C₆₀-containing nanostructured polymeric materials with potential biomedical applications. *Polymer*, 48, 1835–1843.
- Stone, A. J., & Wales, D. J. (1986). Theoretical studies of icosahedral C₆₀ and some related species. *Chemical Physics Letters*, 128, 501–503.
- Strano, M. S. (2003). Probing chiral selective reactions using a revised Kataura plot for the interpretation of single-walled carbon nanotube spectroscopy. *Journal of the American Chemical Society*, 125, 16148–16153.
- Strano, M. S. (2007). Carbon nanotubes: Sorting out left from right. *Nature Nanotechnology*, 2, 340–341.
- Suchanek, W. L., Libera, J. A., Gogotsi, Y., & Yoshimura, M. (2001). Behavior of C₆₀ under hydrothermal conditions: Transformation to amorphous carbon and formation of carbon nanotubes. *Journal of Solid State Chemistry*, 160, 184–188.
- Suenaga, K., Wakabayashi, H., Koshino, M., Sato, Y., Urita, K., & Iijima, S. (2007). Imaging active topological defects in carbon nanotubes. *Nature Nanotechnology*, 2, 358–360.
- Sulman, E., Yanov, I., & Leszczynski, J. (1999). An active site model and the catalytic activity mechanism of the new fullerene-based catalyst - (η^2 -C₆₀)Pd(PPh₃)₂. *Fullerenes, Nanotubes, and Carbon Nanostructures*, 7, 467–484.
- Sun, G. (2003). Assigning the major isomers of fullerene C₈₈ by theoretical ¹³C NMR spectra. *Chemical Physics Letters*, 367, 26–33.
- Sun, G., & Kertesz, M. (2002). ¹³C NMR spectra for IPR isomers of fullerene C₈₆. *Chemical Physics*, 276, 107–114.
- Suzuki, S., & Kobayashi, Y. (2007). Healing of low-energy irradiation-induced defects in single-walled carbon nanotubes at room temperature. *The Journal of Physical Chemistry C*, 111, 4524–4528.
- Sygula, A., & Rabideau, P. W. (1999). Non-pyrolytic syntheses of buckybowls: Corannulene, cyclopentacorannulene, and a semibuckminsterfullerene. *Journal of the American Chemical Society*, 121, 7800–7803.

- Tagmatarchis, N., Arcon, D., Prato, M., & Shinohara, H. (2002). Production, isolation and structural characterization of [92]fullerene isomers. *Chemical Communications*, 2992–2993.
- Tang, A. C., & Huang, F. Q. (1995). Electronic structures of giant fullerenes with I_h symmetry. *Physical Review B*, 51, 13830–13832.
- Tang, A. C., Li, Q. S., Liu, C. W., & Li, J. (1993). Symmetrical clusters of carbon and boron. *Chemical Physics Letters*, 201, 465–469.
- Taylor, R. (1992). The third form of carbon: A new era in chemistry. *Interdisciplinary Science Reviews*, 17, 161–170.
- Taylor, R., Hare, J. P., Abdul-Sada, A. K., & Kroto, H. W. (1990). Isolation, separation and characterisation of the fullerenes C_{60} and C_{70} : The third form of carbon. *Journal of the Chemical Society, Chemical Communications*, 1423–1425.
- Taylor, R., Langley, G. J., Dennis, T. J. S., Kroto, H. W., & Walton, D. R. M. (1992). A mass spectrometric–NMR study of fullerene-78 isomers. *Journal of the Chemical Society, Chemical Communications*, 1043–1046.
- Taylor, R., Langley, G. J., Avent, A. G., Dennis, T. J. S., Kroto, H. W., & Walton, D. R. M. (1993). ^{13}C NMR spectroscopy of C_{76} , C_{78} , C_{84} and mixtures of C_{86} – C_{102} ; Anomalous chromatographic behaviour of C_{82} , and evidence for $C_{70}H_{12}$. *Journal of the Chemical Society, Perkin Transactions*, 2, 1029–1036.
- Terrones, M., Terrones, G., & Terrones, H. (2002). Structure, chirality, and formation of giant icosahedral fullerenes and spherical graphitic onions. *Structural Chemistry*, 13, 373–384.
- Thilgen, C., & Diederich, F. (2006). Structural aspects of fullerene chemistry – A journey through fullerene chirality. *Chemical Reviews*, 106, 5049–5135.
- Thrash, T. P., Cagle, D. W., Alford, J. M., Wright, K., Ehrhardt, G. J., Mirzadeh, S., & Wilson, L. J. (1999). Toward fullerene-based radiopharmaceuticals: High-yield neutron activation of endohedral ^{165}Ho metallofullerenes. *Chemical Physics Letters*, 308, 329–336.
- Troshin, P. A., Avent, A. G., Darwish, A. D., Martsinovich, N., Abdul-Sada, A. K., Street, J. M., & Taylor, R. (2005). Isolation of two seven-membered ring C_{58} fullerene derivatives: $C_{58}F_{17}CF_3$ and $C_{58}F_{18}$. *Science*, 309, 278–281.
- Troyanov, S. I., & Tamm, N. B. (2009). Cage connectivities of C_{88} (33) and C_{92} (82) fullerenes captured as trifluoromethyl derivatives, $C_{88}(CF_3)_{18}$ and $C_{92}(CF_3)_{16}$. *Chemical Communications*, 6035–6037.
- Valsakumar, M. C., Subramanian, N., Yousuf, M., Sahu, P. C., Hariharan, Y., Bharathi, A., Sastry, V. S., Janaki, J., Rao, G. V. N., Radhakrishnan, T. S., & Sundar, C. S. (1993). Crystal structure and disorder in solid C_{70} . *Physical Review B*, 48, 9080–9085.
- Velasco-Santos, C., Martínez-Hernández, A. L., Consultchi, A., Rodríguez, R., & Castaño, V. M. (2003). Naturally produced carbon nanotubes. *Chemical Physics Letters*, 373, 272–276.
- Vostrowsky, O., & Hirsch, A. (2004). Molecular peapods as supramolecular carbon allotropes. *Angewandte Chemie, International Edition*, 43, 2326–2329.
- Wahl, F., Worth, J., & Prinzbach, H. (1993). The pagodane route to dodecahedranes: An improved approach to the $C_{20}H_{20}$ parent framework; partial and total functionalizations – Does C_{20} -fullerene exist? *Angewandte Chemie International Edition in English*, 32, 1722–1726.
- Wanbayor, R., & Ruangpornvisuti, V. (2008). Theoretical study of adsorption of C1–C3 alkoxides on various cap-ended and open-ended armchair (5,5) single-walled carbon nanotubes. *Carbon*, 46, 12–18.
- Wang, G.-W., Zhang, X.-H., Zhan, H., Guo, Q.-X., & Wu, Y.-D. (2003). Accurate calculation, prediction, and assignment of 3He NMR chemical shifts of Helium-3-encapsulated fullerenes and fullerene derivatives. *Journal of Organic Chemistry*, 68, 6732–6738.
- Wang, C., Zhou, G., Liu, H., Wu, J., Qiu, Y., Gu, B.-L., & Duan, W. (2006). Chemical functionalization of carbon nanotubes by carboxyl groups on Stone-Wales defects: A density functional theory study. *The Journal of Physical Chemistry. B*, 110, 10266–10271.
- Wang, X., Tabakman, S. M., & Dai, H. (2008). Atomic layer deposition of metal oxides on pristine and functionalized grapheme. *Journal of the American Chemical Society*, 130, 8152–8153.

- Wang, L., Drahushuk, L. W., Cantley, L., Koenig, S. P., Liu, X., Pellegrino, J., Strano, M. S., & Bunch, J. S. (2015a). Molecular valves for controlling gas phase transport made from discrete ångström-sized pores in graphene. *Nature Nanotechnology*, *10*, 785–790.
- Wang, Y., Díaz-Tendero, S., Manuel Alcamí, M., & Martín, F. (2015b). Cage connectivity and frontier π orbitals govern the relative stability of charged fullerene isomers. *Nature Chemistry*, *7*, 927–934.
- WenXing, B., ChangChun, Z., & WanZhao, C. (2004). Simulation of Young's modulus of single-walled carbon nanotubes by molecular dynamics. *Physica B*, *352*, 156–163.
- Wikipedia – <http://en.wikipedia.org/wiki/Carbon>.
- Woodward, R. B., & Hoffmann, R. (1969). The conservation of orbital symmetry. *Angewandte Chemie (International Edition in English)*, *8*, 781–853.
- Wu, J., & Hagelberg, F. (2008). Computational study on C_{80} enclosing mixed trimetallic nitride clusters of the form $Gd_xM_{3-x}N$ ($M = Sc, Sm, Lu$). *The Journal of Physical Chemistry C*, *112*, 5770–5777.
- Wu, Y.-T., & Siegel, J. S. (2006). Aromatic molecular-bowl hydrocarbons: Synthetic derivatives, their structures, and physical properties. *Chemical Reviews*, *106*, 4843–4867.
- Wu, X., & Zeng, X. C. (2009). Periodic graphene nanobuds. *Nano Letters*, *9*, 250–256.
- Xia, J., Chen, F., Li, J., & Tao, N. (2009). Measurement of the quantum capacitance of grapheme. *Nature Nanotechnology*, *4*, 505–509.
- Xie, S.-Y., Gao, F., Lu, X., Huang, R.-B., Wang, C.-R., Zhang, X., Liu, M.-L., Deng, S.-L., & Zheng, L.-S. (2004). Capturing the labile fullerene[50] as $C_{50}Cl_{10}$. *Science*, *304*, 699.
- Xue, Y., Ding, Y., Niu, J., Xia, Z., Roy, A., Chen, H., Qu, J., Wang, Z. L., & Dai, L. (2015). Rationally designed graphene-nanotube 3D architectures with a seamless nodal junction for efficient energy conversion and storage. *Science Advances*, *1*, 1400198.
- Yakobson, B. I., Brabec, C. J., & Bernholc, J. (1996). Nanomechanics of carbon tubes: Instabilities beyond linear response. *Physical Review Letters*, *76*, 2511–2514.
- Yamada, M., Nakahodo, T., Wakahara, T., Tsuchiya, T., Maeda, Y., Akasaka, T., Kako, M., Yoza, K., Horn, E., Mizorogi, N., Kobayashi, K., & Nagase, S. (2005). Positional control of encapsulated atoms inside a fullerene cage by exohedral addition. *Journal of the American Chemical Society*, *127*, 14570–14571.
- Yamada, M., Akasaka, T., & Nagase, S. (2010). Endohedral metal atoms in pristine and functionalized fullerene cages. *Accounts of Chemical Research*, *43*, 92–102.
- Yang, S. H., Shin, W. H., Lee, J. W., Kim, S. Y., Woo, S. I., & Kang, J. K. (2006a). Interaction of a transition metal atom with intrinsic defects in single-walled carbon nanotubes. *The Journal of Physical Chemistry B*, *110*, 13941–13946.
- Yang, S. H., Shin, W. H., & Kang, J. K. (2006b). Ni adsorption on Stone-Wales defect sites in single-wall carbon nanotubes. *Journal of Chemical Physics*, *125*, 084705-1–084705-5.
- Yang, F. H., Lachawiec, A. J., Jr., & Yang, R. T. (2006c). Adsorption of spillover hydrogen atoms on single-wall carbon nanotubes. *The Journal of Physical Chemistry B*, *110*, 6236–6244.
- Yanov, I., Leszczynski, J., Sulman, E., Matveeva, V., & Semagina, N. (2004). Modeling of the molecular structure and catalytic activity of the new fullerene-based catalyst $(\eta^2-C_{60})Pd(PPh_3)_2$: An application in the reaction of selective hydrogenation of acetylenic alcohols. *International Journal of Quantum Chemistry*, *100*, 810–817.
- Yu, D., Goh, K., Wang, H., Wei, L., Jiang, W., Zhang, Q., Dai, L., & Chen, Y. (2014). Scalable synthesis of hierarchically structured carbon nanotube–graphene fibres for capacitive energy storage. *Nature Nanotechnology*, *9*, 555–562.
- Yuan, H., & He, Z. (2015). Graphene-modified electrodes for enhancing the performance of microbial fuel cells. *Nanoscale*, *2015*, 7022–7029.
- Yumura, T., Nozaki, D., Bandow, S., Yoshizawa, K., & Iijima, S. (2005a). End-cap effects on vibrational structures of finite-length carbon nanotubes. *Journal of the American Chemical Society*, *127*, 11769–11776.
- Yumura, T., Sato, Y., Suenaga, K., & Iijima, S. (2005b). Which do endohedral Ti_2C_{80} metallofullerenes prefer energetically: $Ti_2@C_{80}$ or $Ti_2C_2@C_{78}$? A theoretical study. *The Journal of Physical Chemistry B*, *109*, 20251–20255.

- Yumura, T., Kertesz, M., & Iijima, S. (2007). Local modifications of single-wall carbon nanotubes induced by bond formation with encapsulated fullerenes. *The Journal of Physical Chemistry. B*, *111*, 1099–1109.
- Zhang, J., & Zuo, J. M. (2009). Structure and diameter-dependent bond lengths of a multi-walled carbon nanotube revealed by electron diffraction. *Carbon*, *47*, 3515–3528.
- Zhang, B. L., Wang, C. Z., Ho, K. M., Xu, C. H., & Chan, C. T. (1993). The geometry of large fullerene cages: C_{72} to C_{102} . *Journal of Chemical Physics*, *98*, 3095–3102.
- Zhang, G., Qi, P., Wang, X., Lu, Y., Mann, D., Li, X., & Dai, H. (2006). Hydrogenation and hydrocarbonation and etching of single-walled carbon nanotubes. *Journal of the American Chemical Society*, *128*, 6026–6027.
- Zhang, H., Cao, G., Wang, Z., Yang, Y., Shi, Z., & Gu, Z. (2008). Influence of ethylene and hydrogen flow rates on the wall number, Crystallinity, and length of millimeter-long carbon nanotube array. *The Journal of Physical Chemistry C*, *112*, 12706–12709.
- Zhao, K., & Pitzer, R. M. (1996). Electronic structure of C_{28} , $Pa@C_{28}$, and $U@C_{28}$. *Journal of Physical Chemistry*, *100*, 4798–4802.
- Zhao, Y., & Truhlar, D. G. (2007). Size-selective supramolecular chemistry in a hydrocarbon nanoring. *Journal of the American Chemical Society*, *129*, 8440–8442.
- Zhao, Y., & Truhlar, D. G. (2008). Computational characterization and modeling of buckyball tweezers: Density functional study of concave–convex interactions. *Physical Chemistry Chemical Physics*, *10*, 2813–2818.
- Zhao, X., Slanina, Z., & Goto, H. (2004a). Theoretical studies on the relative stabilities of C_{96} IPR fullerenes. *The Journal of Physical Chemistry. A*, *108*, 4479–4484.
- Zhao, X., Goto, H., & Slanina, Z. (2004b). C_{100} IPR fullerenes: Temperature-dependent relative stabilities based on the Gibbs function. *Chemical Physics*, *306*, 93–104.
- Zhou, Z., Steigerwald, M., Hybertsen, M., Brus, L., & Friesner, R. A. (2004). Electronic structure of tubular aromatic molecules derived from the metallic (5,5) armchair single wall carbon nanotube. *Journal of the American Chemical Society*, *126*, 3597–3607.
- Zhou, L., Gao, C., Zhu, D. D., Xu, W., Chen, F. F., Palkar, A., Echegoyen, L., & Kong, E. S.-W. (2009). Facile functionalization of multilayer fullerenes (carbon nanoions) by nitrene chemistry and grafting from strategy. *Chemistry - European Journal*, *15*, 1389–1396.
- Zhu, Z. H., Hatori, H., Wang, S. B., & Lu, G. Q. (2005). Insights into hydrogen atom adsorption on and the electrochemical properties of nitrogen-substituted carbon materials. *The Journal of Physical Chemistry. B*, *109*, 16744–16749.
- Zhu, Y., Li, L., Zhang, C., Casillas, G., Sun, Z., Yan, Z., Ruan, G., Peng, Z., Raji, A.-R., Kittrell, C., Hauge, R. H., & Tour, J. M. (2012). A seamless three-dimensional carbon nanotube graphene hybrid material. *Nature Communications*, *3*, 1225.
- Zope, R. R., Baruah, T., Pederson, M. R., & Dunlap, B. I. (2008). Static dielectric response of icosahedral fullerenes from C_{60} to C_{2160} characterized by an all-electron density functional theory. *Physical Review B*, *77*, 115452-1–115452-5.

Optical Properties of Quantum Dot Nano-composite Materials Studied by Solid-State Theory Calculations

28

Ying Fu and Hans Ågren

Contents

Introduction	1260
Solid-State k-p Theory	1261
Excitons in Nanostructures	1266
Polarization and Optical Properties of Exciton-Polaritons	1270
Exciton-Polariton Photonic Crystals	1275
Photonic Dispersion of QD Dimer Systems	1276
Lossless Dielectric Constant of QD Dimer Systems	1282
Multiple-Photon and Multiple-Carrier Processes	1283
Multiphoton Process	1283
Impact Ionization and Auger Recombination	1284
Summary	1291
Bibliography	1292

Abstract

This chapter reviews the fundamental concepts of excitons and excitonic polaritons and their extraordinary optical properties in quantum dot nano-composite materials. By starting with the optical excitation of an exciton in the nanostructure we show that the effective dielectric constant of the nanostructure becomes significantly modified due to the exciton generation and recombination, resulting

Y. Fu (✉)

Division of Theoretical Chemistry and Biology, School of Biotechnology, Royal Institute of Technology, Stockholm, Sweden

e-mail: fyg@theochem.kth.se

H. Ågren

Division of Theoretical Chemistry and Biology, School of Biotechnology, Royal Institute of Technology, Stockholm, Sweden

Organic Chemistry, Bogdam Khmel'nitskii National University, Cherkassy, Ukraine

e-mail: agren@theochem.kth.se

in high positive and negative dielectric constants. We also discuss single exciton generation by multiple photons and multiple exciton generation by single photon. All these nonlinear optical properties of quantum dot nano-composite materials offer novel possibilities and are expected to have deep impact in nanophotonics.

Introduction

Materials with nanoscale features have the potential to revolutionize optoelectronics, permitting new and interesting device and system capabilities. The key design element is the geometry, that is, size and shape, which defines the properties of the nanostructured material and has led to the expectation of new and/or significantly improved physical, electrical, and optical properties. To precisely engineer light-matter interaction at the nanoscale, electron-hole pairs, that is, excitons, and their coupling with photons, exciton-polaritons, are becoming increasingly important. Excitons are of great interest in nanoscience because it has been discovered that their properties can be dictated by the size and the shape of a nanostructure in which they are confined, in addition to the constitution of the nanostructure. The new aspect of excitons that is prevalent in or even defining nanoscience is that the physical size and shape of the nanostructure are parameters that significantly influence their properties.

Exciton polaritons were intensively studied in the 1960s and 1970s, and their manifestation in various optical phenomena, including light reflection and transmission, photoluminescence, and resonant light scattering are well established and documented by now, see, for example, the dedicated volume (Sturge and Rashba 1982) and references therein. We refer to Kavokin (2007) that presents detailed discussions of energy dispersion of exciton-polaritons in microcavities (largely about quantum wells, QWs), to Scholes and Rumbles (2006) for a very intuitive description of excitons in nanoscale systems and a detailed review about excitons in polymers and nanotubes, and to Weisbuch et al. (2000) for research and development of excitons and photons in confined structures from three-dimensionally (3D) extended states (bulk) to atom-like quantum dots (QDs) and their applications in light emitting devices.

In semiconductor technology, the geometric miniaturization of individual components and devices have been scaling according to Moore's Law (International Technology Roadmap for Semiconductors; Fu et al. 2006a). Photonics has also been catching up (Thylén et al. 2006) but at a slow pace due to the limitations imposed by the diffraction limit. Resonant excitonic states in nanostructures, after coupling with light to form excitonic polaritons, provide a source for both high and negative dielectric constant (Fu et al. 2006b, 2008; Zia et al. 2006), thus offering novel possibilities to drastically increase the integration density and new functionalities in photonics at the nanoscale. Rapid growth of nanotechnology also greatly expands the clinical opportunities. Colloidal QDs are highly fluorescent with excellent photochemical stability, extreme brightness, and broad excitation spectral range, which have been introduced successfully in many fluorescence-based optical imaging applications (Medintz et al. 2005). Surface-modified and water-soluble

QDs open a new era in cell imaging and bio-targeting as transport vehicles for therapeutic drug delivery to different diseases such as cancer and atherosclerosis (Derfus et al. 2007; Molnár et al. 2010; Vashist et al. 2006).

In this chapter we concentrate on nano-semiconductor systems, where it has become possible during the last 15 years to handle and control single excitons. More specifically we focus on excitons and exciton-polaritons in semiconductor QDs. This chapter presents first a theoretical background of the solid-state theory, the exciton and the coupling between the exciton and the photon (excitonic polariton). Optical properties of exciton polaritons in nanostructures and their photonic engineering are then discussed.

Solid-State k·p Theory

The theoretical description of electronic and optical properties of nanostructures is currently advancing using modern quantum methodologies such as pseudo potential, tight-binding, and atomistic semi-empirical pseudo potential methods (Allan and Delerue 2008; Franceschetti et al. 2006; Jiang et al. 2009; Rabani and Baer 2008). However, these studies have been limited to light atoms and small nanostructure sizes. Our target nanostructure systems are about 10 nm in diameter and even more where macroscopic micrometer scale solid-state theory is well developed.

In isolated atoms the energy levels are sharply defined. When two atoms are brought close to each other their electron wave functions overlap. As a result of the interaction between the electrons, it turns out that each single state of the isolated atom splits into two states with different energies. The degree of splitting increases as the inter-atomic separation decreases. Similarly, if five atoms are placed in close proximity, then each original energy level splits into five new levels. The same process occurs in a solid, where there are roughly 10^{28} atoms/m³: The energy levels associated with each state of the isolated atom spread into essentially continuous energy bands separated from each other by energy gaps.

Before further examining the various properties of semiconductors it is extremely useful to examine the atomic structure of some of the elements which make up commonly used semiconductors as listed in Table 1.

A very important conclusion can be drawn about the elements making up the semiconductors: The outmost valence electrons are made up of electrons in either the s- or p-type orbitals. While this conclusion is strictly true for elements in

Table 1 Atomic structures of typical elements making up common semiconductors

Element	Core electrons	Valence electrons
IV-Si	$1s^2 2s^2 2p^6$	$3s^2 3p^2$
IV-Ge	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10}$	$4s^2 4p^2$
III-Ga	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10}$	$4s^2 4p^1$
V-As	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10}$	$4s^2 4p^3$
II-Cd	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10} 4s^2 4p^6 4d^{10}$	$5s^2$
VI-Se	$1s^2 2s^2 2p^6 3s^2 3p^6 3d^{10}$	$4s^2 4p^4$

the atomic form, it turns out that even in crystalline semiconductors the electrons in the valence band (VB) and conduction band (CB) retain this s- or p-type character. The core electrons are usually not of interest, except of some special characterization-type experiments. Note that II–VI compounds such as CdSe are normally complicated and their lattice structures are hexagonal.

In general, it is found that when atoms exchange or share valence electrons so that the complement of quantum states is completed, they have a lower electrostatic energy for their combined electron patterns than when they are separate. For example, silicon has four valence electrons grouped in two closely spaced energy levels (3 s and 3 p, see Table 1); it can combine with other silicon atoms by sharing four valence electrons with four surrounding silicon atoms in an endless array. The atoms around any one atom are centered at the corners of a regular tetrahedron: the tetrahedral bond. This creates the diamond crystal structure.

Essentially all semiconductors of interest for electronics and optoelectronics have the face-centered cubic (fcc) structure. However, they have two atoms per basis. The coordinates of the two basis atoms are (000) and $(a/4)(111)$, indicated in d_1 and d_2 in Fig. 1b. Here a is normally denoted as the crystal lattice constant. If the two atoms of the basis are identical, the structure is called the diamond structure. Semiconductors such as silicon, germanium, and carbon fall into this category. If the two atoms are different, for example, GaAs, AlAs, and CdS, the structure is called zincblende.

The intrinsic property of a crystal is that the environment around a given atom or group of atoms is exactly the same as the environment around another atom or similar group of atoms. Many of the properties of crystals and many of the theoretical techniques used to describe crystals derive from such a periodicity of crystalline structures. This suggests the use of Fourier analysis in the form of real space–reciprocal space or wave vector space duality for crystal problem discussions. Many concepts are best understood in terms of functions of the wave vector. We prefer to describe a wave with wavelength λ as a plane wave with wave

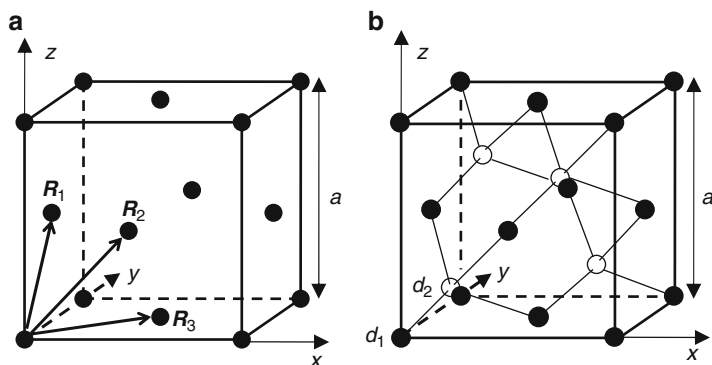


Fig. 1 (a) Face-centered cubic (fcc) lattice with primitive translation vectors R_1 , R_2 , and R_3 (b) Zincblende crystal structure where d_1 and d_2 denote two different atoms (can be identical such as in Si crystal in which case the lattice structure is called diamond). In both (a) and (b), a denotes the lattice constant

vector \mathbf{k} of magnitude $2\pi/\lambda$ and propagation direction perpendicular to the wave front. The space of the wave vectors is called the reciprocal space.

The $\mathbf{k}\cdot\mathbf{p}$ method is much used in analyzing semiconductor nanostructures as it gives a good approximation of the states close to the Γ point ($k = 0$), which is the most relevant region for many device applications. The approximation is given with good accuracy, while in many nanostructure calculations, this is the only feasible numerical model that can be implemented (even though it is still much restricted due to the computational capability).

The goal is essentially to solve the Schrödinger equation for a nanostructure to get the eigenenergies and their associated eigenvectors (wave functions). The wave functions usually have too high frequencies to be feasible to calculate explicitly using numerical methods on computers as an inordinately high number of grid points would be necessary to capture an acceptable numerical representation of the wave function. A solution is to separate the wave functions into an oscillatory part and a modulating part which is of the same scale as the heterostructure. This is the basic idea of the envelope function approximation – the modulating part is called the envelope function. We write the envelope function as a Bloch function:

$$\psi_{nk}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{nk}(\mathbf{r}) \quad (1)$$

where n is the state index and k is the wave vector. The Schrödinger equation for this wave function is simply

$$\left[\frac{\mathbf{p}^2}{2m_0} + V(\mathbf{r}) \right] \psi_{nk}(\mathbf{r}) = E_n(\mathbf{k}) \psi_{nk}(\mathbf{r}) \quad (2)$$

Substituting the factorization into the equation requires some care in evaluating the product of the momentum operator \mathbf{p} and the two parts of the wave function. With $\mathbf{p} = -i\hbar\nabla$ and knowing from vector calculus that $\nabla e^{i\mathbf{k}\cdot\mathbf{r}} = i\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{r}}$, the product of the momentum operator and the wave function becomes

$$-i\hbar\nabla [e^{i\mathbf{k}\cdot\mathbf{r}} u_{nk}(\mathbf{r})] = -i\hbar e^{i\mathbf{k}\cdot\mathbf{r}} (\nabla + i\mathbf{k}) u_{nk}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} (\mathbf{p} + \hbar\mathbf{k}) u_{nk}(\mathbf{r}) \quad (3)$$

Applying the momentum operator a second time gives $e^{i\mathbf{k}\cdot\mathbf{r}} (\mathbf{p} + \hbar\mathbf{k})^2 u_{nk}(\mathbf{r})$, so the Schrödinger equation can be written in the following way that the oscillatory part will cancel out:

$$e^{i\mathbf{k}\cdot\mathbf{r}} \left[\frac{(\mathbf{p} + \hbar\mathbf{k})^2}{2m_0} + V(\mathbf{r}) \right] u_{nk}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} E_n(\mathbf{k}) u_{nk}(\mathbf{r}) \quad (4)$$

Expanding the $(\mathbf{p} + \hbar\mathbf{k})^2$ term gives

$$\left[\frac{\mathbf{p}^2}{2m_0} + V(\mathbf{r}) + \frac{\hbar\mathbf{k} \cdot \mathbf{p}}{m_0} + \frac{\hbar^2 k^2}{2m_0} \right] u_{nk}(\mathbf{r}) = E_n(\mathbf{k}) u_{nk}(\mathbf{r}) \quad (5)$$

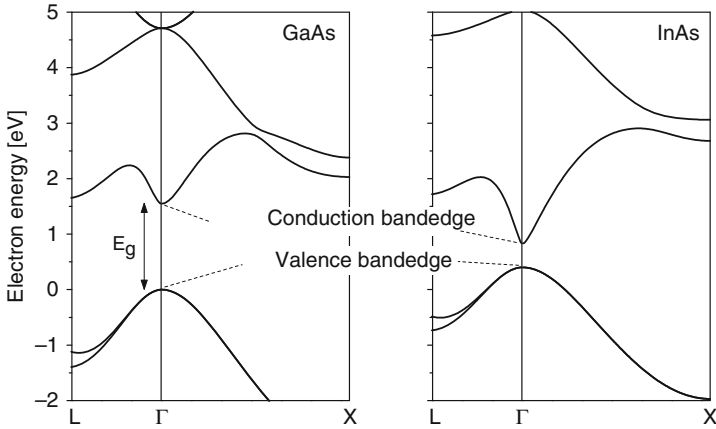


Fig. 2 Energy band structures of electrons in GaAs and InAs bulk materials

The first two terms are identical to the original Hamiltonian, so if the two other terms are treated as a small perturbation the Hamiltonian can be expressed as $(H_0 + H_1 + H_2) u_{nk}(\mathbf{r}) = E_n(\mathbf{k}) u_{nk}(\mathbf{r})$, and treat $H_1 = (\hbar/m_0) \mathbf{k} \cdot \mathbf{p}$ and $H_2 = \hbar^2 k^2 / 2m_0$ as first-order and second-order perturbations, respectively.

If the equation is solved for $\mathbf{k} = 0$ with only H_0 remaining nonzero, the result is a set of eigenvectors $u_{n0}(\mathbf{r})$, typically at the optimal points such as the Γ point in Fig. 2 which shows the energy band structure of most common GaAs and InAs bulk materials. For the following discussion we express the eigenfunction using Dirac notation as $|n\rangle$, and eigenvalues E_n . These form a complete basis set with orthogonal eigenfunctions which can be used in perturbation calculations of states with $\mathbf{k} \neq 0$.

As the atoms of the elements making up the semiconductors are brought together to form the crystal, the valence electronic states are perturbed by the presence of neighboring atoms. While the original atomic functions describing the valence electrons, of course, no longer are eigenstates of the problem, their characteristics can be used as a good approximate set of basis states to describe the “crystalline” electrons. For most semiconductor materials of interest, the atomic functions required to describe the outermost electrons are s , p_x , p_y , and p_z types, as shown by Table 1. A common approach of choosing the above $|n\rangle$, at the Γ point is to follow Kane (1957) and define a set of eight states, that is, $u_{nk}(\mathbf{r})$ in Eq. 1 at $\mathbf{k} = 0$, $|S \uparrow\rangle$, $|X \uparrow\rangle$, $|Y \uparrow\rangle$, $|Z \uparrow\rangle$, $|S \downarrow\rangle$, $|X \downarrow\rangle$, $|Y \downarrow\rangle$, $|Z \downarrow\rangle$, where the arrows indicate spin up and down. $|X, Y, Z\rangle$, denotes degenerate VBs while $|S\rangle$ denotes the CB. Other remote bands are not included unless specifically required.

The eigen function for a given \mathbf{k} is a linear combination of the basis functions:

$$u_{nk}(\mathbf{r}) = \sum_m c_{nm}(\mathbf{k}) |m\rangle \quad (6)$$

so the objective is now to find the coefficients c_{nm} that form our envelope parts together with the basis functions at Γ .

If we substitute this linear combination into the Schrödinger equation, we get

$$H \sum_m c_{nm}(\mathbf{k}) |m\rangle = \sum_m c_{nm}(\mathbf{k}) H |m\rangle = E_n(\mathbf{k}) \sum_m c_{nm}(\mathbf{k}) |m\rangle \tag{7}$$

Multiply this on the left with the conjugate of any, say $|\ell\rangle$, of the basis functions, and we obtain

$$\sum_m c_{nm}(\mathbf{k}) \langle \ell | H | m \rangle = E_n(\mathbf{k}) \sum_m c_{nm}(\mathbf{k}) \langle \ell | m \rangle = E_n(\mathbf{k}) c_{n\ell}(\mathbf{k}) \tag{8}$$

The right-hand part is the result of wave function's orthonormal property.

Inserting the expanded $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian Eq. 5, multiplying with the conjugate, and integrating over an unit cell using the fact that the basis functions are orthonormal gives

$$\sum_m \left[\left(E_m + \frac{\hbar^2 k^2}{2m_0} \right) \delta_{\ell m} + \frac{\hbar}{m_0} \langle \ell | \mathbf{k} \cdot \mathbf{p} | m \rangle \right] c_{nm}(\mathbf{k}) = E_n(\mathbf{k}) c_{n\ell}(\mathbf{k}) \tag{9}$$

The first-order approximation is thus

$$E_n(\mathbf{k}) = E_n + \frac{\hbar^2 k^2}{2m_0}, \quad c_{nm}(\mathbf{k}) = \delta_{nm} \tag{10}$$

The second-order approximation is thus given by

$$E_n(\mathbf{k}) = E_n + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar}{m_0} \sum_{\ell \neq n} \frac{|\langle \ell | \mathbf{k} \cdot \mathbf{p} | n \rangle|^2}{E_n - E_\ell} \tag{11}$$

The result can be expressed in terms of an effective mass m^* :

$$E_n(\mathbf{k}) = E_n + \sum_{i,j} \frac{\hbar^2}{2m_{ij}^*} k_i k_j \tag{12}$$

where $i, j = x, y, z$, and

$$\frac{m_0}{m_{ij}^*} = \delta_{ij} + \frac{2}{m_0} \sum_{\ell \neq n} \frac{\langle n | p_i | \ell \rangle \langle \ell | p_j | n \rangle}{E_n - E_\ell} \tag{13}$$

Note here that the effective mass can be anisotropic which can be utilized for optical coupling in quantum well photodetection (Xu et al. 1994).

It can be easily seen that a narrow bandgap, which leads to two states being close to each other, gives a small effective mass which agrees very well with the experimental data which indicates that InSb has both the smallest bandgap and the lowest effective mass. By a series of experiments the parameters for the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian can be determined, and the final result is a matrix with many material-dependent parameters and no arbitrary parameters to adjust. Details of the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian for III–V and II–VI semiconductor materials are well documented in reviews (Madelung 1991, 1992; Vurgaftman et al. 2001).

Excitons in Nanostructures

After reviewing some solid state theory, we are now capable of modeling excitons and exciton-polaritons in semiconductor nanostructures. We first study and model the exciton in general, then excitons in QD-based nanostructures, and finally applications of QD excitons and exciton-polaritons. Details of exciton theory can be found in Dimmock (1967) and Fu and Willander (1999).

As discussed in the previous section, the energy band structure of electrons in a semiconductor consists of energy bands separated by band gaps. At absolute zero temperature, a pure semiconductor is characterized by having only completely occupied and completely empty energy bands. The highest occupied bands are referred as VBs and the lowest unoccupied band is the CB. Let H_0 be the Hamiltonian of the nanostructure under investigation, then

$$H_0\psi_{vk_h}(\mathbf{r}_h) = E_v(\mathbf{k}_h)\psi_{vk_h}(\mathbf{r}_h)$$

denotes the VB state and

$$H_0\psi_{ck_e}(\mathbf{r}_e) = E_c(\mathbf{k}_e)\psi_{ck_e}(\mathbf{r}_e)$$

the CB state. We consider an optical absorption process in which an allowed electric dipole (not ultra intense excitation) transition raises an electron from a filled VB state to an empty CB state. The properties of a system with an electron missing from a VB state ψ_{vk_h} can be described by considering a single particle of positive charge in the state, that is, the VB hole. The exact expression of the interaction potential between the CB electron and VB hole is rather complicated. For commonly used semiconductor systems, and when the electron and hole are widely separated and moving relatively slowly with respect to one another, the Coulomb interaction between them is screened and given by Dimmock (1967)

$$V(\mathbf{r}_e - \mathbf{r}_h) = -\frac{e^2}{4\pi\epsilon_\infty |\mathbf{r}_e - \mathbf{r}_h|} \quad (14)$$

where ϵ_∞ is the high-frequency dielectric constant.

The exciton state is described by the two-particle wave function

$$\psi_{c\mathbf{k}_e}(\mathbf{r}_e) \psi_{v\mathbf{k}_h}(\mathbf{r}_h)$$

The total wave vector of the excited state is $\mathbf{K} = \mathbf{k}_e + \mathbf{k}_h$. Because of the Coulomb interaction, $\psi_{c\mathbf{k}_e}(\mathbf{r}_e) \psi_{v\mathbf{k}_h}(\mathbf{r}_h)$ does not represent an eigen state of the system. We approximate the eigen state of the exciton by a linear combination of electron-hole-states

$$\psi_{n\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h) = \sum_{c\mathbf{k}_e, v\mathbf{k}_h} A_{n\mathbf{K}}(c\mathbf{k}_e, v\mathbf{k}_h) \psi_{c\mathbf{k}_e}(\mathbf{r}_e) \psi_{v\mathbf{k}_h}(\mathbf{r}_h) \quad (15)$$

where n is the quantum index of the exciton state. By introducing the Fourier transform of $A_{n\mathbf{K}}(c\mathbf{k}_e, v\mathbf{k}_h)$,

$$\psi_{n\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h) = \frac{1}{\Omega} \sum_{\mathbf{k}_e, \mathbf{k}_h} e^{i\mathbf{k}_e \cdot \mathbf{r}_e} e^{i\mathbf{k}_h \cdot \mathbf{r}_h} A_{n\mathbf{K}}(c\mathbf{k}_e, v\mathbf{k}_h) \quad (16)$$

where Ω is the normalization volume, it is easy to show that $\psi_{n\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h)$ satisfies the differential equation

$$[E_c(-i\nabla_e) - E_v(-i\nabla_h) + V(\mathbf{r}_e - \mathbf{r}_h)] \psi_{n\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h) = E \psi_{n\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h) \quad (17)$$

where $E_c(-i\nabla_e)$ is the expression obtained by replacing \mathbf{k}_e in the power-series expansion of $E_c(\mathbf{k}_e)$ by $-i\nabla_e$. The result is valid provided $E_c(\mathbf{k}_e)$ and $E_v(\mathbf{k}_h)$ are analytic and can be expanded in powers of \mathbf{k}_e and \mathbf{k}_h . In the effective mass approximation of Eq. 12, the exciton Hamiltonian becomes

$$H_o = -\frac{\hbar^2 \nabla_e^2}{2m_e^*} - \frac{\hbar^2 \nabla_h^2}{2m_h^*} - \frac{e^2}{4\pi \epsilon_\infty |\mathbf{r}_e - \mathbf{r}_h|} \quad (18)$$

where m_e^* and m_h^* are effective masses of electron and hole.

It is easier to introduce the relative and centre-of-mass coordinates $\mathbf{r}_e - \mathbf{r}_h$ and $\mathbf{R} = (m_e^* \mathbf{r}_e + m_h^* \mathbf{r}_h) / (m_e^* + m_h^*)$ to solve the above exciton Hamiltonian

$$H_o = -\frac{\hbar^2 \nabla_{\mathbf{R}}^2}{2(m_e^* + m_h^*)} - \frac{\hbar^2 \nabla_{\mathbf{r}_e - \mathbf{r}_h}^2}{2m_h^*} - \frac{e^2}{4\pi \epsilon_\infty |\mathbf{r}_e - \mathbf{r}_h|} \quad (19)$$

where $1/m_r^* = 1/m_e^* + 1/m_h^*$ is the reduced effective mass.

For three-dimensionally extended bulk material, the part of the wave function in \mathbf{R} contains a factor of $e^{i\mathbf{K} \cdot \mathbf{R}}$, the part in the relative coordinates contains a hydrogenic $\psi_{n\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h) = \psi_{n\mathbf{K}}(\mathbf{r}_e - \mathbf{r}_h)$ so that the exciton wave function and its energy are

$$e^{i\mathbf{K}\cdot\mathbf{R}}\psi_{n\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h), \quad E_{n\mathbf{K}} = \frac{\hbar^2 K^2}{2(m_e^* + m_h^*)} - \frac{m_r^* e^4}{2\epsilon_\infty^2 n^2}, \quad (20)$$

and the wave function of the ground exciton state ($n = 1$) corresponding to the last two terms of the Hamiltonian Eq. 19 is

$$\psi_{1\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h) = \frac{1}{\sqrt{\pi a_{\text{Br}}^3}} e^{-|\mathbf{r}_e - \mathbf{r}_h|/a_{\text{Br}}} \quad (21)$$

where $a_{\text{Br}} = 4\pi\epsilon_\infty\hbar^2/m_r^*e^2$ is the Bohr radius of the exciton in the semiconductor material. In three-dimensionally confined nanostructures, the center of mass of the exciton is not mobile so that we simply let $\mathbf{K} = 0$ in Eqs. 20 and 21.

We now study the effect of excitons on the optical transition. The probability that the system makes an electronic transition from the ground state ψ_0 to an excited (exciton) state $\psi_{n\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h)$ is proportional to

$$|\langle \psi_{n\mathbf{K}} | e^{i\mathbf{q}\cdot\mathbf{r}} \mathbf{a} \cdot \mathbf{p} | \psi_0 \rangle|^2 \quad (22)$$

where p is the momentum operator, q and a are, respectively, the wave vector and polarization vector of the photon. By Eq. 15 we obtain

$$\langle \psi_{n\mathbf{K}} | \mathbf{p} | \psi_0 \rangle = \sum_{c\mathbf{k}_e, v\mathbf{k}_h} A_{n\mathbf{K}}(c\mathbf{k}_e, v\mathbf{k}_h) \langle \psi_{c\mathbf{k}_e} | \mathbf{p} | \theta \psi_{v\mathbf{k}_h} \rangle \quad (23)$$

where θ is the time-reversal operator and $\theta\psi_{v\mathbf{k}_h}$ is the unoccupied VB state out of which the electron was excited. Denoting $\theta\psi_{v\mathbf{k}_h} = \psi_{v-\mathbf{k}_h}$ as the hole state and making use of the negligible q (the photon wave vector is negligibly small so that $\mathbf{K} = 0$) so that $\mathbf{k}_e = -\mathbf{k}_h \equiv \mathbf{k}$,

$$\langle \psi_{n0} | \mathbf{p} | \psi_0 \rangle = \sum_{\mathbf{k}} A_{n0}(c\mathbf{k}, v-\mathbf{k}) \langle \psi_{c\mathbf{k}} | \mathbf{p} | \psi_{v\mathbf{k}} \rangle \quad (24)$$

where we have neglected the contribution from all other bands except the VB and CB in the vicinity of the Fermi level.

Under normal device operation conditions, $\langle \psi_{c\mathbf{k}} | \mathbf{p} | \psi_{v\mathbf{k}} \rangle \approx \mathbf{p}_{cv}$ which is independent of \mathbf{k} over the range of \mathbf{k} involved,

$$\langle \psi_{n0} | \mathbf{p} | \psi_0 \rangle = \mathbf{p}_{cv} \sum_{\mathbf{k}} A_{n0}(c\mathbf{k} - v\mathbf{k}) \quad (25)$$

By Eq. 16 the transition probability is therefore proportional to

$$|\langle \psi_{n0} | \mathbf{p} | \psi_0 \rangle|^2 = \mathbf{p}_{cv}^2 |\psi_{n0}(\mathbf{r}_e, \mathbf{r}_h)|_{\mathbf{r}_e=\mathbf{r}_h}^2 \quad (26)$$

Table 2 Excitons in common semiconductor (Fu et al. 2000; Madelung 1991, 1992) (low temperature for IV and III–V). denotes the band gap between VB and CB

	Si	GaAs	AlAs	InAs	CdSe	ZnO
m_h^* (heavy hole)	0.537	0.51	0.409	0.35	0.45/1.0	0.59
m_e^* (electron)	1.026	0.067	0.71	0.0239	0.13	0.28
E_g [eV]	1.170	1.519	2.229	0.418	1.842	3.435
a_{Br} [nm]	15.2	11.6	8.6	38.1	5.4	1.8
E_{10} [meV]	14.3	4.2	20.0		13.2	60.0

Note that the optical transition is also constrained to $\mathbf{K} = 0$ since the photon wave vector is negligibly small. For spherical state $\psi_{n0}(0)$ is nonzero only for s-type states of Eq. 21 (the reader can refer to the hydrogen atom in many quantum mechanics textbooks) and $|\psi_{n0}(0)|^2 \propto 1/n^3$ so that

$$|\langle \psi_{n0} | \mathbf{p} | \psi_0 \rangle|^2 \propto \frac{1}{n^3} \quad (27)$$

which shows that only the low-energy exciton states (small n) can be probed optically for bulk materials.

The exciton binding energy E_{10} in Eq. 20 for the ground exciton state ($n = 1$) in commonly used bulk semiconductors is less than 10 meV so that their room temperature luminescence is very weak (exciton thermally dissolved so that n is large). Moreover, exciton Bohr radii, that is, the effective spatial extensions of the excitons, in bulk materials are also quite extended, see Table 2.

In high quality semiconductor nanostructures, quantum size effects prevail and subband formation strongly influences the exciton as the electron and hole are forced to stay in a very confined spatial area. Coulomb interaction between the electron and hole, that is, the exciton binding energy, becomes much increased. Even at room temperature the sharp exciton lines can be detected in the absorption spectra of AlGaAs/GaAs multiple QWs (Miller et al. 1982), which can hardly be observed in bulk GaAs samples.

One of the reasons for the fast development of ZnO and related materials is their large excitonic binding of about 60 meV which binds the electron and hole strongly in bulk ZnO even at room temperature (Sun et al. 2002). A photonic switch operating by controlling exciton excitation in ZnO QWs via an optical near field has been reported, where ZnO QWs were in the form disks with a disk diameter of 80 nm and the heights of 2 nm (in a structure composed of one single-QW and three double-QW) and 3.5 nm in the one double-QW case (Yatsui et al. 2007, 2008). Another development based on the large exciton binding energy of ZnO was the room-temperature polariton laser (Zamfirescu et al. 2002).

Equation 27 tells us that the major reason for studying and developing semiconductor nanostructures for optical applications refers to the forced quantum confinement. The exciton will not dissolve even when the exciton binding is comparable or smaller than the thermal energy. Light–matter interaction in the

forms of optical exciton generation and radiative recombination is much enhanced, resulting in many novel nanostructure-based optoelectronic applications.

Polarization and Optical Properties of Exciton-Polaritons

In this section, we study the optical properties of excitons in nanostructures. We start with the detailed analysis of optical absorption in which an allowed electric dipole transition creates an exciton ψ_{nK} from a filled VB and in this scheme the wave function of the initial state (the filled VB) is simply unity denoted by ψ_0 in the formalism of second quantization. Assume that our electron-hole system is initially in its ground state ψ_0 for time $t < 0$. We switch on an external radiation of $\mathbf{E}(\mathbf{r}, t)$ for $t \geq 0$; the first-order perturbation Hamiltonian is

$$V = \int \mathbf{d}(\mathbf{r}) \cdot \mathbf{E}(\mathbf{r}, t) d\mathbf{r} \quad (28)$$

where $\mathbf{d}(\mathbf{r})$ is the dipole-moment operator given by,

$$\mathbf{d}(\mathbf{r}) = -e\mathbf{r}_e\delta(\mathbf{r} - \mathbf{r}_e) + e\mathbf{r}_h\delta(\mathbf{r} - \mathbf{r}_h) \quad (29)$$

where e is the charge unit. We denote the wave function of the electron-hole system as $|\mathbf{r}_e, \mathbf{r}_h, t\rangle$, its time-dependent Schrödinger equation is then

$$i\hbar \frac{\partial}{\partial t} |\mathbf{r}_e, \mathbf{r}_h, t\rangle = (H_0 + V) |\mathbf{r}_e, \mathbf{r}_h, t\rangle \quad (30)$$

where H_0 is given by Eq. 18.

By the first-order perturbation approximation where we consider only the exciton ground state ψ_0 to one excited excitonic state $\psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t)$ the time-dependent wave function

$$|\mathbf{r}_e, \mathbf{r}_h, t\rangle = |\psi_0(t)\rangle + c_{nK}(t) |\psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t)\rangle \quad (31)$$

Here we simply assume that $|c_{nK}(t)| \ll 1$. We will discuss this more closely in the coming sections.

We now formally denote both $\psi_0(t)$ and $\psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t)$ as eigen functions of H_0 such that

$$\begin{aligned} H_0\psi_0(t) &= E_0\psi_0(t) = \hbar\omega_0\psi_0(t) \\ H_0\psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) &= E_{nK}\psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) = \hbar\omega_{nK}\psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) \\ \psi_0(t) &= \psi_0 e^{-i\omega_0 t} \\ \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) &= \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h) e^{-i\omega_{nK} t} \end{aligned} \quad (32)$$

Substituting Eq. 31 into Eq. 30 leads to

$$\begin{aligned} & i\hbar \frac{\partial \psi_0(t)}{\partial t} + i\hbar \frac{\partial c_{nK}(t)}{\partial t} \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) + i\hbar c_{nK}(t) \frac{\partial \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t)}{\partial t} \\ & = H_0 \psi_0(t) + c_{nK}(t) H_0 \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) + V \psi_0(t) + c_{nK}(t) V \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) \end{aligned}$$

By Eqs. 32,

$$\begin{aligned} & \hbar \omega_0 \psi_0(t) + i\hbar \frac{\partial c_{nK}(t)}{\partial t} \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) + \hbar \omega_{nK} c_{nK}(t) \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) \\ & = E_0 \psi_0(t) + c_{nK}(t) E_{nK} \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) + V \psi_0(t) + c_{nK}(t) V \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) \end{aligned}$$

which is

$$i\hbar \frac{\partial c_{nK}(t)}{\partial t} \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) = V \psi_0(t) + c_{nK}(t) V \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) \quad (33)$$

Multiplying the above equation by $\langle \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) |$ we come to

$$\frac{\partial c_{nK}(t)}{\partial t} = \frac{1}{i\hbar} \left\langle \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h, t) \left| \int \mathbf{d}(\mathbf{r}) \cdot \mathbf{E}(\mathbf{r}, t) \, d\mathbf{r} \right| \psi_0(t) \right\rangle \quad (34)$$

It can be shown (Fu and Willander 1999)

$$\begin{aligned} & \left\langle \psi_{nK}(\mathbf{r}_e, \mathbf{r}_h) \left| \int \mathbf{d} \cdot \mathbf{E}(\mathbf{r}, t) \, d\mathbf{r} \right| \psi_0(\mathbf{r}_e, \mathbf{r}_h) \right\rangle \\ & = \frac{e}{\omega_{nK} m_0} \int \psi_{nK}^*(\mathbf{r}, \mathbf{r}) \mathbf{p}_{cv} \cdot \mathbf{E}(\mathbf{r}, t) \, d\mathbf{r} \end{aligned} \quad (35)$$

where $\psi_{nK}(\mathbf{r}, \mathbf{r})$ is obtained from $\psi_{nK}(\mathbf{r}_e, \mathbf{r}_h)$ by setting $\mathbf{r}_e = \mathbf{r}_h = \mathbf{r}$. Thus,

$$\frac{dc_{nK}(t)}{dt} - \frac{1}{i\hbar} \frac{e}{\omega_{nK} m_0} \int \psi_{nK}^*(\mathbf{r}, \mathbf{r}) \mathbf{p}_{cv} \cdot \mathbf{E}(\mathbf{r}, t) \, d\mathbf{r} e^{i(\omega_{nK} - \omega_0)t} \quad (36)$$

Note that optical couplings among excitonic states are zero because of the quantum selection rules. As well, only a limited number of excitonic states have nonzero optical couplings with the ground state.

In reality $\omega_0 = 0$ since this is the reference state where the VB is completely filled and the CB is empty. Furthermore, we can very well describe the excitation radiation as $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(\mathbf{r}, \omega) e^{-i\omega t}$, so that integrating the above equation we obtain immediately

$$c_{nK}(t) = \frac{e^{i(\omega_{nK} - \omega)t}}{\hbar(\omega_{nK} - \omega + i\gamma)} \frac{e}{\omega_{nK} m_0} \int \psi_{nK}^*(\mathbf{r}, \mathbf{r}) \mathbf{p}_{cv} \cdot \mathbf{E}(\mathbf{r}, \omega) \, d\mathbf{r} \quad (37)$$

Before we move on, let us discuss the spatial distribution of the exciton wave function. Excitons are normally categorized into two types. The Frenkel exciton is an excited electronic state where the electron and the hole are situated in

the same molecule or atom. Because of their small radius, the Frenkel excitons usually are considered to be local. The large-radius Wannier–Mott exciton in bulk semiconductors is relatively weakly bound due to a typically large exciton Bohr radius. A hybrid state of Wannier–Mott exciton and Frenkel exciton in different heterostructure configurations involving QDs was found to exist at the interfaces of the QDs and the surrounding organic medium with complimentary properties of both kinds of excitons (Birman and Huang 2007).

In heterostructures both the CB electrons and the VB holes will experience extra potential energies such as band offsets. For low-energy Frenkel excitons in a QD, the electrons and holes are confined within the same QD volume (type-I exciton) so that we will have a common potential energy $V(|\mathbf{R} - \mathbf{a}|)$ that confines the exciton, where \mathbf{R} is the center of mass of the exciton (see Eq. 19), \mathbf{a} denotes the center of the QD. For common semiconductor QDs, the radius of the QD, R_{QD} , is in the order of the exciton Bohr radius a_{Br} so that one may neglect the free motion of the center of mass of the electron-hole pair in the QD, that is, $\mathbf{K} = 0$ in Eq. 20. For the ground state of the exciton we can approximate $V(|\mathbf{R} - \mathbf{a}|) = 0$ when $|\mathbf{R} - \mathbf{a}| < R_{\text{QD}}$ and $V(|\mathbf{R} - \mathbf{a}|) = \infty$ elsewhere so that the wave function corresponding to the motion of the center of mass becomes (Gasiorowicz 1996)

$$\frac{1}{\sqrt{2\pi R_{\text{QD}} |\mathbf{R} - \mathbf{a}|}} \sin\left(\frac{\pi |\mathbf{R} - \mathbf{a}|}{R_{\text{QD}}}\right) \quad (38)$$

for $|\mathbf{r} - \mathbf{a}| \leq R_{\text{QD}}$. The wave function is zero elsewhere. We thus finally obtain the total wave function of the exciton ground state inside the QD

$$\frac{1}{\sqrt{2\pi R_{\text{QD}} |\mathbf{R} - \mathbf{a}|}} \sin\left(\frac{\pi |\mathbf{R} - \mathbf{a}|}{R_{\text{QD}}}\right) \frac{1}{\sqrt{\pi a_{\text{Br}}^3}} e^{-|r_e - r_h|/a_{\text{Br}}} u_c(\mathbf{r}_e) u_v(\mathbf{r}_h) \quad (39)$$

Here Bloch functions $u_c(\mathbf{r}_e)$ and $u_v(\mathbf{r}_h)$ at CB and VB bandedges are included, see Eqs. 1 and 6.

The polarization of the excited exciton in a QD centered at \mathbf{a} is

$$\mathbf{P}_a(\mathbf{r}, t) = \sum_{n\mathbf{K}} \langle \psi_{n\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h, t) | \mathbf{d}(\mathbf{r}) | \psi_0(\mathbf{r}_e, \mathbf{r}_h) \rangle c_{n\mathbf{K}}^*(t) c_0(t) + \text{c.c.} \quad (40)$$

where \mathbf{d} is the dipole of the exciton given by Eq. 29. Since the wave function Eq. 39 of the exciton is confined and real, the polarization of the exciton in the QD excited by an external electromagnetic field $\mathbf{E}(\mathbf{r}, t)$ is (Fu et al. 1997, 2000)

$$\begin{aligned} \mathbf{P}_a(\mathbf{r}, t) = & \sum_{n\mathbf{K}} \frac{e^2 \mathbf{p}_{cv} e^{-i\omega t}}{\hbar (\omega_{n\mathbf{K}} - \omega + i\gamma) \omega_{n\mathbf{K}}^2 m_0^2} \\ & \times \psi_{n\mathbf{K}}(\mathbf{r}, \mathbf{r}) \int \psi_{n\mathbf{K}}(\mathbf{r}', \mathbf{r}') \mathbf{p}_{cv} \cdot \mathbf{E}(\mathbf{r}', \omega) d\mathbf{r}' \end{aligned} \quad (41)$$

By assuming small spatial variation of the excitation field within the QD, an effective permittivity is defined for the QD exciton polariton by writing

$$\mathbf{D}(\mathbf{r}, \omega) = \epsilon_{\infty} \mathbf{E}(\mathbf{r}, \omega) + \mathbf{P}(\mathbf{r}, \omega) = \epsilon_{\text{QD}}(\mathbf{r}, \omega) \mathbf{E}(\mathbf{r}, \omega) \quad (42)$$

where ϵ_{∞} is the background dielectric constant of the QD material, and

$$\epsilon_{\text{QD}}(\mathbf{r}, \omega) = \epsilon_{\infty} \left(1 + \sum_{n\mathbf{K}} \frac{2\omega_{\text{LT}}}{\omega_{n\mathbf{K}} - \omega + i\gamma} \frac{\sin \alpha_{n\mathbf{K}}}{\alpha_{n\mathbf{K}}} \right) \quad (43)$$

where $\alpha_{n\mathbf{K}} = \pi |\mathbf{r} - \mathbf{a}| / R_{\text{QD}}$. $\epsilon_{\infty} \omega_{\text{LT}} a_{\text{Br}}^3 = e^2 p_{cv}^2 / \pi \hbar \omega_{n\mathbf{K}}^2 m_0^2$. ω_{LT} is normally referred to be the exciton longitudinal-transverse splitting. The above equation was obtained under the perturbation approximation and thus can also be accounted at a semiclassical level by adding a Lorentz-oscillator term to the dielectric function (Andreani et al. 2005).

Notice that $\epsilon_{\text{QD}}(\mathbf{r}, \omega)$ in the above equation is position-dependent. Making an average over the QD volume results in an effective dielectric constant

$$\epsilon_{\text{QD}}(\omega) = \epsilon_{\infty} \left[1 + \sum_{n\mathbf{K}} \frac{6\omega_{\text{LT}}}{\pi^2 (\omega_{n\mathbf{K}} - \omega + i\gamma)} \right] \quad (44)$$

For a normal system without any extra external modifications except the incident electromagnetic field of $\mathbf{E}(\mathbf{r}, t)$ under investigation, the occupation of excited excion state, $|c_{n\mathbf{K}}(t)|^2$, is small (the first-order perturbation condition). Moreover, this occupation of the excited state is also much smaller than the ground state, see Eq. 31, so that an optical absorption dominates. This means that the optical field is absorbed during the exciton excitation following its propagation through the QD. The situation is normally referred to as optical loss, for example, in photodetectors.

Optical gain is achieved in light-emitting and laser devices by population inversion. To achieve optical gain in the QD, we increase the occupation of the excion state. Here we consider such a pumping process that the excion population of the excited excion state $c_{n\mathbf{K}}^{(0)}$ is finite at time $t = 0$ so the time-dependent wavefunction becomes

$$|\mathbf{r}_e, \mathbf{r}_h, t\rangle = \left[c_0^{(0)} + c_0(t) \right] |\psi_0(\mathbf{r}_e, \mathbf{r}_h)\rangle + \sum_{n\mathbf{K}} \left[c_{n\mathbf{K}}^{(0)} + c_{n\mathbf{K}}(t) \right] |\psi_{n\mathbf{K}}(\mathbf{r}_e, \mathbf{r}_h, t)\rangle \quad (45)$$

due to the external electromagnetic field $\mathbf{E}(\mathbf{r}, t)$. We can obtain (Fu et al. 2008)

$$\epsilon_{\text{QD}}(\omega) = \epsilon_{\infty} \left[1 + \sum_{n\mathbf{K}} \frac{6 \left(2 |c_0^{(0)}|^2 - 1 \right) \omega_{\text{LT}}}{\pi^2 (\omega_{n\mathbf{K}} - \omega + i\gamma)} \right] \quad (46)$$

Let us write $\varepsilon = \varepsilon' + i\varepsilon''$,

$$\begin{aligned} \varepsilon'_{\text{QD}}(\omega) &= \varepsilon_{\infty} \left[1 + \sum_{n\mathbf{K}} \frac{6 \left(2|c_0^{(0)}|^2 - 1 \right) (\omega_{n\mathbf{K}} - \omega) \omega_{\text{LT}}}{\pi^2 \left[(\omega_{n\mathbf{K}} - \omega)^2 + \gamma^2 \right]} \right] \\ \varepsilon''_{\text{QD}}(\omega) &= \varepsilon_{\infty} \left[1 + \sum_{n\mathbf{K}} \frac{6 \left(2|c_0^{(0)}|^2 - 1 \right) \gamma \omega_{\text{LT}}}{\pi^2 \left[(\omega_{n\mathbf{K}} - \omega)^2 + \gamma^2 \right]} \right] \end{aligned} \tag{47}$$

Since ε'' represents the energy loss of the incident electromagnetic field, we observe that the QD is lossy when $|c_0^{(0)}|^2 < 1/2$, it becomes transparent when $|c_0^{(0)}|^2 = 1/2$, and optical gain will occur when $|c_0^{(0)}|^2 > 1/2$. At the same time, the polarization changes sign.

Figure 3 shows the calculated dielectric constant for one QD before and after some fraction of the exciton states becomes populated. In this example, II–VI PbSe/ZnSe QD in a CdSe background is assumed and $\hbar\omega_{n\mathbf{K}} = 1.5\text{eV}$. In Fig. 3 and throughout this chapter we write $\varepsilon = \varepsilon' + i\varepsilon''$ where ε' and ε'' are expressed in units of ε_0 . Here we observe the possibility of finding low-loss negative dielectric constant.

As shown in Fig. 3, a huge effective permittivity is expected in the vicinity of $\omega_{n\mathbf{K}}$, with which the energy dispersions of photonic crystals based on QD arrays were derived (Fu et al. 2000) to understand photonic band gaps observed in CdS

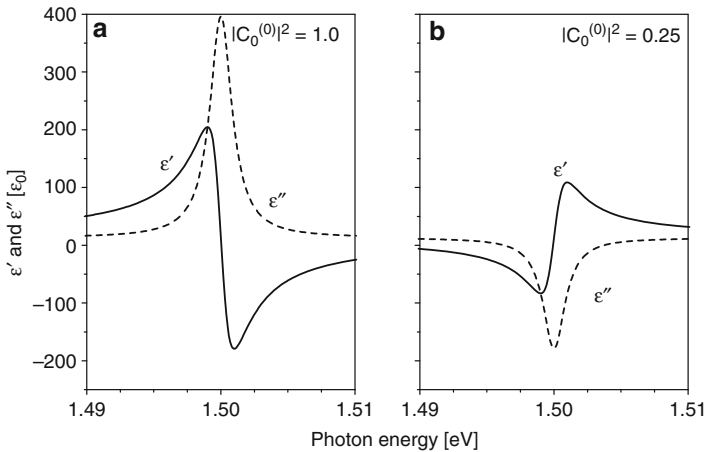


Fig. 3 Dielectric coefficients of the QD before (a) and after (b) the excited exciton state in the QD becomes populated $\varepsilon_{\infty} = 12.8$, $\hbar\omega_{n\mathbf{K}} = 1.5\text{eV}$, $\hbar\omega_{\text{LT}} = 5\text{meV}$, and $\hbar\gamma = 1\text{meV}$ for typical II–VI semiconductor QDs, for example, colloidal PbSe/ZnSe QD

QDs embedded in fcc porous silica matrices (Vlasov et al. 1997). A close-packed 3D array of spherical CuCl QDs in air was shown by extended Maxwell-Garnett theory and rigorous layer-multiple-scattering method to have a negative refractive index within the region of the excitonic resonance (Yannopapas 2007). The structure was proposed for subwavelength imaging (the loss is however formidably high) (Yannopapas 2008).

Exciton-Polariton Photonic Crystals

As shown by Eq. 43, the effective dielectric coefficient of the QD is strongly modified when the exciton state in the QD is resonantly excited by an incident electromagnetic field. The contrast between this exciton effective dielectric coefficient and the background material can thus be utilized to construct a photonic crystal when the QDs are positioned in space in a periodic manner. Equation 43 was coupled to the electromagnetic field to study the optical dispersion of exciton-polariton crystals (Fu et al. 2000; Ivchenko et al. 2000) stimulated by the experimental works of fcc structured silica opals (Vlasov et al. 1997).

Maxwell equations describing the incident electromagnetic field in the QD photonic crystal, free of charges and free of drift-diffusion current, are written as

$$\begin{aligned}\nabla \times [\nabla \times \mathbf{E}(\mathbf{r})] &= \mu_0 \omega^2 \mathbf{D}(\mathbf{r}) \\ \nabla \cdot \mathbf{D}(\mathbf{r}) &= 0\end{aligned}\quad (48)$$

in the MKS unit system, where $\mathbf{E}(\mathbf{r})$ is the electric field, $\mathbf{D}(\mathbf{r})$ is the displacement vector, and ω is the angular frequency of the incident electromagnetic field. Here we have assumed that the QD consists of uniform isotropic linear media, for which $\mathbf{D} = \epsilon_\infty \mathbf{E}$ and $\mathbf{B} = \mu_0 \mathbf{H}$. \mathbf{H} is the magnetizing field and \mathbf{B} is the magnetic field. μ_0 is the magnetic constant (permeability of free space). The nonlocal material equation relating $\mathbf{D}(\mathbf{r})$ and $\mathbf{E}(\mathbf{r})$ is (Fu et al. 1997)

$$\mathbf{D}(\mathbf{r}) = \epsilon_\infty \mathbf{E}(\mathbf{r}) + \sum_a \mathbf{P}_a(\mathbf{r}) \quad (49)$$

where $\mathbf{P}_a(\mathbf{r})$ is the polarization contribution from the excited exciton in a QD centered at \mathbf{a} , that is, Eq. 41. We have neglected the overlap of exciton envelope functions centered at different QDs so that excitons excited in different QDs are assumed to be coupled only via the electromagnetic field.

By denoting $\mathbf{P} = \sum_a \mathbf{P}_a(\mathbf{r})$, it follows from the second Eq. 48 and Eq. 49 that

$$\nabla \cdot \mathbf{E}(\mathbf{r}) = -\frac{1}{\epsilon} \nabla \cdot \mathbf{P}(\mathbf{r}) \quad (50)$$

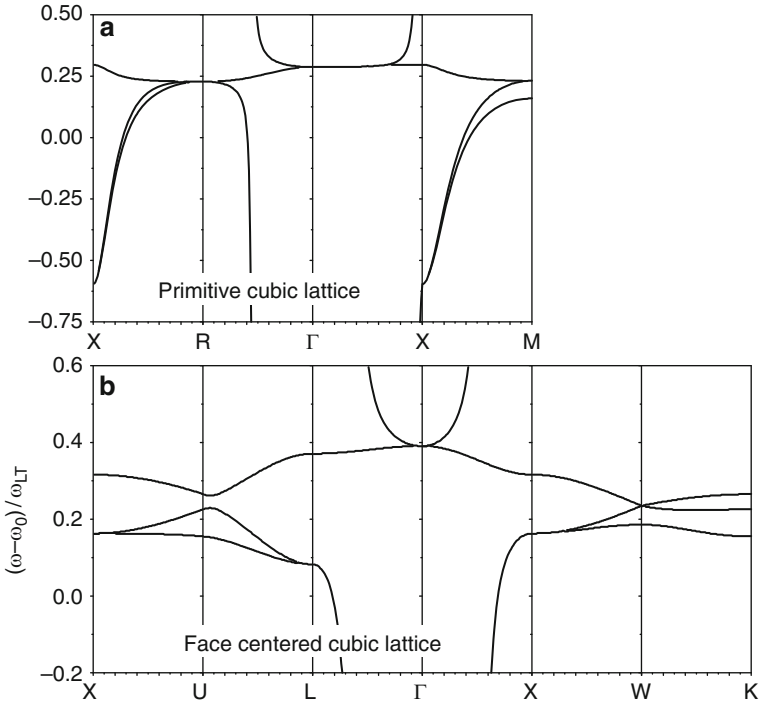


Fig. 4 Photonic dispersion relation of QD photonic crystals. (a) Primitive cubic lattice. (b) Face-centered cubic lattice

so that the first Eq. 48 can be rewritten as

$$\nabla^2 \mathbf{E}(\mathbf{r}) + k^2 \mathbf{E}(\mathbf{r}) = -\frac{k_0^2}{\varepsilon_0} \left\{ \mathbf{P}(\mathbf{r}) + \frac{1}{k^2} \nabla [\nabla \cdot \mathbf{P}(\mathbf{r})] \right\} \quad (51)$$

where $k_0 = \omega/c$, $k = k_0 n = \omega n/c$, and $n = \sqrt{\varepsilon_\infty / \varepsilon_0}$.

Bloch-like solutions of Eq. 51 satisfying

$$\mathbf{E}_q(\mathbf{r} + \mathbf{a}) = e^{iq \cdot \mathbf{a}} \mathbf{E}_q(\mathbf{r}) \quad (52)$$

is the photonic band structure. Figure 4 shows the dispersion relationships of primitive-cubic and fcc QD photonic crystals (Fu et al. 2000). An overall bandgap, most prominent along the $\Lambda(\Gamma-L)$ line (the [111] direction) is observed in the fcc lattice (Zeng et al. 2006a).

Photonic Dispersion of QD Dimer Systems

As demonstrated by Eqs. 47 and 3, ε'_{QD} can be adjusted to be positively large as well as negative by tuning the initial occupation condition of the exciton state in

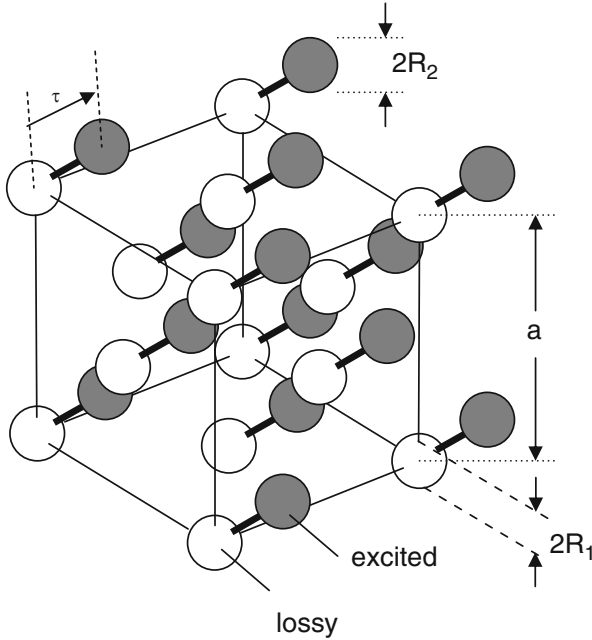


Fig. 5 Schematic QD dimer system in a face-centered cubic (fcc) lattice structure. R_1 and R_2 denote the radii of the QDs and a the lattice constant of the fcc lattice

the QD. However, optical loss is not avoidable when a single type of QDs is used. To tackle the loss issue, we construct a QD dimer photonic crystal by positioning two types of spherical QDs, having radii R_1 and R_2 respectively, in a fcc lattice having a lattice constant a , see Fig. 5. The lossy QDs, that is, type-I QDs (exciton energies are $\hbar\omega_1$) occupy the normal fcc lattice sites, while the already excited QDs, type-II QDs ($\hbar\omega_2$) are displaced from the type-I QDs by τ . Let c_{ai} be the occupation of the ground state of QD i positioned at lattice site a , that is, the valence band is completely filled and the conduction band is completely empty, the contribution of an excited exciton state in a QD ai to the dielectric polarization is given by

$$P_{ai}(\mathbf{r}) = \frac{e^2 \mathbf{p}_{cv} (2|c_{ai}|^2 - 1)}{\hbar (\omega_{ai} - \omega + i\gamma) \omega_{ai}^2 m_0^2} \psi_{ai}^*(\mathbf{r}, \mathbf{r}) \int \psi_{ai}^*(\mathbf{r}', \mathbf{r}') \mathbf{p}_{cv} \cdot \mathbf{E}(\mathbf{r}') d\mathbf{r}' \tag{53}$$

by referring to Eqs. 41 and 46.

We seek again for Bloch-like solutions of Eq. 51 satisfying Eq. 52. We expand the vector function $E_q(\mathbf{r})$ in the Fourier series

$$\mathbf{E}_q(\mathbf{r}) = \sum_{\mathbf{g}} e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}} \mathbf{E}_{q+\mathbf{g}} \quad (54)$$

Where \mathbf{g} are the reciprocal lattice vectors. Denote $a_1 = \sqrt{2|c_1|^2 - 1}$, $a_2 = \sqrt{2|c_2|^2 - 1}$, $v_0 = a^3/4$ is the volume of the primitive unit cell of the fcc lattice, $\phi_i = a_i \psi_i$ and $t = \varepsilon \omega_{\text{LT}} \pi a_{\text{Br}}^3$, the excitonic polarization becomes

$$\mathbf{P}(\mathbf{r}) = \sum_{ai} \frac{t}{\omega_i - \omega + i\gamma} \phi_{ai}(\mathbf{r}, \mathbf{r}) \int \phi_{ai}^*(\mathbf{r}', \mathbf{r}') \mathbf{E}(\mathbf{r}') d\mathbf{r}' \quad (55)$$

Note that the excitonic wave function is real so that we have dropped off the complex conjugation of the wave function in the integrals in Eq. 53. The two integrals in the above equation can be transformed into

$$\begin{aligned} \int \phi_{1,a}(\mathbf{r}) \mathbf{E}(\mathbf{r}) d\mathbf{r} &= e^{i\mathbf{q}\cdot\mathbf{a}} \sum_{\mathbf{g}} I_{1,\mathbf{q}+\mathbf{g}} \mathbf{E}_{q+\mathbf{g}} \equiv e^{i\mathbf{q}\cdot\mathbf{a}} \Lambda_1 \\ \int \phi_{2,a}(\mathbf{r}) \mathbf{E}(\mathbf{r}) d\mathbf{r} &= e^{i\mathbf{q}\cdot\mathbf{a}} \sum_{\mathbf{g}} I_{2,\mathbf{q}+\mathbf{g}} \mathbf{E}_{q+\mathbf{g}} \equiv e^{i\mathbf{q}\cdot\mathbf{a}} \Lambda_2 \end{aligned} \quad (56)$$

where $I_{1,\mathbf{q}+\mathbf{g}} = \int \phi_{1,0}(\mathbf{r}) e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}} d\mathbf{r}$, $I_{2,\mathbf{q}+\mathbf{g}} = \int \phi_{2,0}(\mathbf{r}) e^{i(\mathbf{q}+\mathbf{g})\cdot(\mathbf{r}+\boldsymbol{\tau})} d\mathbf{r}$. The sums $\sum_a \phi_{1,a}(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{a}}$ and $\sum_{a+\boldsymbol{\tau}} \phi_{2,a+\boldsymbol{\tau}}(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{a}}$ satisfy the translational symmetry and can be presented as

$$\begin{aligned} \sum_a \phi_{1,a}(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{a}} &= \sum_{\mathbf{g}} e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}} \frac{I_{1,\mathbf{q}+\mathbf{g}}^*}{v_0} \\ \sum_{a+\boldsymbol{\tau}} \phi_{2,a+\boldsymbol{\tau}}(\mathbf{r}) e^{i\mathbf{q}\cdot\mathbf{a}} &= \sum_{\mathbf{g}} e^{i(\mathbf{q}+\mathbf{g})\cdot\mathbf{r}} \frac{I_{2,\mathbf{q}+\mathbf{g}}^*}{v_0} \end{aligned} \quad (57)$$

The linear equations for the space harmonics $\mathbf{E}_{q+\mathbf{g}}$ can thus be written

$$\begin{aligned} &(|\mathbf{q} + \mathbf{g}|^2 - k^2) \mathbf{E}_{q+\mathbf{g}} \\ &= \frac{k_0^2 t}{v_0} \left[1 - \frac{1}{k^2} (\mathbf{q} + \mathbf{g})^2 \right] \left[\frac{I_{1,\mathbf{q}+\mathbf{g}}^*}{\omega_1 - \omega + i\gamma} \Lambda_1 + \frac{I_{2,\mathbf{q}+\mathbf{g}}^*}{\omega_2 - \omega + i\gamma} \Lambda_2 \right] \end{aligned} \quad (58)$$

Let $S(\mathbf{Q})_{\alpha\beta} = \delta_{\alpha\beta} - Q_\alpha Q_\beta / k^2$, $\alpha, \beta = x, y, z$, $\delta_{\alpha\beta}$ is the Kronecker symbol. Dividing Eq. 58 by $(|\mathbf{q} + \mathbf{g}|^2 - k^2)$, multiplying it by $I_{1,\mathbf{q}+\mathbf{g}}$, and summing over \mathbf{g} :

$$\begin{aligned}
\sum_g I_{1,q+g} \mathbf{E}_{q+g} &= \Lambda_1 \\
&= \frac{k_0^2 t}{v_0} \sum_g \frac{\widehat{S}(\mathbf{q} + \mathbf{g})}{|\mathbf{q} + \mathbf{g}|^2 - k^2} \left[\frac{|I_{1,q+g}|^2}{\omega_1 - \omega + i\gamma} \Lambda_1 + \frac{I_{1,q+g} I_{2,q+g}^*}{\omega_2 - \omega + i\gamma} \Lambda_2 \right] \\
&= \widehat{M}_1(\omega, \mathbf{q}) \Lambda_1 + \widehat{M}_2(\omega, \mathbf{q}) \Lambda_2
\end{aligned} \tag{59}$$

where

$$\begin{aligned}
\widehat{M}_1(\omega, \mathbf{q}) &= \frac{k_0^2 r}{(\omega_1 - \omega + i\gamma) v_0} \sum_g \frac{\widehat{S}(\mathbf{q} + \mathbf{g})}{|\mathbf{q} + \mathbf{g}|^2 - k^2} |I_{1,q+g}|^2 \\
\widehat{M}_2(\omega, \mathbf{q}) &= \frac{k_0^2 r}{(\omega_1 - \omega + i\gamma) v_0} \sum_g \frac{\widehat{S}(\mathbf{q} + \mathbf{g})}{|\mathbf{q} + \mathbf{g}|^2 - k^2} I_{1,q+g} I_{2,q+g}^*
\end{aligned} \tag{60}$$

Similarly, dividing Eq. 58 by $(|\mathbf{q} + \mathbf{g}|^2 - k^2)$, multiplying it by $I_{2,q+g}$, and summing over g ,

$$\begin{aligned}
\sum_g I_{2,q+g} \mathbf{E}_{q+g} &= \Lambda_2 \\
&= \frac{k_0^2 t}{v_0} \sum_g \frac{\widehat{S}(\mathbf{q} + \mathbf{g})}{|\mathbf{q} + \mathbf{g}|^2 - k^2} \left[\frac{I_{2,q+g} I_{1,q+g}^*}{\omega_1 - \omega + i\gamma} \Lambda_1 + \frac{|I_{2,q+g}|^2}{\omega_2 - \omega + i\gamma} \Lambda_2 \right] \\
&= \widehat{M}_3(\omega, \mathbf{q}) \Lambda_1 + \widehat{M}_4(\omega, \mathbf{q}) \Lambda_2
\end{aligned} \tag{61}$$

where

$$\begin{aligned}
\widehat{M}_3(\omega, \mathbf{q}) &= \frac{k_0^2 t}{(\omega_1 - \omega + i\gamma) v_0} \sum_g \frac{\widehat{S}(\mathbf{q} + \mathbf{g})}{|\mathbf{q} + \mathbf{g}|^2 - k^2} I_{2,q+g} I_{1,q+g}^* \\
\widehat{M}_4(\omega, \mathbf{q}) &= \frac{k_0^2 t}{(\omega_1 - \omega + i\gamma) v_0} \sum_g \frac{\widehat{S}(\mathbf{q} + \mathbf{g})}{|\mathbf{q} + \mathbf{g}|^2 - k^2} |I_{2,q+g}|^2
\end{aligned} \tag{62}$$

We arrive at the vector equations

$$\begin{aligned}
\left[\mathbf{I} - \widehat{M}_1(\omega, \mathbf{q}) \right] \Lambda_1 &= \widehat{M}_2(\omega, \mathbf{q}) \Lambda_2 \\
\left[\mathbf{I} - \widehat{M}_4(\omega, \mathbf{q}) \right] \Lambda_2 &= \widehat{M}_3(\omega, \mathbf{q}) \Lambda_1
\end{aligned} \tag{63}$$

where \mathbf{I} is a 3×3 unit matrix. We can rewrite the above equations as

$$D \begin{pmatrix} 1 - M_{1,11} & -M_{1,12} & -M_{1,13} & -M_{2,11} & -M_{2,12} & -M_{2,13} \\ -M_{1,21} & 1 - M_{1,22} & -M_{1,23} & -M_{2,21} & -M_{2,22} & -M_{2,23} \\ -M_{1,31} & -M_{1,32} & 1 - M_{1,33} & -M_{2,31} & -M_{2,32} & -M_{2,33} \\ M_{3,11} & M_{3,12} & M_{3,13} & M_{4,11} - 1 & M_{4,12} & M_{4,13} \\ M_{3,21} & M_{3,22} & M_{3,23} & M_{4,21} & M_{4,22} - 1 & M_{4,23} \\ M_{3,31} & M_{3,32} & M_{3,33} & M_{4,31} & M_{4,32} & M_{4,33} - 1 \end{pmatrix} \quad (64)$$

$$M_{1,\alpha\beta}(\Omega, \mathbf{q}) = \frac{8}{\pi v_0} \frac{\omega_{\text{LT}} R_1^3 |a_1|^2}{(\omega_1 - \omega + i\gamma)} \sigma_{11,\alpha\beta}(\Omega, \mathbf{q})$$

$$M_{2,\alpha\beta}(\Omega, \mathbf{q}) = \frac{8}{\pi v_0} \frac{\omega_{\text{LT}} (R_1 R_2)^{3/2} a_1 a_2^*}{(\omega_2 - \omega + i\gamma)} \sigma_{12,\alpha\beta}(\Omega, \mathbf{q})$$

$$M_{3,\alpha\beta}(\Omega, \mathbf{q}) = \frac{8}{\pi v_0} \frac{\omega_{\text{LT}} (R_1 R_2)^{3/2} a_1^* a_2}{(\omega_1 - \omega + i\gamma)} \sigma_{21,\alpha\beta}(\Omega, \mathbf{q})$$

$$M_{4,\alpha\beta}(\Omega, \mathbf{q}) = \frac{8}{\pi v_0} \frac{\omega_{\text{LT}} R_2^3 |a_2|^2}{(\omega_2 - \omega + i\gamma)} \sigma_{22,\alpha\beta}(\Omega, \mathbf{q})$$

$$\sigma_{st,\alpha\beta}(\Omega, \mathbf{q}) = \sum_{\mathbf{g}} \frac{f(|\mathbf{g} + \mathbf{q}| R_s) f(|\mathbf{g} + \mathbf{q}| R_t) S_{\alpha\beta}(\mathbf{g} + \mathbf{q})}{\Omega^2(\mathbf{g} + \mathbf{q}) - \Omega^2} e^{i(s-t)(\mathbf{g} + \mathbf{q}) \cdot \boldsymbol{\tau}}$$

$$f(x) = \frac{\pi^2 \sin x}{x(\pi^2 - x^2)}$$

$$\Omega = \frac{\omega}{\omega_1}$$

$$\Omega(\mathbf{Q}) = \frac{c|\mathbf{Q}|}{\omega_1 n} \quad (65)$$

where $s, t = 1, 2$. The exciton-polariton dispersion $\omega(\mathbf{q})$ satisfies the equation

$$\text{Det}|D| = 0 \quad (66)$$

Note that by setting $|c_1|^2 = 0$ and $|c_2|^2 = 0.5$ (type-II QDs are totally transparent) we retrieve the case of a photonic crystal composed of only type-I QDs (Fu et al. 2000), see Fig. 4. By setting $|c_1|^2 = |c_2|^2 = 0$ and $\omega_1 = \omega_2$ we retrieve the case of a compound QD photonic crystal (Zeng et al. 2006b).

The dispersion relationship $\omega(\mathbf{q})$ for a QD dimer fcc lattice is presented in Fig. 6. The lattice is denoted by its lattice constant a , which is set to be $(0.95 a_{\text{Bragg}})$ (which is 116 nm for GaAs), where $a_{\text{Bragg}} = c\pi/\omega_1 n$. The lossy type-I QDs ($\hbar\omega_1 = 1.5$ eV) occupy the normal fcc lattice sites, while the excited type-II QDs ($\hbar\omega_2 = 1.503$ eV) are displaced by $\boldsymbol{\tau} = (a/2, a/2, a/2)$. $\hbar\omega_{\text{LT}} = 5$ meV. Here we observe the modification of the photonic band structure of the QD dimer system

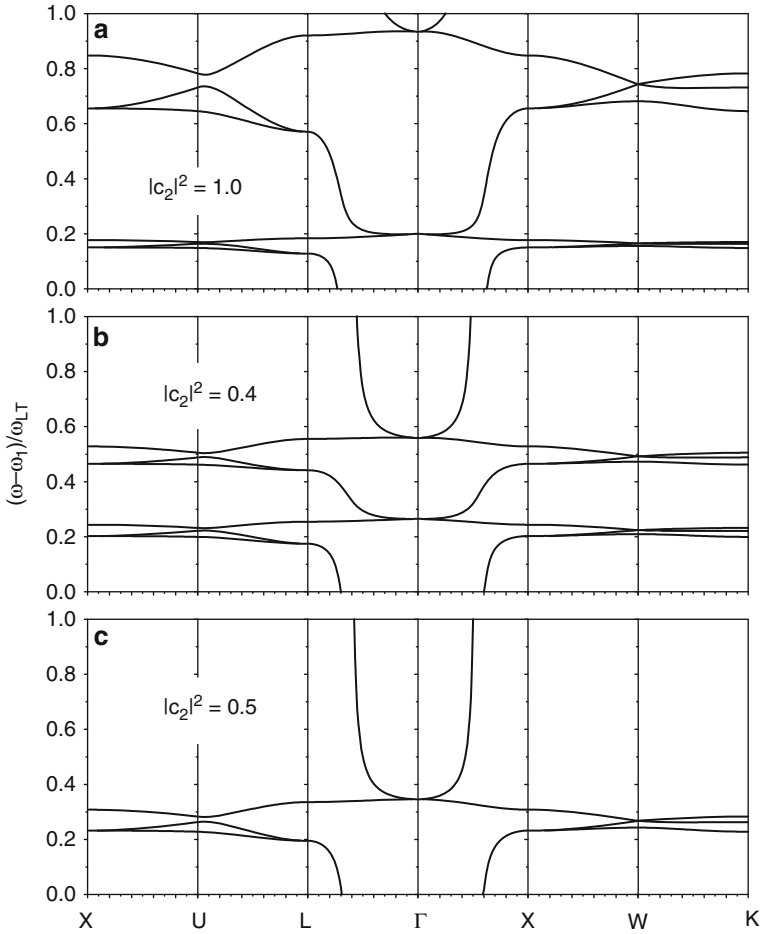


Fig. 6 Energy dispersion relationships of QD dimer systems in the fcc lattice. $R_1 = a/4$, $R_2 = a/5$, $\hbar\omega_1 = 1.5$ eV, $\hbar\omega_2 = 1.503$ eV, $\hbar\omega_{LT} = 5$ meV, $a = 0.95a_{\text{Bragg}}$, $\tau = (a/2, a/2, a/2)$. $|c_1|^2 = 0$. (a) $|c_2|^2 = 1.0$ (complete excitation), (b) $|c_2|^2 = 0.4$, (c) $|c_2|^2 = 0.5$ (type-II QDs are total transparent)

by pumping one type of the QDs (type-II QDs), which evolves from the one of only type-I QDs ($|c_1|^2 = 0$), i.e., at their ground exciton state) when type-II QDs are transparent [$|c_2|^2 = 0.5$, Fig. 6c] to the compound system [$|c_2|^2 = 1.0$, Fig. 6a]. We can observe modified but still characteristic features of the photonic dispersions of individual type-I and type-II QDs in their separate photonic crystal formats in the compound system, see Fig. 4. More specifically, the resonance features of type-I QDs around $(\omega - \omega_1) / \omega_{LT} = 0.3$ in Fig. 6d become compressed by the radiative

interaction between type-I QDs and type-II QDs, and they are also shifted down to around 0.18 in Fig. 6a.

By varying $|c_2|^2$ from 0.0 to 1.0 we find that the solutions of Eq. 66 are symmetric with respect to $|c_2|^2$ and $1 - |c_2|^2$ when $c_1 = 0$. Thus, Fig. 6a represents also the photonic dispersion of the QD dimer system when both type-I and type-II QDs are all initially at their ground exciton states, that is, $|c_1|^2 = |c_2|^2 = 0$. Thus the modification of the exciton state (from ground state to excited state) in one type of QDs in the dimer system does not affect the feature of the dispersion structure.

Lossless Dielectric Constant of QD Dimer Systems

As shown by Fig. 4, the exciton polarization provides possibilities of generating positively high dielectric constant as well as negative dielectric constant at the cost of disturbing the external electromagnetic field (either loss or gain). At least in the photonic crystal composed of one type of QDs (either lossy or gain), the dielectric modulation is always accompanied by loss or gain. The idea of the QD dimer system discussed in the previous sub-section is to compensate the loss of type-I QDs by the gain of type-II QDs, while still maintaining dielectric modulation. More specifically, we require $\varepsilon'' = 0$ at some frequencies in order to achieve lossless dielectric modulation for various optoelectronics applications.

In the following, we consider a PbSe/ZnSe QD assembly such as Fig. 5 immersed in a medium of dielectric constant ε_1 . The macroscopic dielectric constant ε for the ensemble of the QDs can be described by the dielectric theory of Maxwell-Garnett (1906) and Gittleman and Abeles (1977), which for two QD species can be written as

$$\frac{\varepsilon - \varepsilon_i}{\varepsilon + 2\varepsilon_i} = x_1 \frac{\varepsilon_{\text{QD1}} - \varepsilon_i}{\varepsilon_{\text{QD1}} + 2\varepsilon_i} + x_2 \frac{\varepsilon_{\text{QD2}} - \varepsilon_i}{\varepsilon_{\text{QD2}} + 2\varepsilon_i} \quad (67)$$

where x_i is the volume fraction of the i -th QD species. While Eq. 67 works best for x_1 less than 0.4, there is evidence that useful information can be available with higher concentrations (Cohen et al. 1973). Figure 7a shows the spectra of ε'_{QD} and $\varepsilon''_{\text{QD}}$ of the two types of QDs. For the combination of two types of QDs, that is, the QD dimer system, one type is lossy and the other gain, immersed in a conducting polymer with an effective dielectric constant $\varepsilon_1 = 1.8$ the optical spectrum is presented in Fig. 7b.

Figure 7 shows that a mixture of two types of QDs such as Fig. 5 can produce an effective dielectric constant that is lossless and negative, thus permitting, in concept, arbitrarily small scaling of the optical mode volume in the field of nanophotonics (Fu et al. 2008). Another binary mixture of lossy QDs and plasmonic silver nanorods was also shown to have a negative ε and compensated optical loss by using finite-difference time-domain calculations (Bratkovsky et al. 2008).

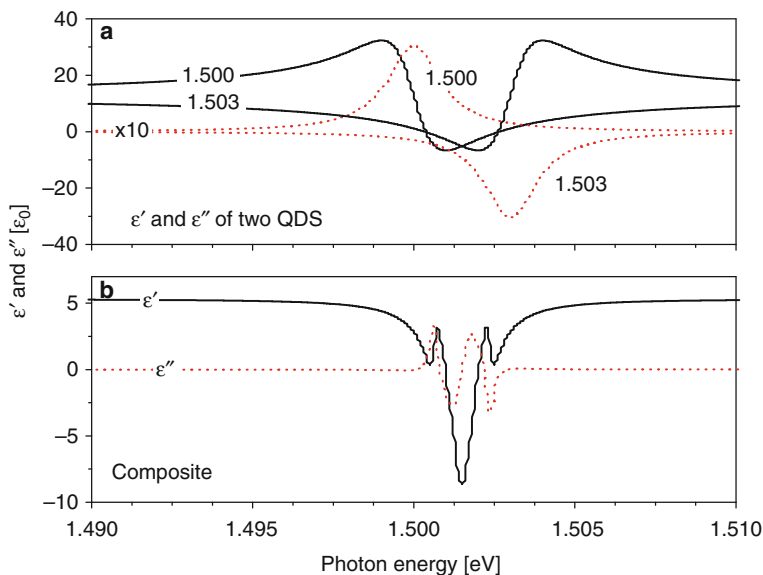


Fig. 7 (a) Dielectric constants of type-I QDs at 1.500 eV and $|c^{(0)}|^2 = 1.0$, and type-II QDs at 1.503 eV and $|c^{(0)}|^2 = 0.0$. *Solid lines: ϵ' ; Dotted lines: ϵ''* . Note that ϵ'' is magnified by 10. (b) QD dimer system after immersing the two types of QDs in $\epsilon_1 = 1.8$. The densities of two types of QDs are $7 \times 10^{16} \text{ cm}^{-3}$ Adapted from Fu et al. (2008)

Multiple-Photon and Multiple-Carrier Processes

Multiphoton Process

Another important aspect about the optical properties of QDs is the multiphoton process which has been widely applied in recent years in biological and medical imaging after the pioneer work of Goepfert-Mayer (1931), Lami et al. (1996), Helmchen et al. (1996), Yokoyama et al. (2006). The multiphoton process has largely been treated theoretically by steady-state perturbation approaches, for example, the scaling rules of multiphoton absorption by Wherrett (1984) and the analysis of two-photon excitation spectroscopy of CdSe QDs by Schmidt et al. (1996). Non-perturbation time-dependent Schrödinger equation was solved to analyze the ultrafast (fs) and ultra-intense (in many experiments the optical power of laser pulse peak can reach GW/cm^2) dynamics of multiphoton processes in QDs that are highly fluorescent with excellent photochemical stability, extreme brightness, and broad excitation spectral range (Fu et al. 2006c, 2007).

For spherical QDs the exciton wave function can be expressed by a radial function and spherical harmonics. The degeneracy of each exciton state is $2(2\ell + 1)$, where ℓ is the quantum number of the angular momentum. Figure 8 shows the exciton states (open circles) confined in the QD as well as their corresponding

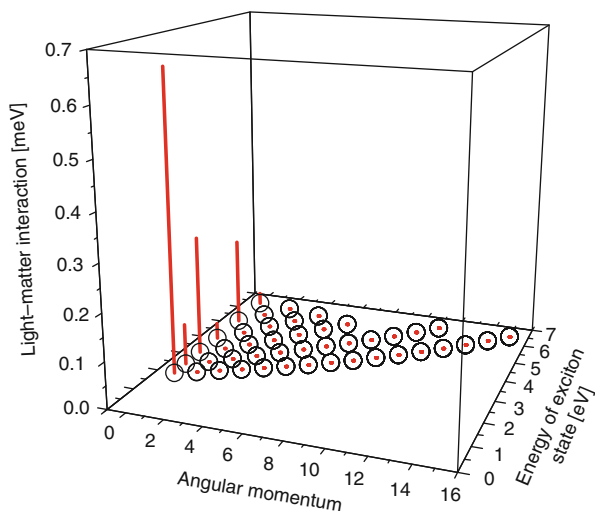


Fig. 8 Exciton states (*open circles*) confined in the QD and their corresponding light-matter interactions (*vertical lines*). Adapted from Fu et al. (2009)

light-matter interactions (vertical lines). For spherical QDs, only exciton states having zero angular momenta ($\ell = 0$) have nonzero relaxation energies. Moreover, only confined exciton states (i.e., both the electron and hole states that compose the exciton state under discussion are confined) have significant light-matter interaction.

In the ultrashort regime, after a continuous-wave optical excitation is switched on at $t = 0$, Fig. 9 shows that the time to reach peak two-photon excitation is much shorter than the one-photon excitation. Moreover, the multiphoton excitation spectral range is very broad. From about 1.0 to 2.5 eV (i.e., two-photon excitation range), the calculated excited exciton rate is much larger than the occupation of the first-excited exciton so that the emission spectrum is expected to be dominated by the recombination of the ground-state exciton. It was concluded by the theoretical simulation that experimentally observed strong multiphoton excitations can be reproduced when optical transitions among all confined states in the QD and an additional few hundred extended states in the barrier are taken into account (Fu et al. 2006c). This agrees with the experimental results of multi-particle (Huxter and Scholes 2006) and multi-exciton dynamics (Suffczyński et al. 2006).

Impact Ionization and Auger Recombination

Nozik proposed that impact ionization might be enhanced in semiconductor QDs to increase the efficiency of solar cells up to about 66 % (Nozik 2002). This was later verified experimentally by Schaller and Klimov (2004). Impact ionization was

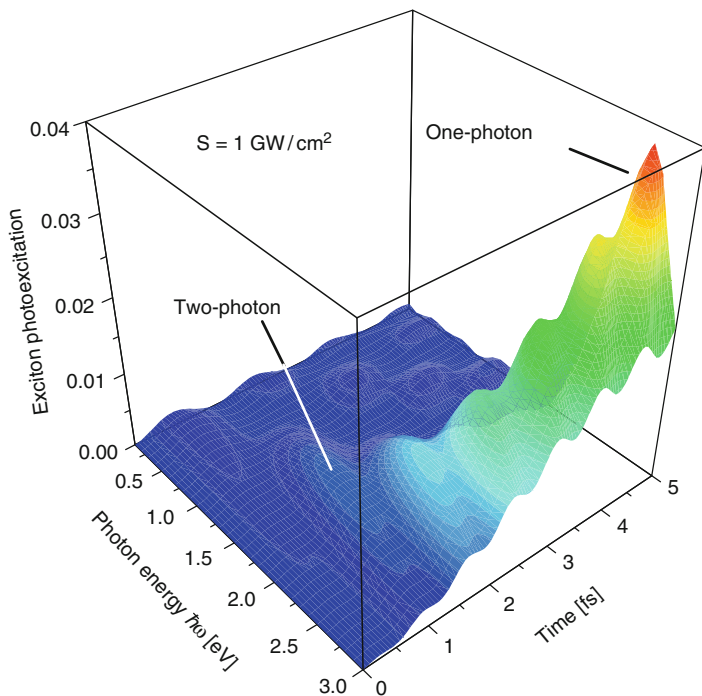


Fig. 9 Multiphoton excitation of QD exciton in an ultrashort period excited by a continuous wave light source switched on at $t = 0$. S denotes the optical power of the excitation light source. Adapted from Fu et al. (2006c)

reported to produce multiple exciton generation (MEG) per photon in a QD that results in a very high quantum yield (up to 300 % when the photon energy reaches four times the QD bandgap) in QD solar cells (Hanna et al. 2004). Multiple carrier extraction (>100 %) was observed at photon energies greater than 2.8 times the PbSe QD bandgap with about 210 % measured at 4.4 times the bandgap (Kim et al. 2008). In a recent communication in Nano Letters, Trinh et al. (2008) showed compelling support for carrier multiplication in PbSe QDs.

We concentrate here on reviewing impact ionization and Auger recombination in QDs using the solid-state theory. As shown in Fig. 10, a highly photoexcited electron and hole pairs can evolve into a multiple-exciton state through impact ionization. The carrier-carrier interaction is expressed by the Coulomb potential in the form of

$$V = \frac{e^2}{4\pi\epsilon |\mathbf{r}_1 - \mathbf{r}_2|} \tag{68}$$

to account for the two-body interaction of two electrons from an initial state $\phi_j(\mathbf{r}_1)\psi_i(\mathbf{r}_2)\exp[-i(E_j + E_i)t/\hbar]$, that is, one electron occupying a VB state

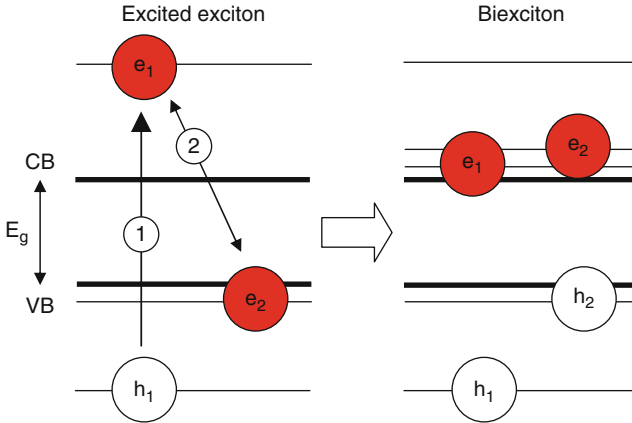


Fig. 10 Schematic depiction of impact ionization of a high-energy electron-hole pair. (1) Electron-hole pair (e_1 and h_1) is generated, for example, by a photon. (2) e_2 gets excited from a VB state to a CB state via Coulomb interaction with e_1 , leaving hole h_2 behind. The reverse process is referred to be Auger recombination

$$\phi_j(\mathbf{r}_1) \exp(-iE_j t/\hbar)$$

and another electron occupying a CB state

$$\psi_i(\mathbf{r}_2) \exp(-iE_i t/\hbar)$$

to a final state

$$\frac{1}{\sqrt{2}} \left| \begin{matrix} \psi_n(\mathbf{r}_1) & \psi_m(\mathbf{r}_1) \\ \psi_n(\mathbf{r}_2) & \psi_m(\mathbf{r}_2) \end{matrix} \right| \exp[-i(E_n + E_m)t/\hbar] \tag{69}$$

in the CB. Notice that the two electrons in the final state are not distinguishable so that we use a Slater determinant.

Other ionization pathways exist such as one electron occupying a low-energy VB state falls into h_1 and the released energy excites another electron from the VB to the CB. Auger recombination can be expressed similarly.

Before discussing the QD case, let us consider the computation of an Auger-type process in solid states that an incident high-energy CB electron with a wave vector k_1 collides with a second electron that occupies a VB state k_2 resulting in two CB electrons k'_1 and k'_2 . The general expression for the scattering matrix element of this process is

$$\iint \psi_{ck_1}^*(\mathbf{r}_1) \Phi_{vk_2}^*(\mathbf{r}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_{ck'_1}(\mathbf{r}_1) \psi_{ck'_2}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \tag{70}$$

where $\psi_{ck}(\mathbf{r})[\Phi_{vk}(\mathbf{r})]$ denotes the total wave function of the CB electron (VB hole) state k . For semiconductor systems and within the effective mass approximation $\psi_{ck}(\mathbf{r}) = \psi_{ck}(\mathbf{r})u_{ck}(\mathbf{r})$ and $\Phi_{vk}(\mathbf{r}) = \phi_{vk}(\mathbf{r})u_{vk}(\mathbf{r})$, where ψ and ϕ are envelope functions of the CB electron and VB hole, respectively, and u 's are periodic Bloch functions, the above expression becomes

$$\int \int \psi_{ck_1}^*(\mathbf{r}_1)u_{ck_1}^*(\mathbf{r}_1)\phi_{vk_2}^*(\mathbf{r}_2)u_{vk_2}^*(\mathbf{r}_2)\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \quad (71)$$

$$\times \psi_{ck'_1}(\mathbf{r}_1)u_{ck'_1}(\mathbf{r}_1)\psi_{ck'_2}(\mathbf{r}_2)u_{ck'_2}(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (72)$$

See, for example, Eq. 6.128 in Ridley (1988). Following overlap integrals are thus involved

$$I_1 = \int_{\text{cell}} u_{ck_1}^*(\mathbf{r}_1)u_{ck'_1}(\mathbf{r}_1) d\mathbf{r}_1 \quad (73)$$

$$I_2 = \int_{\text{cell}} u_{vk_2}^*(\mathbf{r}_2)u_{ck'_2}(\mathbf{r}_2) d\mathbf{r}_2$$

in the evaluation of the Auger-type scattering processes. The first overlap integral can be approximated to unity. Because of the orthogonality of u functions for different bands but the same \mathbf{k} , the second overlap integral I_2 is, in crudest approximation, zero. This, however, is not correct since the periodic Bloch functions are functions of wave vectors. More specifically, we can write

$$u_{ck}(\mathbf{r}) = u_c(\mathbf{r}) + \mathbf{k}\nabla_k u_c(\mathbf{r}) + \dots \quad (74)$$

$$u_{ck}(\mathbf{r}) = u_c(\mathbf{r}) + \mathbf{k}\nabla_k u_c(\mathbf{r}) + \dots$$

for small \mathbf{k} , where $u_c(\mathbf{r})$ and $u_v(\mathbf{r})$ are periodic Bloch functions at the CB and VB bandedges, respectively. $\int_{\text{cell}} u_v^*(\mathbf{r}_2)u_c(\mathbf{r}_2) d\mathbf{r}_2 = 0$. Substituting these expressions into the overlap integrals we obtain the squared overlap integral in terms of the heavy-hole mass m_v (Landsberg and Adams 1973; Ridley 1988)

$$|I_2|^2 = \frac{\hbar^2}{2E_g} \left(\frac{1}{m_0} + \frac{1}{m_v} \right) |\mathbf{k}_2 - \mathbf{k}'_2|^2 \quad (75)$$

where E_g is the energy band gap of the bulk material. By using the inverse Bohr radius of shallow impurities as a measure about the \mathbf{k} values, Landsberg and Adams obtained $|I_2| = 0.223$ for shallow-impurity-assisted Auger-type processes in bulk CdS and $|I_2| = 0.265$ in GaAs (Landsberg and Adams 1973). Note that the inverse Bohr radius of shallow impurities in bulk semiconductors is small. For QDs under investigation, the effective Bohr radius of the electron and hole distribution is largely determined by the QD size, which is about 5 nm, that is, very small compared with

the Bohr radius of shallow impurities in bulk semiconductors (about 100 nm in CdS and GaAs Landsberg and Adams 1973). This results in a large value of $|I_2|$. Under this specific circumstance we approximate $|I_2| = 1$. In other words, the electrons and holes in QDs, described by effective masses in the presence of the QD confinement potentials, interact with each other via the Coulomb force of Eq. 68 (Abrahams 1954). The approach has been adopted for describing carrier interactions in many electronic devices such as tunnel junctions where the kinetic energies of relevant carriers are large, see for examples Takenaka et al. (1979) and Rodina et al. (2002).

By the scattering theory and the generalized Golden rule (Landau and Lifshitz 1965) and denoting $\langle \phi_j \psi_i | \hat{T}(t) | \phi_j \psi_i \rangle \approx e^{-w_{ji}t/2}$ as the temporal development $\hat{T}(t)$ of state ji , it is easy to obtain

$$w_{ji} = \frac{2\pi}{\hbar} \sum_{nm \neq ji} |\langle \phi_j \psi_i | V | \psi_n \psi_m \rangle|^2 \delta(E_j + E_i - E_n - E_m) \tag{76}$$

$\Gamma_{ji} = \hbar w_{ji}/2$ is the relaxation energy and $1/w_{ji}$ the decaying time. In numerical calculations the δ functions in the above equation are replaced by $\Gamma_{ji} / [\Gamma_{ji}^2 + (E_j + E_i - E_n - E_m)^2]$ and Γ_{ji} is to be calculated by the above equation in a self-consistent way that the left and right sides of Eq. 76 equal after knowing the interaction values of $\langle \phi_j \psi_i | V | \psi_n \psi_m \rangle|^2$.

Numerical calculations of the Coulomb interaction Eq. 68 are not trivial. They were simplified in Allan and Delerue (2008) by using a screened Coulomb potential involving one-electron wave functions which were derived for bulk semiconductor materials (Landsberg 1991). To calculate the impact ionization interaction from an initial state of $R_{\ell_1}(\mathbf{r}_1) Y_{\ell_1 m_1}(\theta_1, \phi_1) R_{\ell_2}(\mathbf{r}_2) Y_{\ell_2 m_2}(\theta_2, \phi_2)$ to a final state $R_{\ell_3}(\mathbf{r}_1) Y_{\ell_3 m_3}(\theta_1, \phi_1) R_{\ell_4}(\mathbf{r}_2) Y_{\ell_4 m_4}(\theta_2, \phi_2)$ (the first term one in Eq. 69), we notice

$$\frac{1}{4\pi |\mathbf{r}_1 - \mathbf{r}_2|} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{1}{2\ell + 1} \frac{r_{<}^{\ell}}{r_{>}^{\ell+1}} Y_{\ell m}^*(\theta_2, \phi_2) Y_{\ell m}(\theta_1, \phi_1) \tag{77}$$

where $r_{<} = \min\{|\mathbf{r}_1|, |\mathbf{r}_2|\}$ and $r_{>} = \max\{|\mathbf{r}_1|, |\mathbf{r}_2|\}$ so that the impact ionization energy consists of

$$\begin{aligned} & \frac{r_{<}^{\ell}}{r_{>}^{\ell+1}} R_{\ell_1}(\mathbf{r}_1) R_{\ell_2}(\mathbf{r}_2) R_{\ell_3}(\mathbf{r}_1) R_{\ell_4}(\mathbf{r}_2) Y_{\ell_1 m_1}^*(\theta_1, \phi_1) Y_{\ell_2 m_2}^*(\theta_2, \phi_2) \\ & \times Y_{\ell_m}^*(\theta_2, \phi_2) Y_{\ell m}(\theta_1, \phi_1) Y_{\ell_3 m_3}(\theta_1, \phi_1) Y_{\ell_4 m_4}(\theta_2, \phi_2) \end{aligned} \tag{78}$$

for which we utilize the addition theorem for spherical harmonics (Cohen-Tannoudji et al. 1991)

$$Y_{\ell_1 m_1} Y_{\ell_2 m_2} = \sum_{\ell=|\ell_1-\ell_2|}^{\ell_1+\ell_2} \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)}{4\pi(2\ell+1)}} \times C(\ell 0 | \ell_1 0; \ell_2 0) C(\ell m_1 + m_2 | \ell_1 m_1; \ell_2 m_2) Y_{\ell m_1+m_2} \quad (79)$$

where $C(\ell m_1 + m_2 | \ell_1 m_1; \ell_2 m_2)$ is the CG-coefficient.

A more detailed analysis shows a few selection rules for ℓ 's and m 's such as $m_4 - m_3 = m_1 - m_3$. However, these selection rules can be easily fulfilled in nanoscale QDs because of the large number of available states confined in the QDs (see further discussion below). The most important qualitative selection rules refer to the radial functions that

$$\begin{aligned} & \int \int \frac{r_1^\ell}{r_1^{\ell+1}} R_{\ell_1}(r_1) R_{\ell_2}(r_2) R_{\ell_3}(r_1) R_{\ell_4}(r_2) \\ &= \int_0^a R_{\ell_1}(r_1) R_{\ell_3}(r_1) \left[\int_0^{r_1} \frac{r_2^\ell}{r_1^{\ell+1}} R_{\ell_2}(r_2) R_{\ell_4}(r_2) r_2^2 dr_2 \right. \\ & \left. + \int_{r_1}^a \frac{r_1^\ell}{r_2^{\ell+1}} R_{\ell_2}(r_2) R_{\ell_4}(r_2) r_2^2 dr_2 \right] r_1^2 dr_1 \end{aligned} \quad (80)$$

which shows that a direct spatial overlap between $R_{\ell_2}(r_2)$ (VB hole state) and $R_{\ell_4}(r_4)$ (CB electron state) will result in a large impact ionization. Impact ionization in type-II QDs is thus negligible since the electron-hole wave function overlap is small. Furthermore, to ensure energy conservation, $R_{\ell_1}(r)$ is normally a high-energy CB state, while both $R_{\ell_3}(r)$ and $R_{\ell_4}(r)$ are low-energy CB states. Impact ionization in many core-shell-structured QDs (e.g., Fu et al. 2007) could be small since the high-energy wave functions are much more extended (extended into shells) than the ground-state wave function (deeply confined in the core region).

We now consider a spherical CdSe QD in vacuum with a radius of a . The energy bandgap of bulk CdSe material is 1.74 eV (room temperature) and the exciton Bohr radius is 4.9 nm, the quantum confinement energies for CB electrons is 4.95 and 2.5 eV for the VB holes, and the electron and hole effective masses are $m_c^* = 0.14m_0$ and $m_v = 0.46m_0$ respectively, m_0 is the electron rest mass (Madelung 1992). In numerical calculations we also include a shell of vacuum with a thickness of 20 nm to account for the penetration of wave functions of high-energy states from the QD volume to the vacuum (the results are compared with the ones when a shell of 10 nm thick vacuum shell is included and no significant differences are observed). We include all the confined energy states in the QD (i.e., energy levels below the vacuum state) of all possible combinations of CB electron states and VB hole states that form the initial state of the impact ionization, then all possible combinations of two final CB electron states. Each single-particle state is denoted by its energy, $R(r)$, ℓ and m . Spherical harmonics (i.e., ℓ and m) is calculated using the addition theorem, while the radial integration is to be performed numerically.

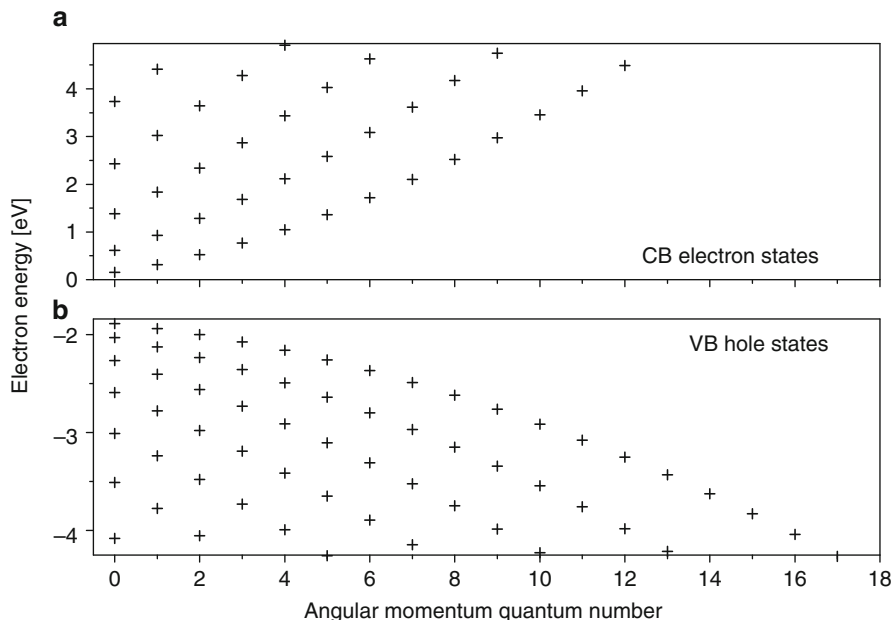


Fig. 11 (a) Conduction-band (CB) electron states and (b) valence-band (VB) hole states confined in a spherical CdSe QD with a radius of 4 nm as functions of angular momentum quantum number ℓ

Figure 11 shows the CB electron states and VB hole states confined in a spherical CdSe QD with a radius of 4 nm as functions of angular momentum quantum number ℓ . Note that states with $\ell \neq 0$ are $(2\ell + 1)$ -fold degenerate. We pick up one of the CB state from Fig. 11a and one VB state from Fig. 11b to calculate the total impact ionization energy to all possible combinations of two final CB states in Fig. 11a, which is shown in Fig. 12a.

Figure 12a conforms with what can be expected intuitively that the initial CB state has to be high while the VB state low in order to fulfil the energy conservation requirement. Furthermore, high-energy states can be highly degenerate because of large ℓ so Γ can be also high. For the case of Fig. 12 where $a = 4$ nm, we observe a maximal Γ of about 7 meV. We further observe tens of Γ over 4 meV while majorities are about 1 meV. These are very high as compared with the light–matter interaction of only 0.065 meV in a similar CdSe QD (Fu et al. 2009), thus we can be very optimistic about MEG processes in colloidal QDs, as has been much reported and also anticipated for significant solar cell applications. For the CdSe QDs under investigation, the MEG effect can be expected when the excitation radiation energy exceeds a threshold of about $2E_g$ (note that E_g is the energy of the ground exciton state, that is, nanocrystal energy gap in Schaller et al. (2005) when the QD radius $a = 4.8$ nm, while the threshold energy will be $2.8 E_g$ when $a = 4.0$ nm, and $3.5 E_g$ for $a = 3.0$ nm).

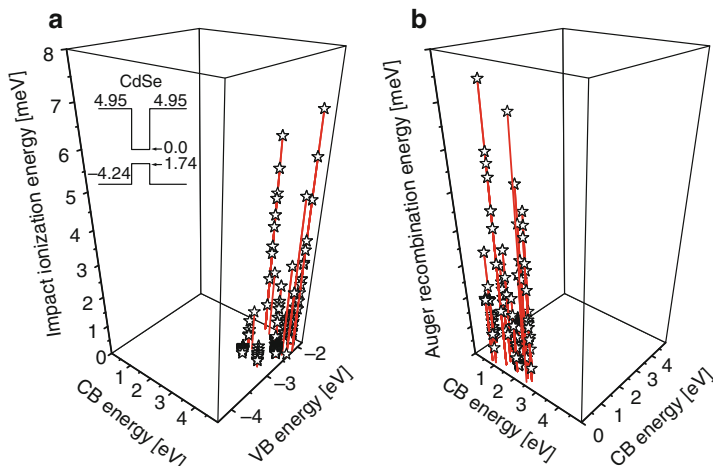


Fig. 12 (a) Impact ionization energy between initial CB state(s) and initial VB state(s). *Inset* shows the band structure of the CdSe QD in vacuum. (b) Auger recombination. The QD radius $a = 4$ nm. Adapted from Fu et al. (2010)

The Auger recombination process (reverse process of the impact ionization) shown in Fig. 12b can be similarly analyzed. Note that the energy range of the initial CB states in the Auger recombination processes is much wider than the impact ionization, especially for small QDs. This is due to the large density of states at high energy so that there are more available final CB states for two initial CB states to interact via the Auger recombination process.

Summary

We have briefly reviewed fundamental theory of excitons and excitonic polaritons in nanostructures. Experimental characterization and application-specific developments using excitons and excitonic polaritons in nanostructures at the electronic and photonic engineering levels have also been briefly discussed. A realization of the potential applications is intimately tied to the control of the geometry and localization of the nanoscale features. Positioning of the nanostructures not only allows for reproducible communication with the environment, but is also necessary for defining the interaction between nanostructures. Solid-state-based nanostructures such as self-assembled QDs provide excellent communication to the environment. The geometry control of these QD size is though relatively poor by the current molecular beam epitaxy (MBE) and metal-organic vapour phase epitaxy (MOVPE) techniques. The size of colloidal II–VI QDs can be very well controlled during the chemical synthetic processes and many colloidal II–VI QDs used in bioimaging applications are core-multishells formed so that possible effects of the surface and external environment on the exciton in the central core region

are minimal. However, the interconnection to the environment becomes a major issue, for example, for photocarrier extraction in solar cell applications. We will, nevertheless, surely witness a further fast development in the research and we hope that this chapter provides some foundation for the understanding and simulation of excitons and exciton polaritons in nanostructures.

Bibliography

- Abrahams, E. (1954). Electron-electron scattering in Alkali metals. *Physical Review*, *95*, 839–910.
- Allan, G., & Delerue, C. (2008). Influence of electronic structure and multiexciton spectral density on multiple-exciton generation in semiconductor nanocrystals: Tight-binding calculations. *Physical Review B*, *77*, 125340(10).
- Andreani, L. C., Gerace, D., & Agio, M. (2005). Exciton-polaritons and nanoscale cavities in photonic crystal slabs. *Physica Status Solidi B*, *242*, 2197–2209.
- Birman, J. L., & Huong, N. Q. (2007). Wannier-Frenkel hybrid exciton in organic-semiconductor quantum dot heterostructures. *Journal of Luminescence*, *125*, 196–200.
- Bratkovsky, A., Ponizovskaya, E., Wang, S.-Y., Holmström, P., Thylén, L., Fu, Y., & Ågren, H. (2008). A metal-wire/quantum-dot composite metamaterial with negative ϵ and compensated optical loss. *Applied Physics Letters*, *93*, 193106(3).
- Cohen, R. W., Cody, G. D., Coutts, M. D., & Abeles, B. (1973). Optical properties of granular silver and gold films. *Physical Review B*, *8*, 3689–3701.
- Cohen-Tannoudji, C., Diu, B., & Laloe, F. (1991). *Quantum mechanics* (Vol. 2, p. 1046). New York: Wiley-Interscience.
- Derfus, A. M., Chen, A. A., Min, D.-H., Ruoslahti, E., & Bhatia, S. N. (2007). Targeted quantum dot conjugates for siRNA delivery. *Bioconjugate Chemistry*, *18*, 1391–1396.
- Dimmock, J. O. (1967). Chapter 7 Introduction to the theory of exciton states in semiconductors. In R. K. Willardson & A. C. Beer (Eds.), *Semiconductors and Semimetals* (Vol. 3, pp. 259–319). New York: Academic.
- Franceschetti, A., An, J. M., & Zunger, A. (2006). Impact ionization can explain carrier multiplication in PbSe quantum dots. *Nano Letters*, *6*, 2191–2195.
- Fu, Y., & Willander, M. (1999). Chapter 1 Elemental and compound semiconductors. In *Physical model of semiconductor quantum devices* (pp. 1–22). Boston: Kluwer.
- Fu, Y., Willander, M., Ivchenko, E. L., & Kiselev, A. A. (1997). Four-wave mixing in microcavities with embedded quantum wells. *Physical Review B*, *55*, 9872–9879.
- Fu, Y., Willander, M., & Ivchenko, E. L. (2000). Photonic dispersions of semiconductor-quantum-dot-array-based photonic crystals in primitive and face-centered cubic lattices. *Superlattices and Microstructures*, *27*, 255–264.
- Fu, Y., Willander, M., & Xu, Q.-X. (2006a). Chapter 5 Quantum effects and nanofabrications in scaling metal-oxide-semiconductor devices. In A. A. Balandin & K. L. Wang (Eds.), *Handbook of semiconductor nanostructures and nanodevices* (Vol. 5, pp. 229–256). Los Angeles: American Scientific Publishers.
- Fu, Y., Berglind, E., Thylén, L., & Ågren, H. (2006b). Optical transmission and waveguiding by excitonic quantum dot lattices. *Journal of the Optical Society of America B*, *23*, 2441–2447.
- Fu, Y., Han, T.-T., Luo, Y., & Ågren, H. (2006c). Multi-photon excitation of quantum dots by ultra-short and ultra-intense laser pulse. *Applied Physics Letters*, *88*, 221114(3).
- Fu, Y., Han, T.-T., Ågren, H., Lin, L., Chen, P., Liu, Y., Tang, G.-Q., Wu, J., Yue, Y., & Dai, N. (2007). Design of semiconductor CdSe-core ZnS/CdS-multishell quantum dots for multiphoton applications. *Applied Physics Letters*, *90*, 173102(3).
- Fu, Y., Thylén, L., & Ågren, H. (2008). A lossless negative dielectric constant from quantum dot exciton polaritons. *Nano Letters*, *8*, 1551–1555.

- Fu, Y., Ågren, H., Kowalewski, J. M., Brismar, H., Wu, J., Yue, Y., Dai, N., & Thylén, L. (2009). Radiative and nonradiative recombination of photoexcited excitons in multi-shell-coated CdSe/CdS/ZnS quantum dots. *Europhysics Letters*, *86*, 37003(6).
- Fu, Y., Zhou, Y.-H., Su, H., Boey, F. Y. C., & Ågren, H. (2010). Impact ionization and Auger recombination rates in semiconductor quantum dots. *Journal of Physical Chemistry C*, *114*, 3743–3747.
- Gasiorowicz, S. (1996). *Quantum physics* (p. 178). New York: Wiley.
- Gittleman, J. I., & Abeles, B. (1977). Comparison of the effective medium and the Maxwell-Garnett predictions for the dielectric constants of granular metals. *Physical Review B*, *15*, 3273–3275.
- Hanna, M. C., Ellingson, R. J., Beard, M., Yu, P., Micic, O. I., & Nozik, A. J. (2004, October 25–28). Quantum dot solar cells: High efficiency through multiple exciton generation. 2004 DOE Solar Energy Technologies Program Review Meeting, Denver, Colorado.
- Helmchen, F., Svododa, K., Denk, W., Kleinfeld, D., & Tank, D. W. (1996). In vivo dendritic calcium dynamics in deep-layer cortical pyramidal neurons. *Nature Neuroscience*, *2*, 989–996.
- Huxter, V. M., & Scholes, G. D. (2006). Nonlinear optical approach to multiexciton relaxation dynamics in quantum dots. *Journal of Chemical Physics*, *125*, 144716–144712.
- International Technology Roadmap for Semiconductors. www.itrs.net.
- Ivchenko, E. L., Fu, Y., & Willander, M. (2000). Exciton polaritons in quantum-dot photonic crystals. *Physics of the Solid State*, *42*, 1756–1765.
- Jiang, J., Gao, B., Han, T.-T., & Fu, Y. (2009). Ab initio study of energy band structures of GaAs nanoclusters. *Applied Physics Letters*, *94*, 092110(3).
- Kane, E. O. (1957). Band structure of indium antimonide. *Journal of Physics and Chemistry of Solids*, *1*, 249.
- Kavokin, A. (2007). Exciton-polaritons in microcavities: Present and future. *Applied Physics A*, *89*, 241–246.
- Kim, S. J., Kim, W. J., Sahoo, Y., Cartwright, A. N., & Prasad, P. N. (2008). Multiple exciton generation and electrical extraction from a PbSe quantum dot photoconductor. *Applied Physics Letters*, *92*, 31107(3).
- Lami, J.-F., Gilliot, P., & Hirlimann, C. (1996). Observation of interband two-photon absorption saturation in CdS. *Physical Review Letters*, *77*, 1632–1635.
- Landau, L. D., & Lifshitz, E. M. (1965). *Quantum mechanics* (2nd ed., p. 129). Oxford: Pergamon Press.
- Landsberg, P. T. (1991). *Recombination in semiconductors*. London: Cambridge University Press.
- Landsberg, P. T., & Adams, M. J. (1973). Theory of donor-acceptor radiative and Auger recombination in simple semiconductors. *Proceedings of the Royal Society of London A*, *334*, 523–539.
- Madelung, O. (Ed.). (1991). *Semiconductors group IV elements and III-V compounds*. Berlin: Springer.
- Madelung, O. (Ed.). (1992). *Data in science and technology: Semiconductors other than group IV elements and III-V compounds*. Boston: Springer.
- Maxwell-Garnett, J. C. (1906). Colours in metal glasses, in metallic films, and in metallic solutions. II. *Philosophical Transactions of the Royal Society of London*, *205*, 237–288.
- Mayer, M. G. (1931). Elementary processes with two quantum jumps. *Annalen Der Physik (Leipzig)*, *9*, 273–294.
- Medintz, I. L., Uyeda, H. T., Goldman, E. R., & Mattoussi, H. (2005). Quantum dot bioconjugates for imaging, labelling and sensing. *Nature Materials*, *4*, 435–446.
- Miller, D. A. B., Chemla, D. S., Eilenberg, D. J., Smith, P. W., Gossard, A. C., & Tsang, W. T. (1982). Large room-temperature optical nonlinearity in GaAs/Ga_{1-x}Al_xAs multiple quantum well structures. *Applied Physics Letters*, *41*, 679–681.
- Molnár, M., Fu, Y., Friberg, P., & Chen, Y. (2010). Optical characterization of colloidal CdSe quantum dots in endothelial progenitor cells. *Journal of Nanobiotechnology*, *8*, 2. doi:[10.1186/1477-3155-8-2](https://doi.org/10.1186/1477-3155-8-2).CrossRef.

- Nozik, A. J. (2002). Quantum dot solar cells. *Physica E: Low-dimensional Systems and Nanostructures*, 14, 115–120.
- Rabani, E., & Baer, R. (2008). Distribution of multiexciton generation rates in CdSe and InAs nanocrystals. *Nano Letters*, 8, 4488–4492.
- Ridley, B. K. (1988). *Quantum processes in semiconductors* (pp. 269–278). Oxford: Clarendon Press.
- Rodina, P., Ebert, U., Hundsdoerfer, W., & Grekhov, I. (2002). Tunneling-assisted impact ionization fronts in semiconductors. *Journal of Applied Physics*, 92, 958–964.
- Schaller, R. D., & Klimov, V. I. (2004). High efficiency carrier multiplication in PbSe nanocrystals: Implications for solar energy conversion. *Physical Review Letters*, 92, 186601(4).
- Schaller, R. D., Agranovich, V. M., & Klimov, V. I. (2005). Mechanism for high-efficiency carrier multiplication in semiconductor nanocrystals: Direct photogeneration of multiexcitons via virtual single-exciton states. *Nature Physics*, 1, 189–194.
- Schmidt, M. E., Blanton, S. A., Hines, M. A., & Guyot-Sionnest, P. (1996). Size-dependent two-photon excitation spectroscopy of CdSe nanocrystals. *Physical Review B*, 53, 12629–12632.
- Scholes, G. D., & Rumbles, G. (2006). Excitons in nanoscale systems. *Nature Materials*, 5, 683–696.
- Sturge, M. D., & Rashba, E. I. (Eds.). (1982). *Excitons*. Amsterdam: North-Holland.
- Suffczyński, J., Kazimierzczuk, T., Goryca, M., Piechal, B., Trajnerowicz, A., Kowalik, K., Kossacki, P., Golnik, A., Korona, K. P., Nawrocki, M., & Gaj, J. A. (2006). Excitation mechanisms of individual CdTe/ZnTe quantum dots studied by photon correlation spectroscopy. *Physical Review B*, 74, 085319(7).
- Sun, H. D., Makino, T., Segawa, Y., Kawasaki, M., Ohtomo, A., Tamura, K., & Koinuma, H. (2002). Enhancement of exciton binding energies in ZnO/ZnMgO multiquantum wells. *Journal of Applied Physics*, 91, 1993–1997.
- Takenaka, N., Inoue, M., & Inuishi, Y. (1979). Influence of inter-carrier scattering on hot electron distribution function in GaAs. *Journal of the Physical Society of Japan*, 47, 861–868.
- Thylén, L., He, S., Wosinski, L., & Dai, D. (2006). The Moore's Law for photonic integrated circuits. *Journal of Zhejiang University Science A*, 7, 1961–1967.
- Trinh, M. T., Houtepen, A. J., Schins, J. M., Hanrath, T., Pirus, J., Knulst, W., Goossens, A. P. L. M., & Siebbeles, L. D. A. (2008). In spite of recent doubts carrier multiplication does occur in PbSe nanocrystals. *Nano Letters*, 8, 1713–1718.
- Vashist, S. K., Tewari, R., Bajpai, R. P., Bharadwaj, L. M., & Raiteri, R. (2006). Review of quantum dot technologies for cancer detection and treatment. *Azozono Journal of Nanotechnology Online*, 2, 1–14. doi:10.2240/azozono0113.
- Vlasov, Y. A., Astratov, V. N., Karimov, O. Z., Kaplyanskiy, A. A., Bogomolov, V. N., & Prokofiev, A. V. (1997). Existence of a photonic pseudogap for visible light in synthetic opals. *Physical Review*, B55, R13357–13360.
- Vurgaftman, I., Meyer, J. R., & Ram-Mohan, L. R. (2001). Band parameters for III-V compound semiconductors and their alloys. *Journal of Applied Physics*, 89, 5815–5875.
- Weisbuch, C., Benisty, H., & Houdré, R. (2000). Overview of fundamentals and applications of electrons, excitons and photons in confined structures. *Journal of Luminescence*, 85, 271–293.
- Wherrett, B. S. (1984). Scaling rules for multiphoton interband absorption in semiconductors. *Journal of the Optical Society of America B-Optical Physics*, 1, 67–72.
- Xu, W.-L., Fu, Y., Willander, M., & Shen, S. C. (1994). Theory of normal incident absorption for the intersubband transition in n-type indirect-gap semiconductor quantum wells. *Physical Review B*, 49, 13760–13766.
- Yannopapas, V. (2007). Artificial magnetism and negative refractive index in three-dimensional metamaterials of spherical particles at near-infrared and visible frequencies. *Applied Physics A*, 87, 259–264.
- Yannopapas, V. (2008). Subwavelength imaging of light by arrays of metal-coated semiconductor nanoparticles: A theoretical study. *Journal of Physics: Condensed Matter*, 20, 255201–255208.

- Yatsui, T., Sangu, S., Kawazoe, T., Ohtsu, M., An, S. J., Yoo, J., & Yi, G.-C. (2007). Nanophotonic switch using ZnO nanorod double-quantum-well structures. *Applied Physics Letters*, *90*, 223110(3).
- Yatsui, T., Yib, G.-C., & Ohtsu, M. (2008). Progress in developing nanophotonic integrated circuits. *Proceedings of SPIE*, *7007*, 700703(8).
- Yokoyama, H., Guo, H., Yoda, T., Takashima, K., Sato, K.-I., Taniguchi, H., & Ito, H. (2006). Two-photon bioimaging with picosecond optical pulses from a semiconductor laser. *Optics Express*, *14*, 3467–3471.
- Zamfirescu, M., Kavokin, A., Gil, B., Malpuech, G., & Kaliteevski, M. (2002). ZnO as a material mostly adapted for the realization of room-temperature polariton lasers. *Physical Review B*, *65*, 161205(4).
- Zeng, Y., Fu, Y., Chen, X., Lu, W., & Ågren, H. (2006a). Complete band gaps in three-dimensional quantum-dot photonic crystals. *Physical Review B*, *74*, 115325(5).
- Zeng, Y., Chen, X.-S., Lu, W., & Fu, Y. (2006b). Exciton polaritons of nano-spherical-particle photonic crystals in compound lattices. *The European Physical Journal B*, *49*, 313–318.
- Zia, R., Schuller, J. A., Chandran, A., & Brongersma, M. L. (2006). Plasmonics: The next chip-scale technology. *Materials Today*, *9*, 20.

Modeling of Quasi-One-Dimensional Carbon Nanostructures with Density Functional Theory

29

Veronica Barone, Oded Hod, and Juan E. Peralta

Contents

Density Functional Theory with Periodic Boundary Conditions	1298
Structure–Property Relations in Single-Walled Carbon Nanotubes and Graphene Nanoribbons	1300
Single-Walled Carbon Nanotubes	1300
Graphene Nanoribbons	1303
Modeling the Optical Spectrum of Single-Walled Carbon Nanotubes and Graphene Nanoribbons	1308
Single-Walled Carbon Nanotubes	1308
Graphene Nanoribbons	1311
Chemistry at the Edges of Graphene	1312
Finite Size Effects in Low-Dimensional Graphitic Materials	1315
Quantum Confinement in Graphitic Systems	1315
Edge Effects in Graphitic Systems	1318
Electromechanical Properties of One-Dimensional Graphitic Structures	1323
Carbon Nanotubes in NEMS Applications	1324
Graphene Nanoribbons in NEMS Applications	1327
Concluding Remarks	1329
Bibliography	1330

Abstract

The purpose of this chapter is to describe and review examples of how theoretical investigations can be applied to elucidate the behavior of carbon nanostructures

V. Barone (✉) • J.E. Peralta
Department of Physics, Central Michigan University, Mount Pleasant, MI, USA
e-mail: v.barone@cmich.edu; juan.peralta@cmich.edu

O. Hod
Department of Chemical Physics, School of Chemistry, The Sackler Faculty of Exact Sciences,
Tel Aviv University, Tel Aviv, Israel
e-mail: odedhod@post.tau.ac.il

and to understand the physical mechanisms taking place at the molecular level. We will place a special emphasis in theoretical works utilizing density functional theory. We assume that the reader is familiar with the basics of density functional theory as well as the electronic properties of single-walled carbon nanotubes and graphene nanoribbons (GNRs). We do not intend to present an extensive review; instead, we focus on several examples to illustrate the powerful predictive capabilities of current computational approaches.

Density Functional Theory with Periodic Boundary Conditions

The focus of this chapter is on carbon structures that are elongated, predominantly one-dimensional systems. For most modeling purposes, these structures can be considered as periodic in one dimension. There are many software packages either free or commercially available to perform density functional theory (DFT) calculations using periodic boundary conditions. In fact, these type of calculations are routine in condensed matter physics and became very popular recently in the quantum chemistry community. One of the most widely used approaches for DFT calculations with periodic boundary conditions is the combination of the local density approximation (LDA) or the generalized gradient approximation (GGA) and plane waves. This approach usually involves only valence electrons, while the effect of core electrons is represented with pseudopotentials. Many modern pseudopotential calculations use ultrasoft pseudopotentials, which were first developed by Vanderbilt in the early 1990s (Vanderbilt 1990). Full-potential all-electron calculations can also be performed using plane waves, although they are somewhat computationally expensive for routine calculations. On the other hand, another class of software packages employ localized basis functions such as Gaussians. These basis functions allow not only an all-electron treatment of the system but also present the flexibility of pseudopotential calculations when heavy atoms are involved. In most cases, dealing with systems involving first- and second-row elements in an all-electron fashion can be routinely done. Another advantage of localized basis functions is that periodic boundary conditions can be explicitly imposed in either one, two, or three dimensions, in contrast to plane waves, which are naturally periodic in three dimensions and therefore approaches such as the super-cell are commonly used to reduce the dimensionality from three to two or one at the price of increasing the computational cost. The main advantage of plane waves over Gaussian basis functions is that the basis set convergence is more controlled. The basis set convergence of Gaussian-type orbitals is well known for molecules and there are tens of available basis sets developed for different purposes (see for instance the EMSL library Feller 1996, 2007). However, the basis set convergence of Gaussian functions is less known for periodic systems, especially for properties like the bandgap. Another disadvantage of Gaussians is that available basis for molecular systems cannot always be straightforwardly employed in periodic boundary conditions calculations due to linear dependencies that occur in periodic structures, limiting in practice the use of large exponents (Gruneich and Hess 1998).

Table 1 Bandgaps (in eV) for the (10,0) SWNT obtained with different functionals (LDA, PBE, and HSE, shown row-wise) and basis sets (3-21G, 3-21G*, and 6-31G*, shown column-wise)

	LDA optimized			PBE optimized			HSE optimized		
	3-21G	3-21G*	6-31G*	3-21G	3-21G*	6-31G*	3-21G	3-21G*	6-31G*
LDA	0.79	0.79	0.78	0.77	0.77	0.76	0.79	0.79	0.77
PBE	0.80	0.80	0.79	0.78	0.78	0.76	0.79	0.79	0.78
HSE	1.00	1.00	0.98	0.97	0.97	0.96	0.99	0.99	0.98

The structure was optimized using the functional and basis set indicated in each column. Calculations were performed with the Gaussian development program (Frisch et al. 2006)

When using localized functions, in general, the basis set needs to be carefully chosen in order to obtain sensible results. It is important to point out that using either localized basis or plane waves, different properties can have different basis set convergence behavior, and the fact that one property such as the energy is converged (within a given criterion) does not imply that all properties are converged. In single-walled carbon nanotubes (SWNTs), for instance, the choice of the Gaussian basis set does not seem to have a significant impact on the bandgap. In Table 1, we show the bandgap (obtained as Kohn–Sham band energy differences) of the (10,0) SWNT calculated with different functionals and standard Gaussian basis sets. The initial SWNT geometry was obtained with the *Tubegen* program (Frey and Doren 2005) and then relaxed using the different functionals and basis sets indicated in each column. For each optimized structure, we calculate the bandgap of the system with the different functionals and basis sets. As shown in Table 1, the bandgap does not depend significantly on the level of theory utilized for the geometry optimization, or the basis set. Table 1 illustrates, however, the dependence of the bandgap upon the choice of the exchange–correlation (XC) functional. This shows that for SWNTs, the choice of the XC functional has a larger impact on the calculated bandgap and related properties.

There are several options available in DFT software packages for the choice of the approximate XC functional. The most widely used XC functionals depend either on the electronic density (local[–spin] density approximation or L[S]DA) or on the electronic density and its gradient (generalized gradient approximation or GGA). For example, in Table 1 we employ the combination of Dirac exchange and the parameterization of Vosko et al. (1980) for correlation in the LSDA functional and the GGA functional of Perdew, Burke, and Ernzerhof (PBE) (1996). Comparing with available experimental data, these families of functionals usually perform well for structural properties but have some deficiencies for energetics and electronic properties. Other type of approximate XC functionals include an orbital dependency either through quantities such as the kinetic energy density or Hartree–Fock (HF) type of exchange (for a thorough review, see Kümmel and Kronik 2008). Examples of orbital-dependent functionals are the meta-GGA TPSS (Tao et al. 2003), or the hybrid functionals PBE0 (Adamo and Barone 1999; Ernzerhof and Scuseria 1999; Perdew et al. 1997) (also known as PBEh), and B3LYP (Becke 1993). These type of functionals generally show some improvement over the LSDA and GGA but

are computationally more demanding. In particular, due to the nature of the HF approximation, hybrid functionals in extended systems can be extremely demanding for metallic and small bandgap systems (Paier et al. 2007). An alternative to standard hybrid functionals for extended systems are the short-range hybrid functionals such as HSE (Heyd et al. 2003, 2006). These functionals truncate the long-range tail of the electron–electron interaction in the exchange contribution to the electronic energy. In this way, the remaining (short-range) part of the HF exchange can be efficiently evaluated (Adamson et al. 1996, 1999). It has been shown that this truncation has little effect on the properties of finite systems and at the same time provides an efficient route for hybrid DFT calculations in extended systems (Heyd and Scuseria 2004).

Most commonly available functionals, either derived from nonempirical grounds or using fitting procedures, are meant to approximate only the electronic ground state, and not excited states. Moreover, although in “static” DFT, Kohn–Sham eigenvalues are often used to evaluate the energy bandgap, there is no formal justification for this: The formal theoretical framework is provided by time-dependent DFT (Runge and Gross 1984). For a review, see Onida et al. (2002). Therefore, the evaluation of the energy bandgaps and related properties from DFT calculations needs to be carried out with care. From a pragmatic viewpoint, Kohn–Sham eigenvalues are often used to distinguish between metals and semiconductors and to estimate excitation energies. LSDA and GGA Kohn–Sham gaps usually underestimate experimental data and sometimes small-gap semiconductors are erroneously predicted to be metallic. Kohn–Sham gaps from hybrid functional calculations, on the other hand are, in general, in better agreement with experimental gaps (Barone et al. 2005a, b; Heyd et al. 2005; Kümmel and Kronik 2008; Peralta et al. 2006), although they tend to overestimate, in general, the available experimental data. Short-range hybrid functionals improve upon regular hybrids and provide good agreement with experiments and more sophisticated many-electron approaches (Barone et al. 2005b; Batista et al. 2006).

Structure–Property Relations in Single-Walled Carbon Nanotubes and Graphene Nanoribbons

Single-Walled Carbon Nanotubes

Carbon nanotubes are one of the most fascinating carbon nanostructures. Although soon after the discovery of fullerenes in 1985 (Kroto et al. 1985) there were many speculations about the possibility of synthesizing long and narrow graphitic cage-like structures, the synthesis of the first carbon nanotubes was reported by Iijima in 1991. After that seminal work, several new routes toward the synthesis of single- and multi-walled carbon nanotubes were developed.

The electronic properties of SWNT were first studied within the so-called zone-folding scheme (ZF) (Hamada et al. 1992; Saito et al. 1992, 1998). This

scheme is based on the tight-binding (TB) approximation for the two-dimensional honeycomb lattice and the subsequent quantization of the wave vector associated with the radial direction of the nanotube. In this framework, the electronic structure of SWNTs is determined by the allowed values of the quantized wave vector. Within this approach, each nanotube, characterized by the indices (n, m) , can be either metallic if $(2n + m)/3 = k$, with k integer, or semiconducting for k non-integer (with a bandgap that depends on the inverse diameter, $1/d$). This versatility in the electronic behavior, that depends *only* on the geometric structure of the tubes, ignited a large amount of experimental and theoretical efforts with the promise of novel technological applications based on these unique nanostructured materials.

In 1998, Tang et al. produced SWNTs grown in zeolite channels with diameters of about 4 Å (Tang et al. 1998). From the optical spectrum and the possible diameter distributions, these tubes were identified as the (3,3), (4,2), and (5,0) (Li et al. 2001). In 2004, the thinnest SWNTs, with diameters of about 3 Å were grown inside multi-walled nanotubes. Because of the small diameter, the possible (n, m) indices of these narrow SWNT were predicted to be (2,2), (4,0), and (3,1) (Zhao et al. 2004). DFT calculations pointed out the importance of curvature effects, especially for narrower nanotubes which present a significant $\sigma - \pi$ hybridization not accounted for in the ZF scheme (Blase et al. 1994; Gülsiren et al. 2002; Mintmire et al. 1992; Reich et al. 2002). Although the ZF scheme works reasonably well for the average diameter SWNT of about 10 Å, several studies confirmed the profound effect of the $\sigma - \pi$ hybridization in small diameter SWNT on their electronic properties.

DFT calculations have shown that the zigzag (5,0) SWNT, which should be a semiconductor according to the zone-folding scheme, is indeed metallic (Barone and Scuseria 2004; Cabria et al. 2003; Li et al. 2001; Liu and Chan 2002; Machón et al. 2002; Springborg and Satpathy 1994). The same holds true for the zigzag (4,0) SWNT (Barone and Scuseria 2004; Cabria et al. 2003). Although it is tempting to assume that narrower tubes tend to be metallic due to the $\sigma - \pi$ hybridization, it has been shown that the narrowest chiral tubes (4,3) and (3,1) present the largest bandgaps of all SWNTs (1.7 and 1.3 eV, respectively, obtained with the HSE functional) (Barone and Scuseria 2004). Due to curvature effects, most narrow semiconducting SWNTs become indirect gap semiconductors and largely deviate from the ZF $1/d$ predictions.

On the modeling side, we note that the calculated electronic properties of narrow nanotubes strongly depend on the exchange-correlation functional used. In Fig. 1 we show the band structure and density of states (DOS) for the three narrow (5,0), (3,3), and (4,2) nanotubes obtained with LDA and the hybrid PBE0 functionals. The largest differences are found for the chiral semiconducting tube for which PBE0 predicts a bandgap as large as 1.36 eV. However, as we show in section “[Modeling the Optical Spectrum of Single-Walled Carbon Nanotubes and Graphene Nanoribbons](#),” while the bandgap of SWNTs is largely underestimated by LSDA, it is also significantly overestimated by PBE0.

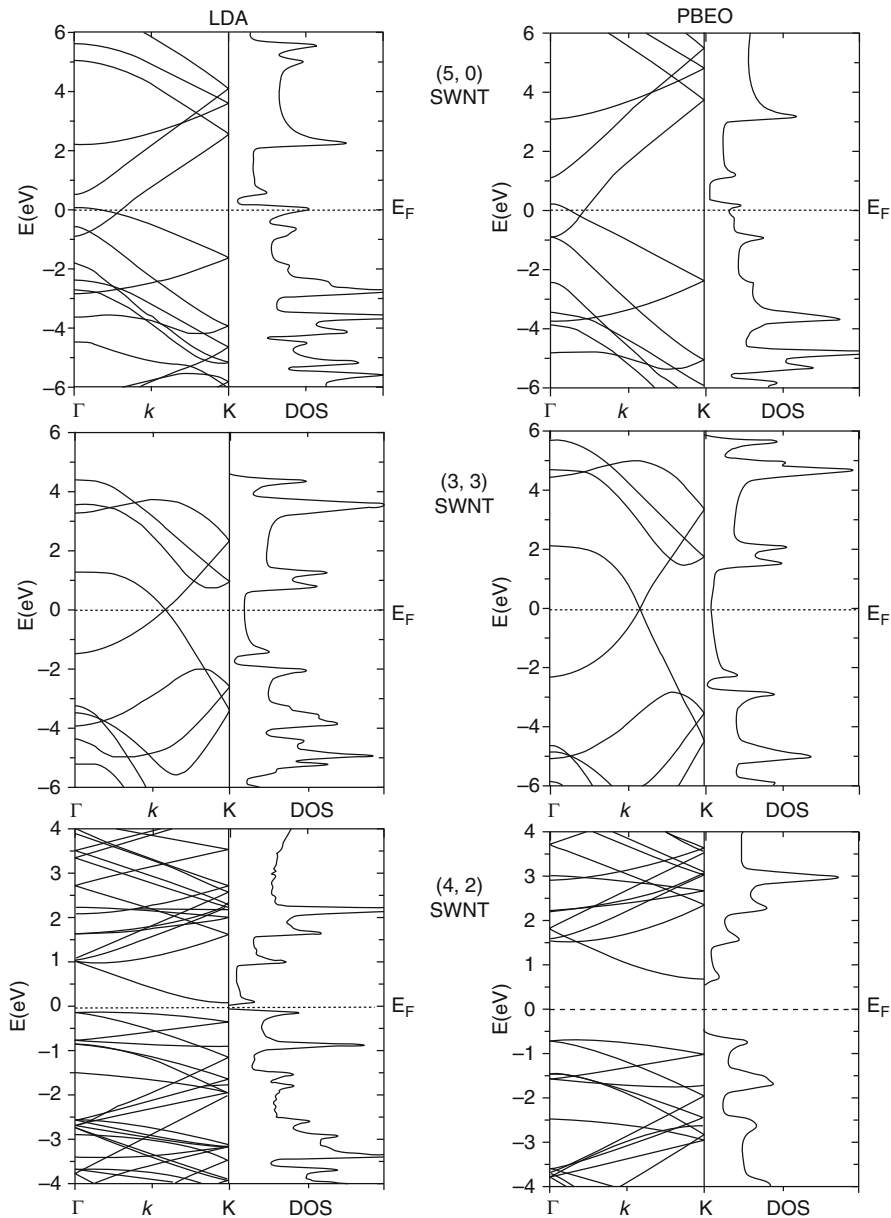


Fig. 1 Band structure and electronic density of states for the (5,0), (3,3), and (4,2) SWNT. (a) Obtained at the LDA/6-31G* level. (b) Obtained at the PBE0/6-31G* level

Graphene Nanoribbons

Graphene, a single layer of graphite, has attracted a lot of attention since 2004 when it was isolated for the first time as a stand-alone two-dimensional all-carbon network (Novoselov et al. 2004). Since then, major accomplishments in their production have been achieved and now, lower dimensional graphene derivatives are routinely created in the lab (Berger et al. 2006; Cai et al. 2010; Han et al. 2007; Ritter and Lyding 2009; Wang et al. 2008; Yang et al. 2008). Recent studies point to graphene and its derivatives as one of the most promising materials for technological applications ranging from spintronic devices to energy storage media (Bhardwaj et al. 2010; Son et al. 2006b).

By cutting a quasi-one-dimensional structure out of a two-dimensional graphene sheet, it is possible to obtain a flat structure usually referred to as a graphene nanoribbon (GNR). These ribbons can also be thought of as unwrapped nanotubes of a given chirality. In Fig. 2 we present a scheme of the two extreme crystallographic orientations of the GNRs axis: armchair (AGNR) and zigzag (ZGNR). Dangling bonds at the edges are usually passivated with hydrogen atoms for the purpose of

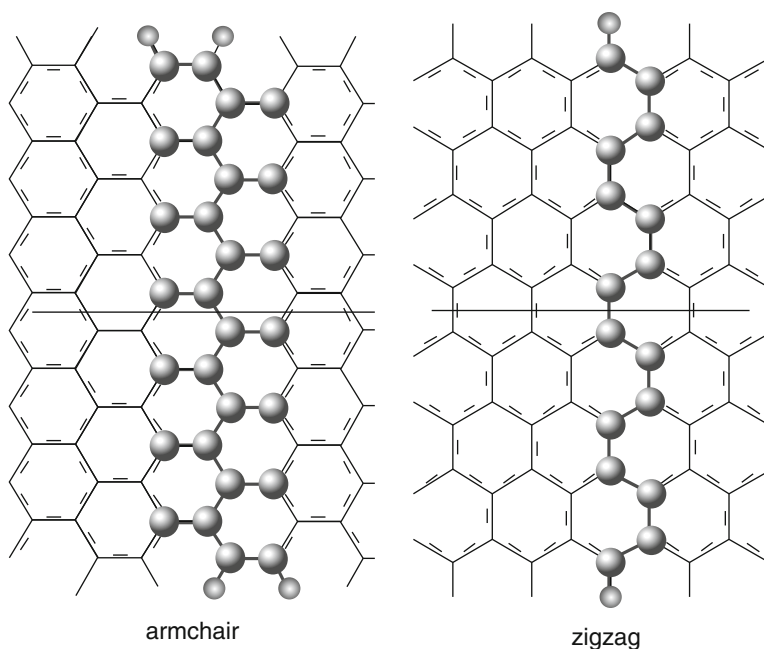


Fig. 2 Schematic representation of the fundamental cell of an armchair and a zigzag graphene nanoribbon. The horizontal line represents the translational vector

performing computational studies as it is very difficult to experimentally determine the chemical nature of the edges.

One of the earliest theoretical investigations of the electronic properties of these materials was reported by Fujita et al. (1996). In that work, the authors report striking differences between armchair and zigzag GNRs. Most notably, the authors found that magnetism can arise in nanometer-scale graphitic fragments in the zigzag configuration due to the appearance of localized edge states. According to TB calculations these edge states are manifested as flat bands near the Fermi energy in the band structure of the system. We will return to the case of ZGNRs later in this section. AGNRs do not present such edge states. Instead, the bandgap of these ribbons is predicted to vary with a threefold periodicity as a function of their width, changing from metallic to semiconducting as one consecutively increases their width (Ezawa 2006). When electronic correlation effects are taken into account via DFT, similar bandgap oscillations appear. Nevertheless, a major difference is identified. While in the TB calculations, at certain widths, AGNRs appear to become metallic, DFT calculations predict all AGNRs to be semiconducting (Barone et al. 2006; Son et al. 2006a). In Fig. 3, we present DFT results for the bandgap of armchair ribbons as a function of their width. As shown in this figure, ultra-narrow graphene nanoribbons with widths up to 1 nm present the largest bandgaps in

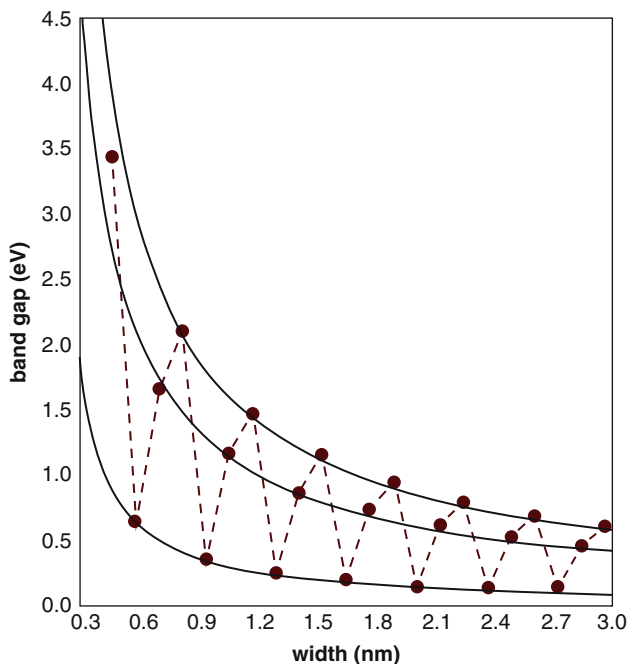


Fig. 3 Bandgap of H-terminated armchair GNRs as a function of the width obtained with the HSE functional (Reprinted (Adapted) with permission from Barone et al. (2006). © (2006) American Chemical Society)

the range of 1–3 eV. For the range of widths covered by the DFT calculations (up to 3 nm), no significant quenching of the energy gap oscillations is observed (Barone et al. 2006). However, GNRs are expected to reach the graphene limit of zero bandgap for sufficiently large widths.

We note that the armchair GNRs bandgaps presented in Fig. 3 can be separated into three groups, namely, the points forming the envelope of the maxima of the oscillations, those forming the envelope of the minima of the oscillations, and the remaining intermediate points. It is possible then to extrapolate the behavior of each of these subgroups independently to larger widths. Such an extrapolation can be performed using an inverse power law with two fitting parameters (Barone et al. 2006). This simple rule presents the correct asymptotic behavior and provides qualitative information on the electronic structure of ribbons beyond the range of widths studied by first-principles calculations. From these calculations, it is shown that in order to obtain an AGNR with a bandgap comparable to that of Ge (0.67 eV) or InN (0.7 eV), it will be necessary to go to a range of widths between 2 and 3 nm. If a larger band gap material is needed (like Si, 1.14 eV, or GaAs, 1.43 eV), the width of the AGNR must be reduced to as low as 1–2 nm. The extrapolation to wider ribbons also shows that an AGNR of about 8 nm will present a bandgap smaller than 0.2 eV. It is interesting to note that when the first DFT calculations on these systems were carried out, it seemed impossible to produce nanoribbons as narrow as a few nanometers. Only 3 years later, the fabrication and electrical measurement of nanoribbons with widths down to a few nanometers are performed routinely in many laboratories around the world (Wang et al. 2008; Yang et al. 2008; Berger et al. 2006; Cai et al. 2010; Han et al. 2007; Ritter and Lyding 2009).

Experimentally, it has been demonstrated that the energy gap in graphene nanoribbons can be tuned during fabrication and that the energy gap scales inversely with the ribbon width (Chen et al. 2007; Han et al. 2007). Suspended GNRs with widths below 10 nm have been synthesized by a chemical route (Wang et al. 2008). However, the uncontrollable character of these methods restrict the quality of the GNR's edges and consequently limit their applications. Novel, organic synthetic protocols leading to graphene type molecules with different sizes have been presented recently (Yang et al. 2008; Cai et al. 2010). This route should provide perfect edged narrow graphene nanoribbons with widths up to 12 nm.

In zigzag graphene nanoribbons, quantum confinement is accompanied by the existence of pronounced edge states localized around the zigzag edges. As mentioned before, the existence of these edge states have been suggested theoretically more than a decade ago (Fujita et al. 1996; Klein 1994; Kobayashi 1993; Nakada et al. 1996; Tanaka et al. 1987). Recently, such states along the zigzag edges of graphite have been experimentally observed (Kobayashi et al. 2005, 2006; Niimi et al. 2005). One of the most interesting characteristics of these edge states in graphite is their predicted magnetic character (Fujita et al. 1996; Kusakabe and Maruyama 2003; Son et al. 2006a; Wakabayashi et al. 1999). Early theoretical predictions of graphene edge states were based on TB and Hubbard models

(Fujita et al. 1996; Nakada et al. 1996; Tanaka et al. 1987), Huckel theory (Klein 1994), and the DV- $X\alpha$ method (Kobayashi 1993; Tsukada et al. 1983). A more chemically oriented interpretation of the structure and electronic character of the different edges has recently been given using Clar's sextet theory, a well-known tool for the study of aromaticity in organic materials (Balaban and Klein 2009; Baldoni et al. 2008; Wassmann et al. 2008).

The magnetic character carried by the zigzag edge states has a dominant effect on the electronic structure of the system. If one views a ZGNR as unrolling an armchair SWNT, one would naively expect both systems to have similar electronic character while replacing the symmetry of a "particle in a ring" type of boundary conditions with that of a "particle in a box." Therefore, since all armchair SWNTs are metallic in nature, it can be assumed that all ZGNRs would be metallic as well. When performing spin-restricted DFT calculations, this indeed turns out to be the case (Son et al. 2006b). Nevertheless, if the spin degree of freedom is taken into account via the unrestricted DFT scheme, the spin-polarized character of the electronic ground state of the ZGNR is revealed (Son et al. 2006b). Here, the electronic edge state on one zigzag edge has one spin flavor while the corresponding state on the other zigzag edge has the opposite spin flavor. This antiparallel edge spin alignment is a direct consequence of the antiferromagnetic coupling of spins on adjacent sites within the hexagonal carbon lattice. In contrast to the case of the metallic armchair SWNTs, the magnetic ground state exhibits a finite bandgap, and therefore, all ZGNRs become semiconducting (Son et al. 2006b). This is an excellent example showing how edge states may dominate the electronic character of the system.

The existence of such spin-polarized edge states opens a venue for controlling the electronic properties of GNRs. One possible scheme was recently suggested where an in-plane electric field, applied perpendicular to the axis of the ribbon, drives the system into a half-metallic state (Son et al. 2006b). Due to the field-induced charge separation, a local gating with opposite charge polarization occurs at the edges of the ribbons. This in turn, shifts the local DOS at both edges with respect to one another thus increasing the bandgap of electrons with one spin flavor and reducing the bandgap of the opposite spins (see Fig. 4). By changing the intensity of the applied electric field, one can control the ratio of the bandgaps of the two spin components up to a point where the bandgap of one spin flavor completely vanishes while the other spin flavor presents a large bandgap. A system in such a half-metallic state may serve as a perfect spin filter in nanospintronic devices.

The half-metallic behavior exhibited by zigzag ribbons under an external electric field was first obtained using the LSDA functional. Shortly after, calculations in a large cluster model of a zigzag GNR performed using the hybrid B3LYP challenged those findings arguing that the inclusion of HF exchange prevents the half-metallic behavior (Rudberg et al. 2007). However, Kan et al. (2007) and Hod et al. (2007a) independently found half-metallicity in periodic zigzag nanoribbons using the B3LYP and HSE functionals, respectively.

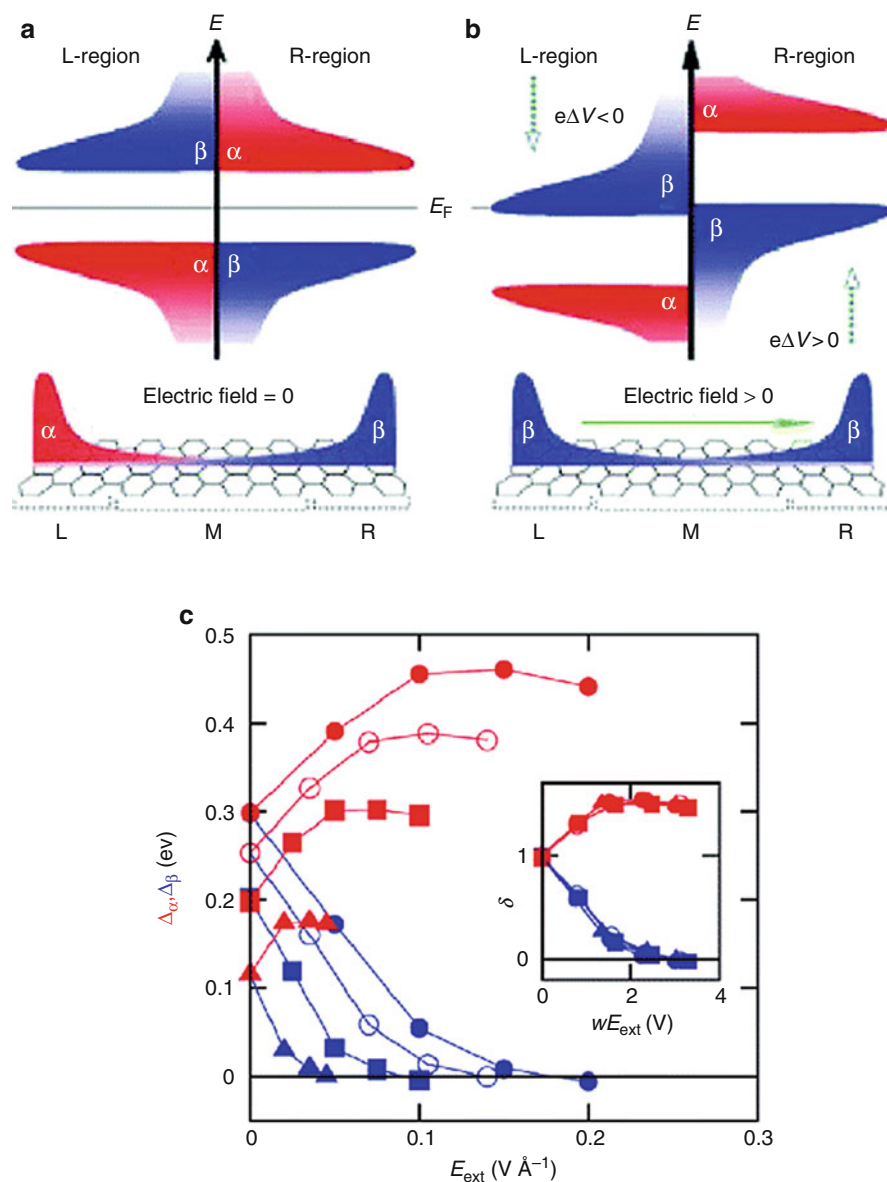


Fig. 4 (a) Schematic density of states diagram of the electronic states of a zigzag GNR in the absence of an applied electric field. (b) Schematic density of states diagram in the presence of a transverse electric field. The electrostatic potential on the *left edge* is lowered, whereas the one on the *right edge* is raised. (c) Dependence of half-metallicity on system size. *Red* denotes the bandgap of α -spin, and *blue* the gap of β -spin as function of E_{ext} for the 8-GNR (*filled circles*), 11-GNR (*open circles*), 16-GNR (*squares*), and 32-GNR (*triangles*). The rescaled gaps for the various widths collapse to a single function as shown in the inset (Reprinted with permission from Macmillan Publishers Ltd: Nature (Son et al. 2006b), © (2006))

Modeling the Optical Spectrum of Single-Walled Carbon Nanotubes and Graphene Nanoribbons

Single-Walled Carbon Nanotubes

Characterization methods for single-walled carbon nanotubes based on optical absorption have attracted much attention during the past few years. Experiments based on photoluminescence (Bachilo et al. 2002; Weisman and Bachilo 2003), resonant Raman spectroscopy (Fantini et al. 2004; Telg et al. 2004), and Rayleigh scattering (Sfeir et al. 2006) have been reported in which optical transitions are obtained as fingerprints of a certain (n, m) SWNT. Optical transitions were first studied theoretically within the tight-binding approach considering the excitations as inter-band transitions. Within this framework, the optical spectra is generally obtained by utilizing the random-phase approximation (RPA) for the imaginary part of the dielectric function ϵ (for a review, see Onida et al. 2002)

$$\text{Im}(\epsilon) = \frac{1}{\omega^2} \sum_k \sum_{o,u} \left| \langle \psi_o^k | \mathbf{p} | \psi_u^k \rangle \right|^2 \delta(\epsilon_0^k - \epsilon_u^k - \omega), \quad (1)$$

where \mathbf{p} is the linear momentum operator and the indices o and u stand for occupied and unoccupied Bloch orbitals, respectively. An example of the band structure and optical spectrum for a semiconducting tube obtained from DFT calculations is shown in Fig. 5. Allowed optical transitions, marked with arrows in the band structure of panel (a), produce a peak in the optical spectrum (panel (b)). Within the RPA the first-order optical transition corresponds to the fundamental gap of the semiconducting SWNT (generally dipole-allowed).

The relation between the optical transitions and the diameter of the nanotubes (Kataura plot) (Kataura et al. 1999) was used as a useful guide for experimentalists in characterizing nanotubes samples. However, Kataura plots based on conventional TB calculations present serious limitations (Bachilo et al. 2002). DFT calculations utilizing local and semi-local functionals are well known to underestimate the bandgap of semiconductors significantly. This holds also true for higher order optical transitions. Hybrid functionals have been shown to improve significantly the description of the bandgap in semiconducting materials (Heyd and Scuseria 2004; Heyd et al. 2005). In the specific case of semiconducting nanotubes, commonly employed hybrid functionals, like B3LYP or PBE0 significantly overestimate the bandgap (Barone et al. 2005b).

Despite the excitonic character of optical transitions in SWNTs (Spataru et al. 2004), the hybrid meta-GGA TPSS and the screened-exchange hybrid HSE provide excellent agreement with experiments for the optical gap (Barone et al. 2005b) as depicted in Fig. 6 where the mean error for the first-order optical transition with respect to experimental values in a set of ten semiconducting nanotubes is shown. It is worth pointing out, however, that DFT-based approaches cannot

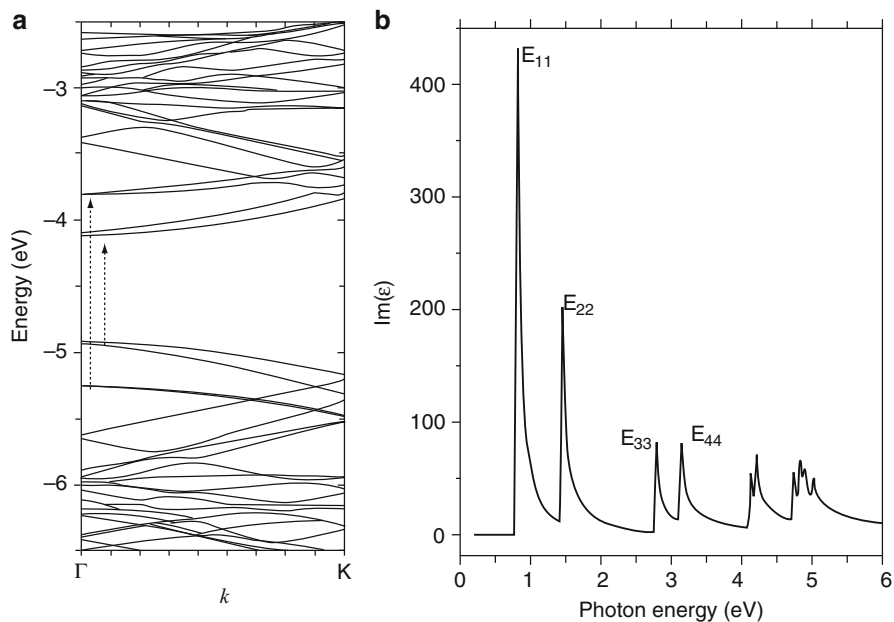


Fig. 5 Band structure (a) and optical spectrum (b) of the semiconducting (10,8) SWNT (Reprinted (Adapted) with permission from Barone et al. (2005b). © (2005) American Chemical Society)

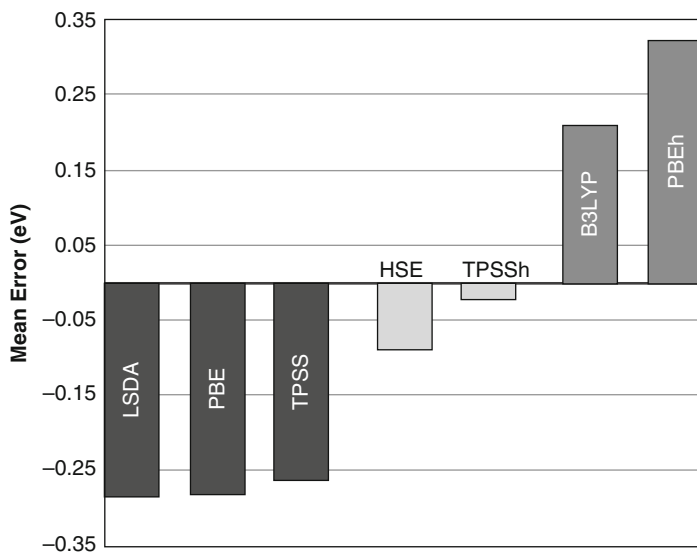


Fig. 6 E_{11} mean errors in a set of ten semiconducting SWNTs calculated with different functionals

predict the strength of exciton binding energies due to the mean-field nature of the approximation.

First-principles calculations including electron–electron interactions beyond the mean field theory have shown the excitonic character of optical transitions in SWNTs with large exciton binding energies (of up to 1 eV for the (8,0) SWNT) (Spataru et al. 2004). These predictions have been corroborated by experiments (Shaver et al. 2007; Wang et al. 2005). Unfortunately, these calculations are too demanding for routine calculations in large diameter tubes, and unlikely to be practical to study the effect of defects and functionalization.

The success of the hybrid functionals HSE and TPSSh in predicting the peak position of optical transitions has been attributed to unknown error cancellations (Kümmel and Kronik 2008; Spataru et al. 2008). However, excitonic effects in metallic nanotubes are up to two orders of magnitude smaller than in semiconducting tubes (Deslippe et al. 2007) and remarkably, the same hybrid functionals that are able to describe the optical peaks in semiconducting tubes also produce excellent results in metallic tubes (Barone et al. 2005a).

As a first example of the predictive capabilities of these functionals, we show in Fig. 7 calculated first-order optical transitions, E_{11} as a function of the corresponding experimental values in a set of five semiconducting and five metallic chiral nanotubes. All non-hybrid functionals employed here (LDA, PBE,

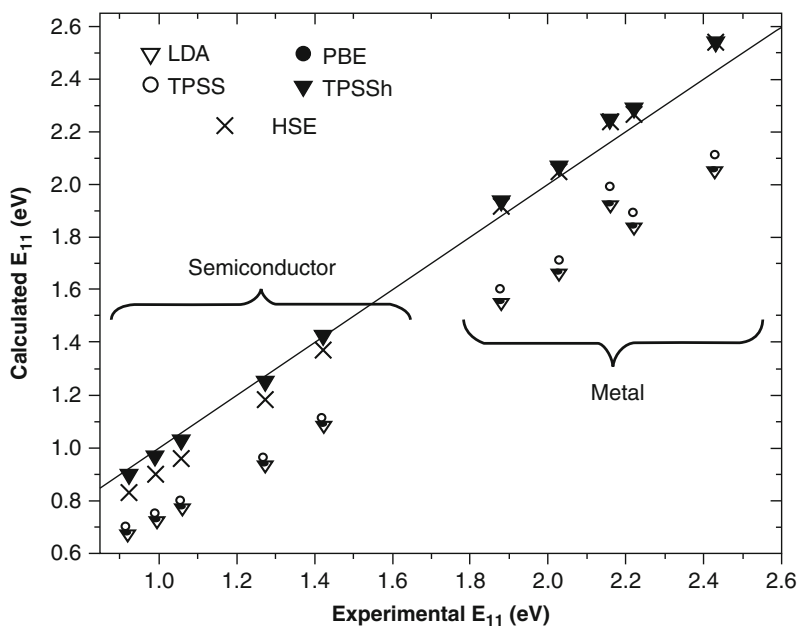


Fig. 7 Calculated versus experimental first optical excitation energies for semiconducting and metallic SWNTs (Reprinted (Adapted) with permission from Barone et al. (2005a). © (2005) American Chemical Society)

Table 2 First-order optical transitions (eV) in metallic and semiconducting tubes calculated using the hybrid TPSSh and HSE functionals (Barone et al. 2005a, b), GW plus electron–hole interactions (GW + e–h) (Spataru et al. 2008), and experimental values

Tube	TPSSh	HSE	GW + e–h	Exp.	Reference (Exp.)
Semiconductor					
(10,0)	1.04	0.97	1.00	1.07	Bachilo et al. (2002)
(11,0)	1.19	1.12	1.21	1.20	Bachilo et al. (2002)
Metallic					
(12,0)	2.25	2.24	2.25	2.16	Fantini et al. (2004)
(10,10)	1.89	1.86	1.84	1.89	Fantini et al. (2004)

and TPSS) underestimate E_{11} in metallic tubes by approximately 0.3 eV. This error is comparable to the error for E_{11} in semiconducting tubes. The best overall performance is achieved by the hybrids TPSSh and HSE, which yield comparable first-order transitions in the case of metallic SWNTs.

As a second example, we compare in Table 2 first-order transitions calculated using the hybrid TPSSh and HSE functionals (Barone et al. 2005a), and calculations considering GW plus electron–hole interactions (GW + e–h) (Spataru et al. 2008), with experimental values. Here, it is worth to point out the results obtained with hybrid functionals, that predict peak positions in agreement with more complex quasiparticle and excitonic effects approaches. An explanation for this behavior has been recently presented by Brothers et al. (2008).

Graphene Nanoribbons

Theoretical and experimental studies reveal that the optical peaks of AGNRs might be utilized as tools to determine the nature of their edges (Barone et al. 2006; Pimenta et al. 2007). The first calculations of the optical spectrum of GNRs was presented by Barone et al. (2006) by means of DFT using the screened-exchange hybrid HSE functional. As expected from an inter-band transitions framework, first optical excitations present the corresponding oscillations as a function of the width. Second-order transitions also exhibit these oscillations, as shown in Fig. 8.

For armchair ribbons with hydrogen terminations there is a shift between the oscillation of the first and second optical transition energies such that the local maxima of E_{11} coincide with the local minima of E_{22} . This is expected to give rise to a doublet in the optical spectrum. A similar doublet is expected to appear for bare GNRs with widths smaller than 1.2 nm. Nevertheless, for larger widths, the bandgap oscillations of the first and second optical transition energies of bare ribbons are in phase and the doublet is expected to disappear. This effect should provide a practical way of revealing information on the size and the nature of the edges of GNRs. Many-body approaches later pointed out that due to the one-dimensional nature of GNRs, and like in the case of SWNTs, their optical spectrum is dominated by excitonic

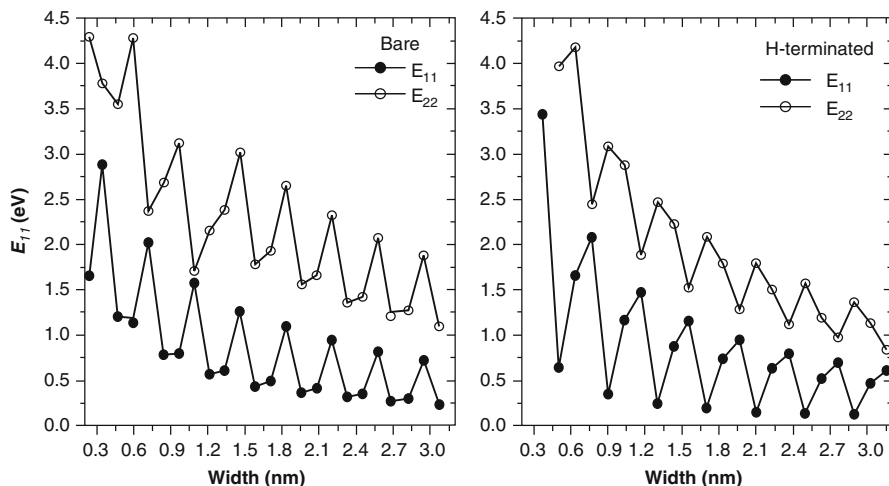


Fig. 8 Dependence of the first (E_{11}) and second (E_{22}) optical transition energies on the width of bare (*left panel*) and hydrogen-terminated (*right panel*) GNRs, at the HSE/6-31G* level of theory (Reprinted (Adapted) with permission from Barone et al. (2006). © (2006) American Chemical Society)

effects with binding energies as high as 1.4 eV for a 1.2 nm wide ribbon (Prezzi et al. 2008; Yang et al. 2007).

Chemistry at the Edges of Graphene

The question of what is at the edges of graphene has fascinated researchers in the area (Radovic and Bockrath 2005), especially, since the possibility of finding magnetic edge states was openly discussed in terms of the type of chemical bond that is expected to appear for bare and hydrogen-terminated edges (Radovic and Bockrath 2005). The question of how magnetic and electronic properties of nanoribbons depend upon chemical functionalization and doping has been partially addressed by several studies. One of the first studies along this line of research was presented by Hod et al. (2007a). In this study, the authors assumed that in most synthetic scenarios GNRs edges will most likely be oxidized with an unknown effect on their electronic properties. The oxidation schemes considered in these calculations included hydroxyl, lactone, ketone, and ether groups. The authors have shown that these oxidized ribbons are, in general, more stable than hydrogen-terminated GNRs (see Fig. 9). These configurations maintain a spin-polarized ground state with antiferromagnetic ordering localized at the edges, similar to the fully hydrogenated counterparts. Edge oxidation has been found to lower the onset electric field required to induce half-metallic behavior (see section “Graphene Nanoribbons”) and extend the overall field range at which the systems remain half-metallic. When the edges of the ribbon are fully or partially hydrogenated, the field intensity needed to switch the

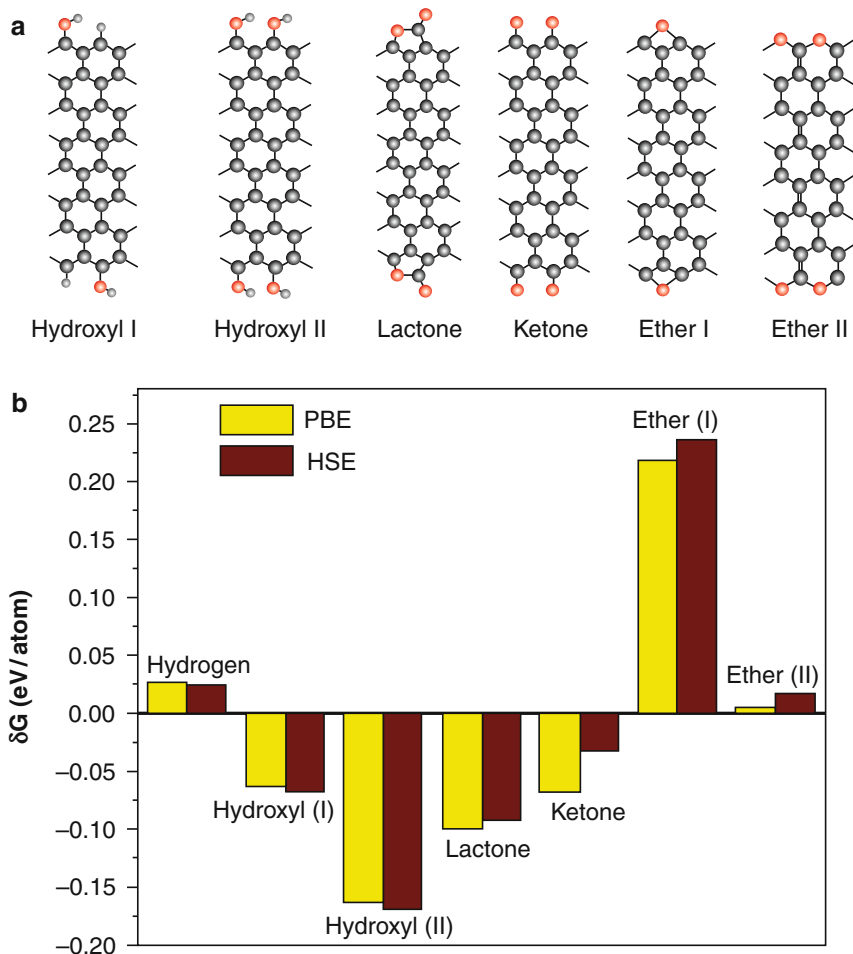


Fig. 9 (a) Scheme of the oxygen-containing groups considered. (b) Relative stability of the different chemical groups at the edges as calculated using the enthalpy of formation with respect to molecular oxygen and hydrogen, and graphene as shown in Hod et al. (2007a) (Reprinted (Adapted) with permission from Hod et al. (2007a). © (2007) American Chemical Society)

system to the half-metallic regime is about 0.4 V/\AA and the range at which the half-metallic behavior is maintained is of 0.3 V/\AA . Nevertheless, when the edges are fully oxidized, the system turns half-metallic at a lower field intensity (0.2 V/\AA) and the range of half-metallic behavior is doubled to 0.6 V/\AA . Unfortunately, it is found that oxygen-containing groups at the edges have a minor effect on the energy difference between the antiferromagnetic ground state and the above-lying ferromagnetic state. This indicates a weak site-to-site exchange coupling and therefore oxygen terminations do not increase the spin coherence length in these systems.

Gunlycke et al. (2007) also studied the effect of oxygen and imine groups and found that transport and magnetic properties are greatly affected by the nature of the chemical groups at the edges. The electronic properties of ribbons with edge carbons presenting two hydrogen or two fluorine atoms were also studied and it was found that in narrow GNRs the sp^3 type of edge carbon results in a non-spin-polarized state (Kudin 2008). Kan et al. (2007) studied the effect of combining donor and acceptor groups at different edges of zigzag nanoribbons using DFT in the GGA. Within this approach, the authors found a half-metallic behavior without an external field for some particular chemical decorations. Cervantes-Sodi et al. (2008a, b) investigated the electronic properties of chemically modified ribbons and observed that chemical modifications of zigzag ribbons can lift the spin degeneracy which promotes a semiconducting-metal transition, or a half-semiconducting state, with the two spin channels having a different bandgap, or a spin-polarized half-semiconducting state, where the spins in the valence and conduction bands are oppositely polarized. The authors find that edge functionalization studied in their work gives electronic states a few eV away from the Fermi level in armchair ribbons and does not significantly affect their bandgap. Lee and Cho (2009) presented calculations based on the local spin density approximation and found that edge-oxidated ZGNRs present a metallic behavior. This was rationalized in terms of the electronegativity of O with respect to C. However, it needs to be stressed that while LSDA or PBE calculations might yield a metallic solution, hybrid functionals still predict all oxidation schemes to be semiconducting (Hod et al. 2007a).

Chemical functionalization of rectangular graphene nanodots has been studied in detail using the DFT-B3LYP approach (Zheng and Duley 2008). This study shows that edge chemical modifications in finite ribbons significantly alter their electronic structure. Finite size effects in this type of dots will be further discussed in the next section.

It is interesting to note that besides chemical functionalization at the edges of graphene it is possible to introduce adatoms on the graphene surface. Usually, this type of interaction is governed by a charge transfer mechanism. Rigo et al. (2009) presented DFT calculations on Ni adsorption on graphene nanoribbons. The adsorption takes place preferentially along the edges of zigzag nanoribbons and the interaction of the adatom with the carbon backbone quenches the magnetization of the latter in the neighborhood of the adsorption site. Other transition metal atoms adsorbed on GNRs were studied using DFT (Sevincli et al. 2008). Interestingly, Fe or Ti adsorption makes certain armchair GNRs half-metallic with a 100% spin polarization at the Fermi level. These results indicate that the properties of graphene nanoribbons can be strongly modified through the adsorption of 3d transition metal atoms. Alkaline and alkali metal adsorption on graphene nanoribbons were found to exhibit a strong site-dependent interaction (Choi and Jhi 2008). Similar results have been obtained by Uthaisar et al. (2009) when considering the interaction of Li adatoms and graphene nanoribbons. The strength of the interaction is much larger in zigzag GNRs than in armchairs and occurs preferentially along the edges. This enhancement is rationalized in terms of the larger number of electron acceptor states in ZGNRs compared to AGNRs. Energy barriers for Li migration also present

important characteristics along the edges which can result in faster kinetics than in regular graphene (Uthaisar and Barone 2010). In addition, recent computational studies suggest lithium doping as a possible route for bandgap engineering of graphitic systems (Krepel and Hod 2011).

Finite Size Effects in Low-Dimensional Graphitic Materials

Finite size effects play a central role in dictating the electronic properties of materials at the nanoscale. Due to their unique electronic structure, quasi-zero-dimensional (quantum dots) graphitic structures may exhibit fascinating physical phenomena, which are absent in their quasi-one-dimensional (nanowires, nanotubes, and nanoribbons) counterparts. Many factors govern the effect of reduced dimensions on the electronic properties of nanoscale materials. Here we focus on two such important factors, which are strongly manifested in the electronic characteristics of graphitic materials, namely, quantum confinement and edge effects:

1. Quantum confinement is related to the boundary conditions enforced on the electronic wave function by the finite size of the system. When the typical de-Broglie wavelength associated with the Fermi electrons becomes comparable to the dimensions of the system, its electronic and optical properties deviate substantially from those of the bulk system. As the confining dimension decreases and reaches this limit (which is typically within the nanometer regime) the energy spectrum turns discrete and the energy gap becomes size dependent.
2. Edge effects in graphitic materials are dominated by localized states, which are physically located at the boundaries of the system and energetically positioned in the vicinity of the Fermi energy. These states influence not only the electronic properties of these systems but also their chemical reactivity.

Quantum Confinement in Graphitic Systems

As shown in previous sections, combining the unique electronic structure of the two-dimensional graphene sheet with the quantum confinement in quasi-one-dimensional graphene derivatives, results in unique electronic properties, which are governed by the specific geometry and dimensions of the relevant system. In section “[Structure–Property Relations in Single-Walled Carbon Nanotubes and Graphene Nanoribbons](#),” it was shown that one of the most notable examples of such effects is found in the strong dependence of the electronic character of carbon nanotubes on their specific diameter and chirality. This diversity in the electronic structure obtained from a single material just by changing its spatial symmetry, is one of the most promising characteristics of carbon nanotubes for applications as basic components in future nanoscale electronic devices.

Recent experimental procedures have allowed the production of ultra-short carbon nanotubes (Chen et al. 2006a, b; Gu et al. 2002; Javey et al. 2004; Khabashesku et al. 2002; Mickelson et al. 1998; Nakamura et al. 2003). The electronic structure of these quasi-zero-dimensional systems is expected to be considerably different from their elongated counterparts, since the reduction of dimensionality implies additional confinement restrictions which may result in the emergence of new and interesting physical phenomena. Several theoretical investigations have addressed the importance of quantum confinement on the electronic properties of finite carbon nanotubes (Baldoni et al. 2007; Li et al. 2002; Liu et al. 2001; Rochefort et al. 1999b). Using a variety of methods including the HF approximation, semiempirical calculations, and the GGA, Rochefort et al. (1999b) have investigated the change in bandgap, density of states, and binding energies as a function of the length of armchair carbon nanotubes. In contrast with the metallic character of the infinite tubes, finite segments shorter than 10 nm are predicted to present a considerable bandgap, which vanishes as the length of the tube increases. Interestingly, the convergence to the infinite metallic system is non-monotonic, and pronounced threefold bandgap oscillations occur as the length of the tube is extended (see Fig. 10). These bandgap oscillations were associated with the periodic changes in the bonding characteristics of the HOMO (highest occupied molecular orbital) and LUMO (lowest unoccupied molecular orbital), which are a direct consequence of the quantum confinement of the π electrons along the tube axis.

As one may expect, the calculated bandgaps strongly depend on the specific computational method, where the HF and semiempirical approximations predict bandgap values which are four to six times larger than those obtained by the generalized gradient DFT and extended Huckel calculations. Nevertheless, the general characteristics of the bandgap behavior and the existence of the bandgap oscillations are predicted by all methods. Similar results have been obtained using the hybrid B3LYP functional for open-ended and capped finite nanotube sections (Li et al. 2002). Expanding the study to the case of zigzag and chiral nanotubes, Liu et al. (2001) have used extended Huckel calculations to show the dependence of the bandgap oscillations on the chirality of the finite nanotube segment. As shown in Fig. 11, as the chirality of the tube changes gradually from armchair to zigzag, the amplitude of the oscillations reduces, and almost vanishes for the case of (12,0) zigzag nanotube segments. This leads to the interesting conclusion that the metallic character of the infinite armchair nanotubes is replaced by HOMO–LUMO gap oscillations for finite armchair nanotube segments, while the oscillatory bandgap nature of infinite zigzag nanotubes is replaced by a vanishing HOMO–LUMO gap for the finite zigzag nanotube segments. An interesting interpretation of the behavior of the HOMO and LUMO levels of finite carbon nanotube segments as a function of their length was recently given using Clar sextet theory (Baldoni et al. 2007).

As discussed above, when an infinite graphene sheet is cut to form a quasi-one-dimensional graphene nanoribbon with a finite width and infinite length, the π -electrons wave function is confined along the direction perpendicular to the axis of the ribbon and is forced to vanish at large distances along this direction. These “particle in a box” like boundary conditions induce discretization of the

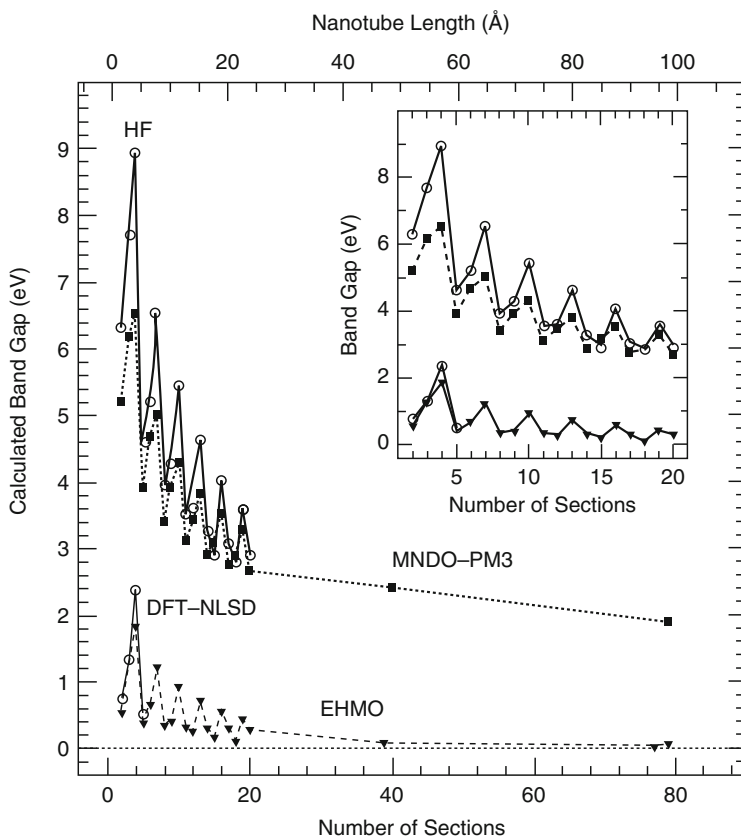


Fig. 10 Variation of the band gap of a (6,6) nanotube as a function of its length determined using different computational techniques (Reprinted (Adapted) with permission from Rochefort et al. (1999b). © (1999) American Chemical Society)

two-dimensional dispersion relation into a set of one-dimensional bands. This discretization induces bandgap dependence on the width of the obtained ribbons.

The dimensionality of these systems may be further reduced to form graphene nanodots, which can be viewed as molecular derivatives of graphene. This extra confinement has been recently shown to strongly impact the electronic structure of these systems (Hod et al. 2008; Shemella et al. 2007). Figure 12 presents the HOMO-LUMO gaps as a function of the length and width of a large number of graphene quantum dots calculated using the local density approximation (upper left panel), the PBE flavor of the generalized gradient correction (upper right panel), and the screened-exchange hybrid HSE functional (lower left panel). The studied graphene derivatives are rectangular in shape and denoted by $N \times M$ where N and M are the number of hydrogen atoms passivating the armchair and zigzag edge, respectively. As in the case of infinite armchair graphene nanoribbons, an oscillatory

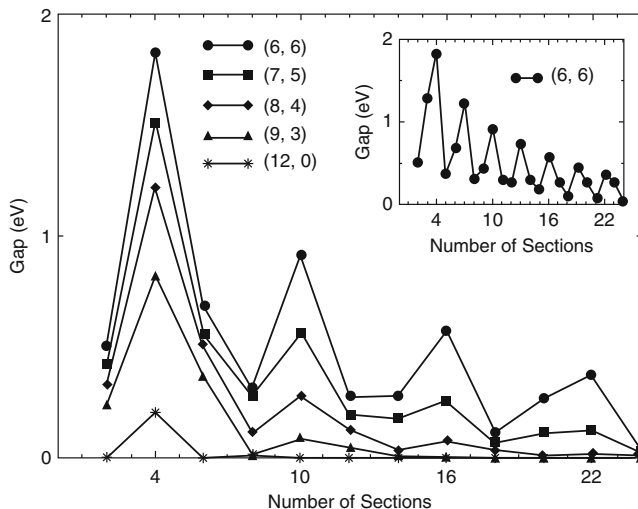


Fig. 11 The bandgap as a function of the (even) number of SWNTs sections (Reprinted with permission from Liu et al. (2001). © (2001) by the American Physical Society)

behavior of the energy gap as a function of the length of the ribbon is observed. The periodicity of these oscillations appears to be somewhat different than the threefold period obtained for the infinitely long counterparts. This, however, is a result of the fact that in order to prevent dangling carbon bonds, the width step taken for the finite systems is twice as large than the one taken in the infinitely long armchair ribbon calculations. The amplitude of the oscillations is found to be considerably damped due to the finite size of the ribbons. As expected, when the length of the armchair edge (N) is increased, the oscillation amplitude increases as well. It is interesting to note that, in general, the HOMO–LUMO gap is inversely proportional to the width (N) and the length (M) of the finite GNR in accordance with the semimetallic graphene sheet limit. Therefore, in order to obtain energy gap tailoring capability, one will have to consider GNRs with long armchair edges (large N values) and short zigzag edges (small M values). This will increase the amplitude of the energy gap oscillations while maintaining overall higher gap values.

Edge Effects in Graphitic Systems

The effects of localized edge states on the electronic properties of quasi-one-dimensional systems have been discussed in previous sections. The question arises of whether similar effects can be observed in quasi-zero-dimensional systems. It is well established that small molecular derivatives of graphene, such as different types of polyaromatic hydrocarbons, have a closed shell nonmagnetic ground state. On the other hand, infinite ZGNRs present a spin-polarized ground state. This suggests

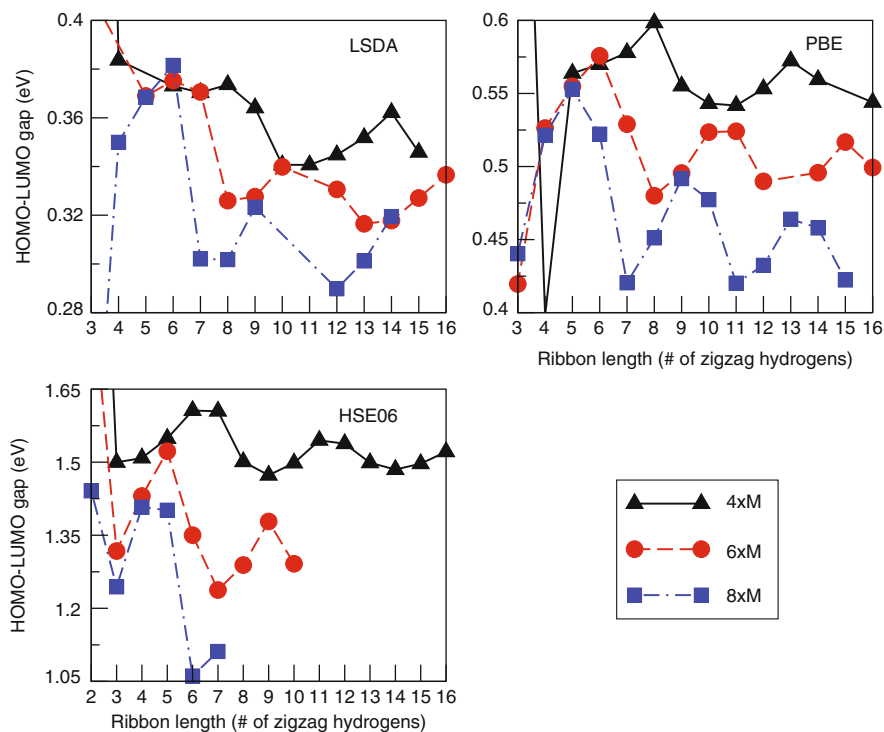


Fig. 12 HOMO–LUMO gap values for three sets of graphene nanodots, as calculated by the local spin density approximation (*upper left panel*), PBE functional (*upper right panel*), and the HSE06 functional (*lower left panel*) (Reprinted with permission from Hod et al. (2008). © (2008) by the American Physical Society)

that there exists a critical size at which molecular graphene derivatives become spin polarized. One of the earliest studies addressing this issue dates back to more than two decades ago (Stein and Brown 1987). Using Huckel theory it was found that finite graphene flakes exhibit electronic edge states along the zigzag edges. More recently, several studies, based on DFT calculations have investigated this question in detail. As a first step for obtaining quasi-zero-dimensional graphene derivatives, it is convenient to consider truncated graphene nanoribbons in the form of rectangular nanodots (Hod et al. 2008; Jiang et al. 2007; Kan et al. 2007; Rudberg et al. 2007; Shemella et al. 2007). Surprisingly, it was found that even molecular scale graphene derivatives, such as the bisanthrene (phenanthro[1,10,9,8-opqra]perylene) isomer of the $C_{28}H_{14}$ molecule and $C_{36}H_{16}$ (tetrabenzo[bc,ef,kl,no]coronene), are predicted to present a spin polarized ground state as shown in Fig. 13 (Hod et al. 2008; Jiang et al. 2007; Kan et al. 2007). For the smaller molecules, this was further verified using a complete active space self-consistent field many-body-wave function approach (Jiang et al. 2007). Furthermore, the application of an in-plane electric field perpendicular to the zigzag edge was found to turn the finite systems

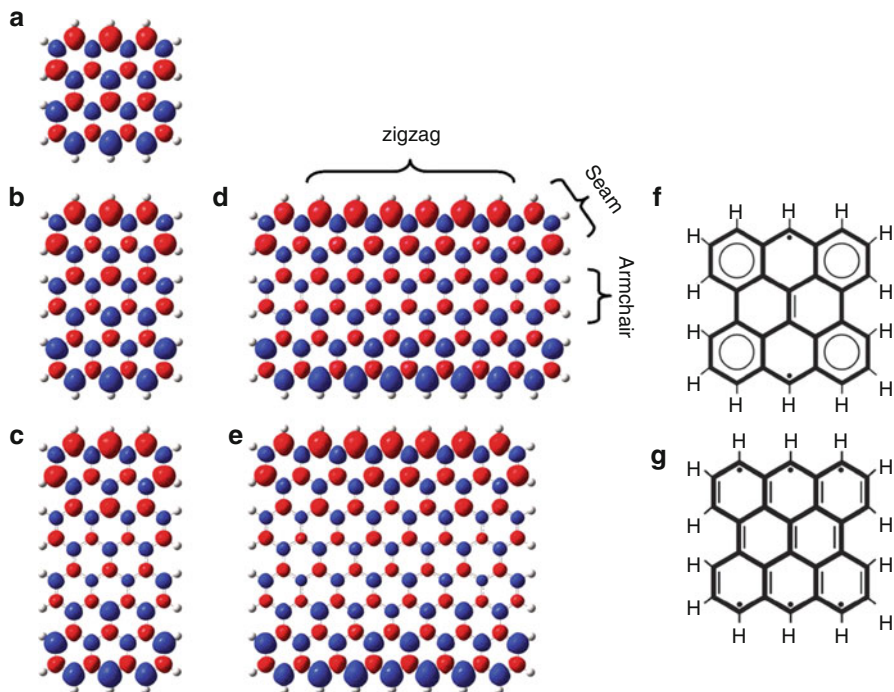


Fig. 13 (a)–(e) Isosurface spin densities of the antiferromagnetic ground state of several molecular graphene derivatives as obtained using the HSE functional and the 6-31G** basis set. (f) and (g) represent diradical and hexaradical Clar structures of bisanthrene (Reprinted with permission from Hod et al. (2008). © (2008) by the American Physical Society)

into molecular scale half-metals (Hod et al. 2008; Kan et al. 2007), similar to the case of infinite ZGNRs.

When considering more complicated graphene derivatives such as triangular graphene flakes (Ezawa 2007; Fernandez-Rossier and Palacios 2007), the combination of the local antiferromagnetic spin ordering on adjacent carbon sites and the edge geometry of the triangular structure results in spin frustration. This leads to a metallic ground state with an overall ferromagnetic character and a finite magnetic moment. The results of tight-binding Ising model (Ezawa 2007) and Hubbard model (Fernandez-Rossier and Palacios 2007) Hamiltonians are consistent with the predictions of DFT (Fernandez-Rossier and Palacios 2007). These can be further rationalized by Lieb's theorem (Lieb 1989) regarding the total spin S of the exact ground state of the Hubbard model in bipartite lattices. The honeycomb lattice of graphene is formed by two triangular interpenetrating sublattices, A and B . Since triangular nanostructures have more atoms in one sublattice, $N_A > N_B$ the total spin S of the ground state is $2S = N_A - N_B > 0$. Here, the main contribution to the magnetic moment comes from edge states around the zigzag edges of

the triangular structure. Such molecular graphene derivatives were suggested to function as permanent magnets in future nanoscale memory devices.

Naturally, the effect of edge states reduces as the surface to edge ratio becomes larger. An interesting question to address is at what length would the effect of the edges disappear and the electronic structure of the system becomes essentially identical to that of the infinite system? In a recent study, this question was addressed using a divide-and-conquer (D&C) DFT approach, which enables the efficient and accurate calculation of the electronic properties and charge transport through finite elongated systems (Hod et al. 2006). Within this approach, the Hamiltonian \mathbf{H} is given in a localized basis set representation by a block-tridiagonal matrix, where the first and last diagonal blocks correspond to the two terminating units of the ribbon (see upper panel of Fig. 14). The remaining diagonal blocks correspond to the central part of the GNR which is composed of a replicated unit cell. The terminating units and the replicated central part unit cell are chosen to be long enough such that the block-tridiagonal representation of \mathbf{H} (and the overlap matrix \mathbf{S}) is valid.

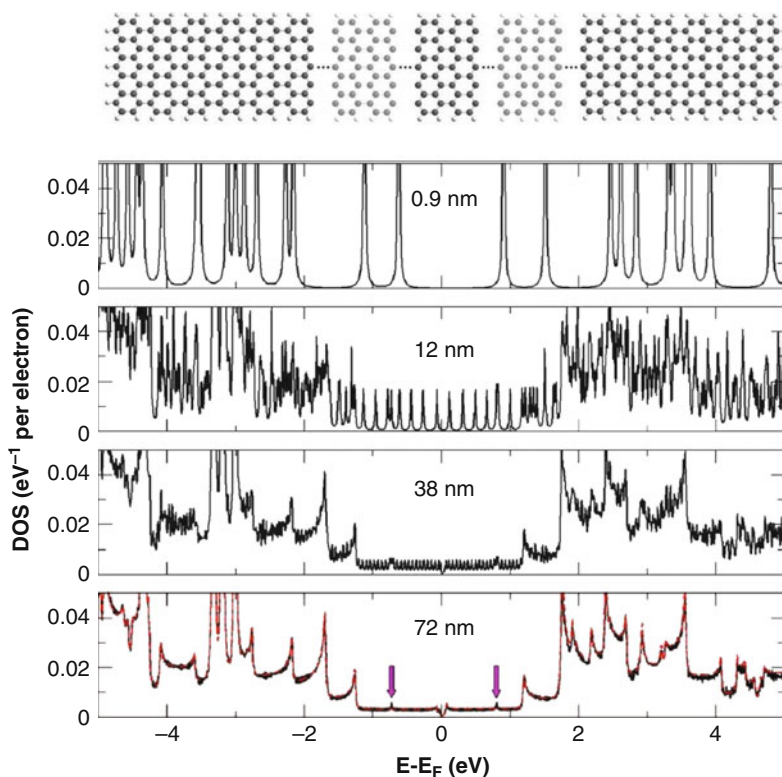


Fig. 14 *Upper panel:* A finite elongated graphene nanoribbon. *Lower panels:* DOS of the finite GNR at different lengths. The red curve in the lowermost panel is the DOS of the infinite system (Reprinted with permission from Hod et al. (2007b). © (2006) by the American Physical Society)

The terminating unit diagonal Hamiltonian blocks and their coupling to the central part are evaluated via a molecular calculation involving the two terminating units and one unit cell cut out of the central part. The replicated unit cell blocks of the central part and the coupling between two such adjacent blocks are approximated to be constant along the GNR and obtained from a periodic boundary conditions calculation. The resulting block-tridiagonal matrix $ES - \mathbf{H}$ is then partially inverted for each value of the energy E using an efficient algorithm, to obtain the relevant Green's function blocks needed for the DOS and transport calculations using the following formula

$$\rho(E) = -\frac{1}{\pi} \text{Im} \{ \text{Tr} [G^r(E)\mathbf{S}] \}, \quad (2)$$

where $G^r(E) = [ES - \mathbf{H}]^{-1}$. A detailed account of the D&C method can be found in Hod et al. (2006).

To study these edge effects in graphene, three graphene nanoribbons with consecutive widths were considered and their DOS as a function of the ribbons' length were calculated using this D&C approach (Hod et al. 2007b). As an example, Fig. 14 presents the DOS of a finite graphene nanoribbon of different lengths. It can be seen that for graphene nanodots up to 12 nm in length, the DOS resembles that of a finite molecular system characterized by a discrete set of energy levels. When the ribbon is further elongated to 38 nm, a constant DOS around the Fermi energy of the infinite system arises and typical Van Hove singularities start to build up. For this specific system, at a length of 72 nm most of the features that are related to the edge states disappear, and the DOS almost perfectly matches that of an infinite armchair graphene nanoribbon. The rate of convergence of the electronic properties of finite systems to those of the infinite counterparts will depend on the type of the system and the specific interactions that govern its electronic character. Therefore, the divide-and-conquer method allows the quantitative estimation of this rate based on state-of-the-art DFT approximations.

We now turn back to the case of carbon nanotubes. When a zigzag nanotube is cut into finite segments, the zigzag edges are exposed. We have shown in the previous section that quantum confinement effects may play an important role in determining the electronic character of such structures. Similar to the case of graphene nanoribbons, these effects are now accompanied by the formation of spin-polarized edge states which put their own fingerprints on the electronic structures of the system (see Fig. 15). There has been a debate in the literature regarding the energetic stability of these edge states, and whether they would appear for all finite zigzag nanotubes regardless of their diameter (Higuchi et al. 2004; Kim et al. 2003; Okada and Oshiyama 2003). It is now well accepted that all finite zigzag SWNT segments are expected to present a spin-polarized ground state (Hod and Scuseria 2008; Kim et al. 2003; Mananes et al. 2008). Furthermore, as for ZGNRs, the application of an axial electric field will drive the system into a half-metallic state, thus forming a perfect spin filter. If one now eliminates only one zigzag edge of the tube by capping it with a half-fullerene, spin frustration results in a

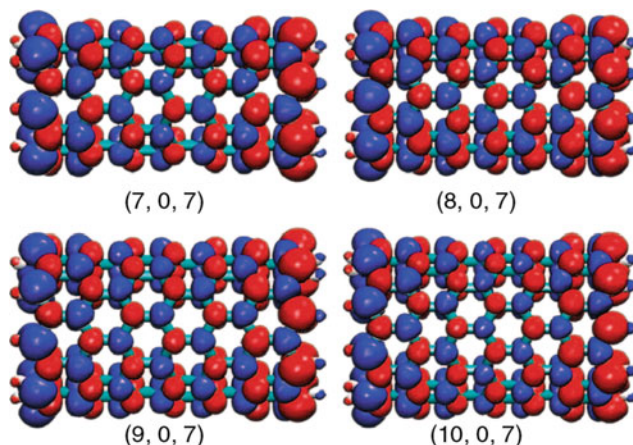


Fig. 15 Antiferromagnetic-type ground state spin density maps of the (7,0) (*upper left panel*), (8,0) (*upper right panel*), (9,0) (*lower left panel*), (10,0) (*lower right panel*) finite zigzag SWNT segments as obtained using the HSE functional with the 6-31G** basis set. Red and blue isosurfaces indicate the two spin flavors with an isovalue of $0.0015 a_0^{-3}$ (Reprinted (Adapted) with permission from Hod and Scuseria (2008). © (2008) American Chemical Society)

spin-polarized ferromagnetic ground state (Kim et al. 2003), which resembles the case of triangular graphene derivatives discussed above and forms a molecular magnet bearing a permanent magnetic moment.

Two-dimensional graphene and its lower dimensional derivatives present a diversity of electronic behaviors, making them particularly attractive as building blocks for future nano-devices. Simplified model Hamiltonian approaches may give important insights on the general physical trends in these systems. Nevertheless, density functional theory in general, and the screen exchange hybrid functional approximation, in particular, seem to be excellent tools to quantitatively study the structure–function relations in these systems and the effects of external perturbations such as chemical substitutions and electric and magnetic fields.

Electromechanical Properties of One-Dimensional Graphitic Structures

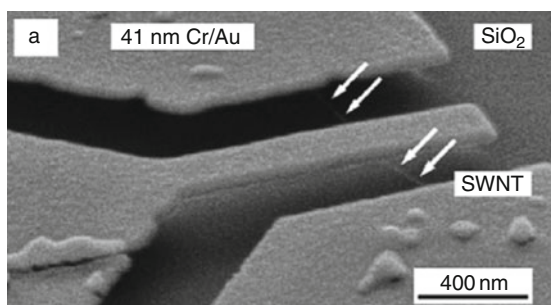
Electromechanical devices are based on systems for which the mechanical properties can be controlled via the application of external electric potentials and/or the electronic properties may be altered via induced mechanical deformations. Such devices can be scaled down to form microelectromechanical systems (MEMS), which are small integrated devices or systems that combine electrical and mechanical components. Their dimensions may vary from the submicron level up to a few millimeters. By fabricating miniature mechanical elements such as beams, gears, diaphragms, and springs MEMS enabled the realization of diverse applications

including ink-jet-printer cartridges, accelerometers, miniature robots, microengines, locks, inertial sensors, microtransmissions, micromirrors, micro actuators, optical scanners, fluid pumps, transducers, and chemical, pressure, and flow sensors. Nanoelectromechanical (NEMS) systems present the ultimate miniaturization of such devices, scaling them further down to the molecular level. Apart from the reduced dimensions, these systems are characterized by lower energy consumption and increased sensitivity toward external perturbations. Furthermore, due to their quantum-mechanical nature, molecular-sized NEMS may present unique physical properties that cannot be obtained from their microscopic and macroscopic counterparts. The unique electronic and mechanical properties of carbon nanotubes have marked them as promising candidates for key components in NEMS. In recent years, several theoretical investigations as well as experimental realization of NEMS have been presented (Cao et al. 2003; Cohen-Karni et al. 2006; Fennimore et al. 2003; Gomez-Navarro et al. 2004; Hall et al. 2007; Maiti 2003; Minot et al. 2003; Nagapriya et al. 2008; Paulson et al. 1999; Rueckes et al. 2000; Sazonova et al. 2004; Semet et al. 2005; Stampfer et al. 2006; Tombler et al. 2000).

Carbon Nanotubes in NEMS Applications

The key components in most of the suggested setups are suspended carbon nanotubes bridging the gap between two conducting electrodes without making contact with the underlying surface (see Fig. 16). One may then induce mechanical deformations via manipulation of an external nanoscale tip while simultaneously measuring the changes in the conductance of the system. Strong reversible electromechanical response of carbon nanotubes under combined bending and stretching deformations have been recorded by several experimental groups (Cao et al. 2003; Maiti 2003; Minot et al. 2003; Semet et al. 2005; Stampfer et al. 2006; Tombler et al. 2000). In Fig. 17, the change in the SWNT's resistance as a function of the depression applied to the top cantilever (Fig. 16) is plotted. The resistance of the system increases by more than an order of magnitude over a depression range of less than 4 nm, suggesting that such setups may be used for highly sensitive and reliable displacement sensors. A similar setup, where a floating pedal is attached to a suspended

Fig. 16 SEM image of a cantilever-SWNT-based NEMS sensing device. The *white arrows* indicate the location of the suspended SWNT (Reprinted (Adapted) with permission from Stampfer et al. (2006). © (2006) American Chemical Society)



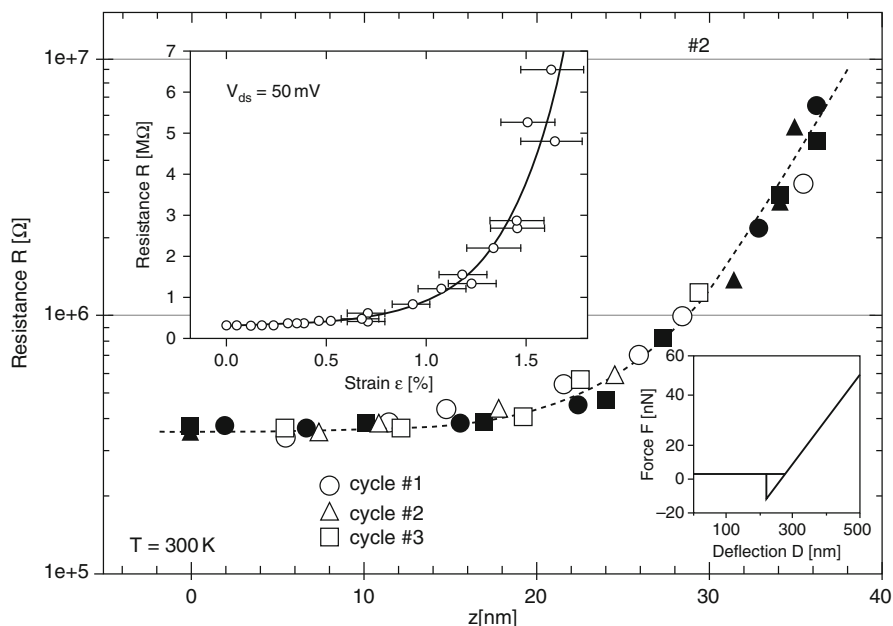


Fig. 17 Electromechanical measurements of an electromechanical device based on a metallic SWNT (Fig.16). The *upper insert* shows the resistance plotted as a function of strain. The *lower insert* shows a force versus deflection measurement performed on the cantilever-SWNT contact point (Reprinted(Adapted) with permission from Stampfer et al. (2006). © (2006) American Chemical Society)

multi-walled nanotube, has enabled the investigation of the electromechanical response of single (Hall et al. 2007) and multi-walled (Cohen-Karni et al. 2006) nanotubes under the application of torsional deformations. In this setup, an atomic force microscope (AFM) tip presses down on the floating pedal. Since the SWNT is fixed at both edges to the conducting electrodes, the central section experiences torsion. This change in helicity of the system induces considerable variations in the electronic conductance. Interestingly, pronounced bandgap oscillations are observed as a function of the deflection angle of the pedal. The origin of these oscillations has been associated with the distortion of the first Brillouin zone (fBZ), which results in the shifting of the Dirac points (Cohen-Karni et al. 2006; Nagapriya et al. 2008). As the Dirac points shift within the fBZ they may approach an allowed electronic band thus increasing the conductance of the system. Once the Dirac point crosses the allowed band the conductance reaches a maximum value. Further deformation results in shifting the Dirac points away from the allowed band and thus a decrease in the conductance. This process is repeated anytime the Dirac point crosses one of the many parallel allowed bands. Other prototype devices such as nanotube actuators (Fennimore et al. 2003; Gomez-Navarro et al. 2004), and nonvolatile memory components based on SWNTs junctions (Rueckes et al. 2000) have been

successfully fabricated as well, demonstrating the potential of SWNTs as nanoscale electromechanical devices. Aiming to obtain a microscopic understanding of these experimental observations and to guide further experiments with carbon nanotubes under various mechanical deformation modes, many theoretical investigations have been performed (Maiti 2009). Due to the complexity of the deformed systems, most of the computational treatments were based on semiempirical or DFT-based TB calculations.

The effects of axial stretching (Heyd et al. 1997; Jiang et al. 2004; Kane and Mele 1997; Kleiner and Eggert 2001; Yang et al. 1999), bending (Kane and Mele 1997; Liu et al. 2000; Nardelli 1999; Nardelli and Bernholc 1999; Rochefort et al. 1998, 1999a; Yang and Han 2000), torsion (Jiang et al. 2004; Kleiner and Eggert 2001; Rochefort et al. 1999a; Yang and Han 2000; Yang et al. 1999; Zhang et al. 2009), and radial compression (Lammert et al. 2000; Lu et al. 2003; Peng and Cho 2002; Svizhenko et al. 2005) were studied extensively showing high sensitivity of the electronic properties and transport characteristics of the deformed nanotubes to the applied deformations.

Several DFT calculations were conducted to study the effect of axial strain (Guo et al. 2005), bending (Maiti 2001; Maiti et al. 2002), radial compression (Mehrez et al. 2005; Wu et al. 2004), and torsional deformations in carbon nanotubes (Nagapriya et al. 2008):

1. *Axialstrain*: While isolating the effect of axial stretching and compression is a challenging experimental task (Cao et al. 2003), pure axial deformations are easy to study using modern computational tools (Guo et al. 2005). Bandgap variations due to axial compression and stretching were studied using the local density approximation. As it can be seen in Fig. 18 for strains up to 3%, linear variations in the bandgap are predicted. The slope of the bandgap response depends on the tube type and on the sign of the axial deformation. For all systems studied, the relative changes in bandgap are of the order of 0.5 eV and are definitely measurable in experimental conditions.
2. *Bending*: The effect of pure bending and AFM tip-induced bending was studied using the Harris functional approximation (Maiti 2001; Maiti et al. 2002). It was shown that both pure bending and tip-induced compression maintain the hexagonal sp^2 lattice structure up to relatively high bending angles. Armchair tubes remain significantly conducting even at large deformations. However, metallic zigzag tubes display a dramatic drop in conductance, in particular under tip-induced deformations.
3. *Compression*: Radial compression was studied via a combined molecular dynamics and density functional theory-based nonequilibrium Green's function approach (Wu et al. 2004). Reversible pressure-induced metal-to-semiconductor transitions of armchair SWNTs were predicted suggesting that SWNTs may be used as miniature sensitive pressure detectors (Mehrez et al. 2005; Wu et al. 2004).
4. *Torsion*: While the torsional electromechanical response was successfully explained via the deformation of the Brillouin zone, another possible explanation

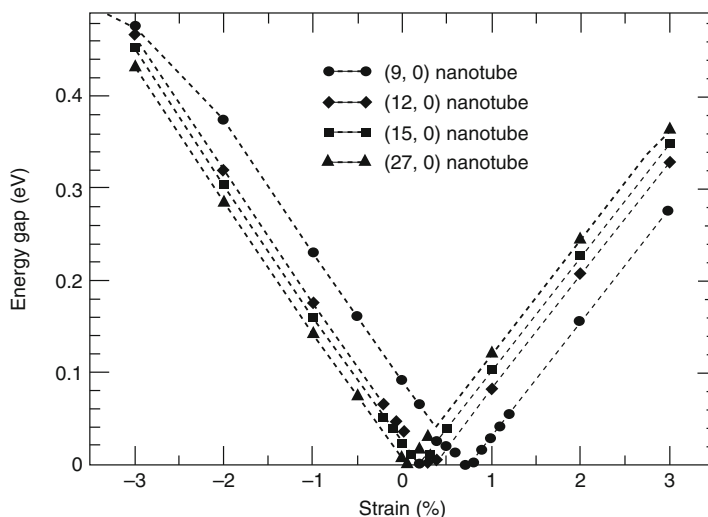


Fig. 18 Calculated bandgap variation of zigzag (9,0), (12,0), (15,0), and (27,0) SWNTs as a function of axial strain (Reprinted with permission from Guo et al. (2005). © (2005) American Institute of Physics)

for the bandgap oscillations may be the periodic formation of Moire patterns due to registry mismatch between the different walls of the multi-walled tubes, which have a different curvature. In order to rule out this possibility, DFT calculations within the LSDA were used to show that the interlayer coupling is weak and the electronic structure of the individual walls resembles that of single-walled nanotubes for different intertube orientations (Nagapriya et al. 2008).

As can be seen from all of the examples given above, the study of the physical properties of nanotubes in general and their electromechanical behavior in particular requires an intimate relation between theory and experiment, where the theoretical tools enable an atomic scale understanding of the experimental observations and provide guidelines for the design of new experiments. As we have seen above, when considering finite size effects and chemical functionalization, DFT presents new insights that are not captured within the simple TB approximations. When considering the electromechanical response of SWNTs to different mechanical deformations, TB calculations seem to give a reliable description. Nevertheless, calculations based on DFT are required for comparison and validation purposes.

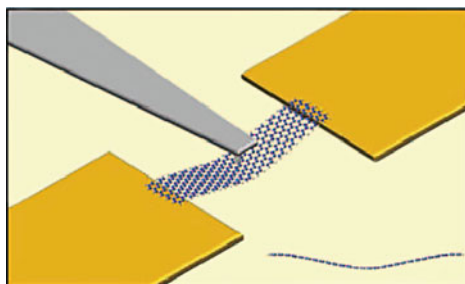
Graphene Nanoribbons in NEMS Applications

Soon after the successful isolation of a single graphene sheet (Novoselov et al. 2004) the first experimental realization of an electromechanical device based on graphene nanoribbons was presented (Bunch et al. 2007). In this setup, atomic

thick layers of graphene were suspended above pre-designed trenches carved in the underlying silicon oxide surface while bridging the gap between gold electrodes. Both optical and electrical actuation procedures were demonstrated resulting in vibrational frequencies in the MHz regime. High charge sensitivities suggested the possibility of using similar devices as ultrasensitive mass and force detection (Bunch et al. 2007). Following this novel experimental fabrication and manipulation of graphene nanoribbons as nanoelectromechanical components, several other studies have explored the mechanical (Frank et al. 2007; Poot and van der Zant 2008) and electromechanical response of these systems (Milaninia et al. 2009). The effects of uniaxial strains in isolated graphene nanoribbons were recently studied in details using the generalized gradient approximation of the exchange correlation functional within the framework of DFT (Faccio et al. 2009; Sun et al. 2008). It was found that the electronic properties of zigzag GNRs are not sensitive to uniaxial strain, while the energy gap of armchair nanoribbons displays an oscillatory pattern as a function of the applied strain. By comparison to TB calculations, it was deduced that the nearest-neighbor hopping terms between atomic sites within the carbon hexagonal lattice are responsible for the observed electromechanical response. To simulate the effects of bending and torsional deformations in suspended graphene nanoribbons, DFT calculations utilizing the screened-exchange HSE density functional were performed (Hod and Scuseria 2009). A large set of short armchair graphene nanoribbons was considered, where narrow strips of atoms close to the zigzag edges of the ribbon were fixed to simulate a doubly clamped suspended nanoribbon (see Fig. 19).

The effect of a driven deformation due to an external tip were simulated by constraining a strip of hexagons at the central part of the ribbon to be either depressed or rotated with respect to the fixed edge atoms. Extreme mechanical deformations, way beyond the linear response regime, were applied to the full set of nanoribbons studied. Most of the nanoribbons considered stayed stable under depressions of 1 nm as well as torsional angles of 90° maintaining their spin-polarized ground state character. As it can be seen in Fig. 20, pronounced electromechanical responses for both bending and torsion (torsion not presented in Fig. 20) have been obtained including evidence of dimension-dependent bandgap oscillations similar to the case of SWNTs (Cohen-Karni et al. 2006; Nagapriya et al. 2008).

Fig. 19 An artist view of an electromechanical device based on graphene nanoribbons. An external tip may induce bending and torsion of the nanoribbon (Reprinted (Adapted) with permission from Hod and Scuseria (2009). © (2009) American Chemical Society)



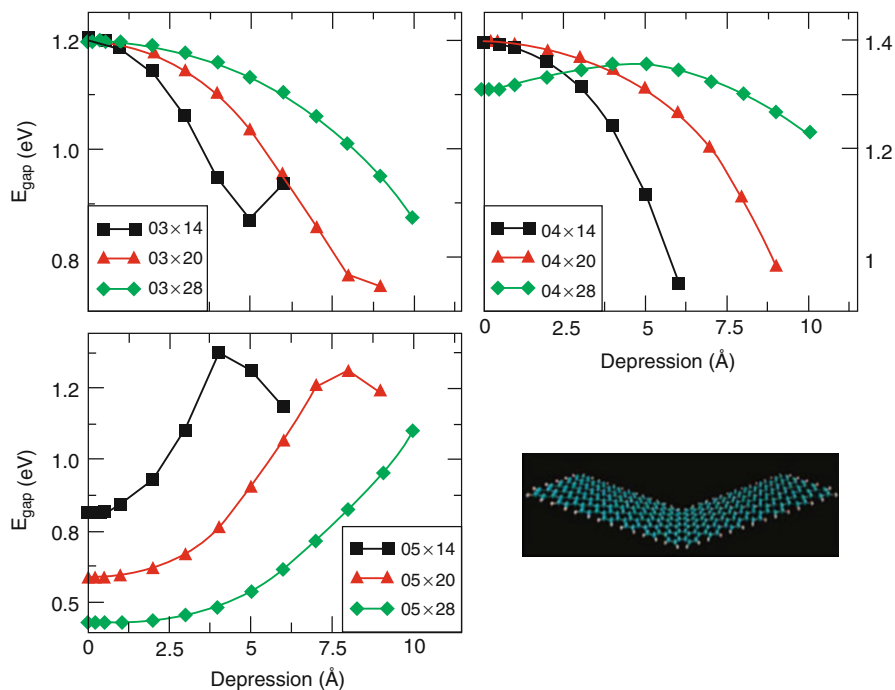


Fig. 20 HOMO–LUMO gap variations as a function of the depression applied to the central region of a doubly clamped GNR. Different panels and *line colors* represent results for GNRs of different widths and lengths, respectively (Reprinted (Adapted) with permission from Hod and Scuseria (2009). © (2009) American Chemical Society)

Apart from indicating the promise that graphene nanoribbons hold from an electromechanical perspective, such calculations based on state-of-the-art density functional theory approaches, enhance the molecular scale understanding of the physical processes governing the behavior of these intriguing materials. Furthermore, theory and computations may guide future experiments to pursue promising scientific and technological routes toward which efforts and resources should be directed.

Concluding Remarks

In this chapter, we have given several examples of how theoretical investigations can be applied to elucidate the behavior of carbon nanostructures with emphasis on the understanding of the basic physical mechanisms that take place at the molecular level. In particular, we have shown that density functional theory is a powerful tool to this end, and we have provided several examples where density functional theory has been utilized to investigate electronic and structural properties of graphene

nanoribbons, carbon nanotubes, and other low-dimensional graphitic materials. We hope that this chapter will furnish the reader with an ample background on this growing field and that it will serve as a starting point for readers interested in pursuing research in this exciting field.

Acknowledgments VB acknowledges the donors of The American Chemical Society Petroleum Research Fund for support through the award ACS PRF#49427-UNI6. OH acknowledges the support of the Israel Science Foundation (Grant 1313/08), the Center for Nanoscience and Nanotechnology at Tel-Aviv University, the Lise Meitner-Minerva Center for Computational Quantum Chemistry, and the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 249225. JEP acknowledges support from NSF DMR Award #DMR-0906617.

Bibliography

- Adamo, C., & Barone, V. (1999). Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of Chemical Physics*, *110*, 6158–6170.
- Adamson, R., Dombroski, J., & Gill, P. (1996). Chemistry without Coulomb tails. *Chemical Physics Letters*, *254*, 329–336.
- Adamson, R., Dombroski, J., & Gill, P. (1999). Efficient calculation of short-range Coulomb energies. *Journal of Computational Chemistry*, *20*, 921–927.
- Bachilo, S. M., Strano, M. S., Kittrell, C., Hauge, R. H., Smalley, R. E., & Weisman, R. B. (2002). Structure-assigned optical spectra of single-walled carbon nanotubes. *Science*, *298*, 2361–2366.
- Balaban, A. T., & Klein, D. J. (2009). Claromatic carbon nanostructures. *Journal of Physical Chemistry C*, *113*, 19123–19133.
- Baldoni, M., Sgamellotti, A., & Mercuri, F. (2007). Finite-length models of carbon nanotubes based on clar sextet theory. *Organic Letters*, *9*, 4267–4270.
- Baldoni, M., Sgamellotti, A., & Mercuri, F. (2008). Electronic properties and stability of graphene nanoribbons: An interpretation based on Clar sextet theory. *Chemical Physics Letters*, *464*, 202–207.
- Barone, V., & Scuseria, G. E. (2004). Theoretical study of the electronic properties of narrow single-walled carbon nanotubes: Beyond the local density approximation. *The Journal of Chemical Physics*, *121*, 10376–10379.
- Barone, V., Peralta, J. E., & Scuseria, G. E. (2005a). Optical transitions in metallic single-walled carbon nanotubes. *Nano Letters*, *5*, 1830–1833.
- Barone, V., Peralta, J. E., Wert, M., Heyd, J., & Scuseria, G. E. (2005b). Density functional theory study of optical transitions in semiconducting single-walled carbon nanotubes. *Nano Letters*, *5*, 1621–1624.
- Barone, V., Hod, O., & Scuseria, G. E. (2006). Electronic structure and stability of semiconducting graphene nanoribbons. *Nano Letters*, *6*, 2748–2754.
- Batista, E. R., Heyd, J., Hennig, R. G., Uberuaga, B. P., Martin, R. L., Scuseria, G. E., Umrigar, C. J., & Wilkins, J. W. (2006). Comparison of screened hybrid density functional theory to diffusion Monte Carlo in calculations of total energies of silicon phases and defects. *Physical Review B*, *74*, 121102.
- Becke, A. D. (1993). Density-functional thermochemistry III. The role of exact exchange. *The Journal of Chemical Physics*, *98*, 5648–5652.
- Berger, C., Song, Z., Li, X., Wu, X., Brown, N., Naud, C., Mayou, D., Li, T., Hass, J., Marchenkov, A. N., Conrad, E. H., First, P. N., & de Heer, W. A. (2006). Electronic confinement and coherence in patterned epitaxial graphene. *Science*, *312*, 1191.

- Bhardwaj, T., Antic, A., Pavan, B., Barone, V., & Fahlman, B. D. (2010). Enhanced electrochemical lithium storage by graphene nanoribbons. *Journal of the American Chemical Society*, *132*, 12556–12559.
- Blase, X., Benedict, L. X., Shirley, E. L., & Louie, S. G. (1994). Hybridization effects and metallicity in small radius carbon nanotubes. *Physical Review Letters*, *72*, 1878–1881.
- Brothers, E. N., Izmaylov, A. F., Normand, J. O., Barone, V., & Scuseria, G. E. (2008). Accurate solid-state band gaps via screened hybrid electronic structure calculations. *The Journal of Chemical Physics*, *129*, 011102.
- Bunch, J. S., van der Zande, A. M., Verbridge, S. S., Frank, I. W., Tanenbaum, D. M., Parpia, J. M., Craighead, H. G., & McEuen, P. L. (2007). Electromechanical resonators from graphene sheets. *Science*, *315*, 490–493.
- Cabria, I., Mintmire, J. W., & White, C. T. (2003). Metallic and semiconducting narrow carbon nanotubes. *Physical Review B*, *67*, 121406.
- Cai, J., Ruffieux, P., Jaafar, R., Bieri, M., Braun, T., Blankenburg, S., Muoth, M., Seitsonen, A. P., Saleh, M., Feng, X., Müllen, K., & Fasel, R. (2010). Atomically precise bottom-up fabrication of graphene nanoribbons. *Nature*, *466*, 470–473.
- Cao, J., Wang, Q., & Dai, H. (2003). Electromechanical properties of metallic, quasimetallic, and semiconducting carbon nanotubes under stretching. *Physical Review Letters*, *90*, 157601.
- Cervantes-Sodi, F., Csányi, G., Piscanec, S., & Ferrari, A. C. (2008a). Edge-functionalized and substitutionally doped graphene nanoribbons: Electronic and spin properties. *Physical Review B*, *77*, 165427.
- Cervantes-Sodi, F., Csányi, G., Piscanec, S., & Ferrari, A. C. (2008b). Electronic properties of chemically modified graphene ribbons. *Physica Status Solidi B*, *245*, 2068–2071.
- Chen, Z., Kobashi, K., Rauwald, U., Booker, R., Fan, H., Hwang, W.-F., & Tour, J. M. (2006a). Soluble ultra-short single-walled carbon nanotubes. *Journal of the American Chemical Society*, *128*, 10568–10571.
- Chen, Z., Ziegler, K., Shaver, J., Hauge, R., & Smalley, R. (2006b). Cutting of single-walled carbon nanotubes by ozonolysis. *Journal of Physical Chemistry B*, *110*, 11624–11627.
- Chen, Z., Lin, Y.-M., Rooks, M. J., & Avouris, P. (2007). Graphene nano-ribbon electronics. *Physica E*, *40*, 228–232.
- Choi, S.-M., & Jhi, S.-H. (2008). Self-assembled metal atom chains on graphene nanoribbons. *Physical Review Letters*, *101*, 266105.
- Cohen-Karni, T., Segev, L., Srur-Lavi, O., Cohen, S. R., & Joselevich, E. (2006). Torsional electromechanical quantum oscillations in carbon nanotubes. *Nature Nanotechnology*, *1*, 36–41.
- Deslippe, J., Spataru, C. D., Prendergast, D., & Louie, S. G. (2007). Bound excitons in metallic single-walled carbon nanotubes. *Nano Letters*, *7*, 1626–1630.
- Ernzerhof, M., & Scuseria, G. E. (1999). Assessment of the Perdew-Burke-Ernzerhof exchange-correlation functional. *The Journal of Chemical Physics*, *110*, 5029–5036.
- Ezawa, M. (2006). Peculiar width dependence of the electronic properties of carbon nanoribbons. *Physical Review B*, *73*, 045432.
- Ezawa, M. (2007). Metallic graphene nanodisks: Electronic and magnetic properties. *Physical Review B*, *76*, 245415.
- Faccio, R., Denis, P. A., Pardo, H., Goyenola, C., & Mombro, A. W. (2009). Mechanical properties of graphene nanoribbons. *Journal of Physics: Condensed Matter*, *21*, 285304.
- Fantini, C., Jorio, A., Souza, M., Strano, M. S., Dresselhaus, M. S., & Pimenta, M. A. (2004). Optical transition energies for carbon nanotubes from resonant Raman spectroscopy: Environment and temperature effects. *Physical Review Letters*, *93*, 147406.
- Feller, D. (1996). The role of databases in support of computational chemistry calculations. *Journal of Computational Chemistry*, *17*, 1571–1586.
- Feller, D. (2007). Basis set exchange: v1.2.2. <https://bse.pnl.gov/bse/portal>
- Fennimore, A., Yuzvinsky, T., Han, W., Fuhrer, M., Cumings, J., & Zettl, A. (2003). Rotational actuators based on carbon nanotubes. *Nature*, *424*, 408–410.

- Fernandez-Rossier, J., & Palacios, J. J. (2007). Magnetism in graphene nanoislands. *Physical Review Letters*, *99*, 177204.
- Frank, I. W., Tanenbaum, D. M., Van der Zande, A. M., & McEuen, P. L. (2007). Mechanical properties of suspended graphene sheets. *Journal of Vacuum Science & Technology B*, *25*, 2558–2561.
- Frey, J. T., & Doren, D. J. (2005). TubeGen Online - Version 3.3 - Web-Accessible nanotube structure generator. <http://turin.nss.udel.edu/research/tubegenonline.html>.
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, N. J., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., & Fox, D. J. (2006). *GAUSSIAN development version, revision f.02*. Wallingford CT: Gaussian Inc.
- Fujita, M., Wakabayashi, K., Nakada, K., & Kusakabe, K. (1996). Peculiar localized state at zigzag graphite edge. *Journal of the Physical Society of Japan*, *65*, 1920–1923.
- Gomez-Navarro, C., de Pablo, P., & Gomez-Herrero, J. (2004). Radial electromechanical properties of carbon nanotubes. *Advanced Materials*, *16*, 549.
- Gruneich, A., & Hess, B. (1998). Choosing GTO basis sets for periodic HF calculations. *Theoretical Chemistry Accounts*, *100*, 253–263.
- Gu, Z., Peng, H., Hauge, R., Smalley, R., & Margrave, J. (2002). Cutting single-wall carbon nanotubes through fluorination. *Nano Letters*, *2*, 1009–1013.
- Gülseren, O., Yildirim, T., & Ciraci, S. (2002). Systematic ab initio study of curvature effects in carbon nanotubes. *Physical Review B*, *65*, 153405.
- Gunlycke, D., Li, J., Mintmire, J., & White, C. (2007). Altering low-bias transport in zigzag-edge graphene nanostrips with edge chemistry. *Applied Physics Letters*, *91*, 112108.
- Guo, G., Liu, L., Chu, K., Jayanthi, C., & Wu, S. (2005). Electromechanical responses of single-walled carbon nanotubes: Interplay between the strain-induced energy-gap opening and the pinning of the Fermi level. *Journal of Applied Physics Letters*, *98*, 044311.
- Hall, A. R., Falvo, M. R., Superfine, R., & Washburn, S. (2007). Electromechanical response of single-walled carbon nanotubes to torsional strain in a self-contained device. *Nature Nanotechnology*, *2*, 413–416.
- Hamada, N., Sawada, S., & Oshiyama, A. (1992). New one-dimensional conductors – Graphitic microtubules. *Physical Review Letters*, *68*, 1579.
- Han, M. Y., Oezylmaz, B., Zhang, Y., & Kim, P. (2007). Energy band-gap engineering of graphene nanoribbons. *Physical Review Letters*, *98*, 206805.
- Heyd, J., & Scuseria, G. E. (2004). Efficient hybrid density functional calculations in solids: Assessment of the Heyd-Scuseria-Ernzerhof screened Coulomb hybrid functional. *The Journal of Chemical Physics*, *121*, 1187–1192.
- Heyd, R., Charlier, A., & McRae, E. (1997). Uniaxial-stress effects on the electronic properties of carbon nanotubes. *Physical Review B*, *55*(11), 6820–6824.
- Heyd, J., Scuseria, G. E., & Ernzerhof, M. (2003). Hybrid functionals based on a screened Coulomb potential. *The Journal of Chemical Physics*, *118*, 8207–8215.
- Heyd, J., Peralta, J., Scuseria, G., & Martin, R. (2005). Energy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional. *The Journal of Chemical Physics*, *123*, 174101.
- Heyd, J., Scuseria, G. E., & Ernzerhof, M. (2006). Erratum: Hybrid functionals based on a screened Coulomb potential [*The Journal of Chemical Physics*, *118*, 8207 (2003)]. *The Journal of Chemical Physics*, *124*, 219906

- Higuchi, Y., Kusakabe, K., Suzuki, N., Tsuneyuki, S., Yamauchi, J., Akagi, K., & Yoshimoto, Y. (2004). Nanotube-based molecular magnets with spin-polarized edge states. *Journal of Physics. Condensed Matter*, *16*, S5689–S5692.
- Hod, O., & Scuseria, G. E. (2008). Half-metallic-zigzag carbon nanotube dots. *ACS Nano*, *2*, 2243–2249.
- Hod, O., & Scuseria, G. E. (2009). Electromechanical properties of suspended graphene nanoribbons. *Nano Letters*, *9*, 2619–2622.
- Hod, O., Peralta, J. E., & Scuseria, G. E. (2006). First-principles electronic transport calculations in finite elongated systems: A divide and conquer approach. *The Journal of Chemical Physics*, *125*, 114704.
- Hod, O., Barone, V., Peralta, J. E., & Scuseria, G. E. (2007a). Enhanced half-metallicity in edge-oxidized zigzag graphene nanoribbons. *Nano Letters*, *7*, 2295–2299.
- Hod, O., Peralta, J. E., & Scuseria, G. E. (2007b). Edge effects in finite elongated graphene nanoribbons. *Physical Review B*, *76*, 233401.
- Hod, O., Barone, V., & Scuseria, G. E. (2008). Half-metallic graphene nanodots: A comprehensive first-principles theoretical study. *Physical Review B*, *77*, 035411.
- Iijima, S. (1991). Helical microtubules of graphitic carbon. *Nature*, *354*, 56–58.
- Javey, A., Qi, P., Wang, Q., & Dai, H. (2004). Ten- to 50-nm-long quasi-ballistic carbon nanotube devices obtained without complex lithography. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 13408–13410.
- Jiang, H., Bu, W., Jiang, J., & Dong, J. (2004). Electronic structure in finite-length deformed metallic carbon nanotubes. *The European Physical Journal B*, *42*, 503–508.
- Jiang, D. E., Sumpter, B. G., & Dai, S. (2007). First principles study of magnetism in nanographenes. *The Journal of Chemical Physics*, *127*, 124703.
- Kan, E. J., Li, Z., Yang, J., & Hou, J. G. (2007). Will zigzag graphene nanoribbon turn to half metal under electric field? *Applied Physics Letters*, *91*, 243116.
- Kane, C. L., & Mele, E. J. (1997). Size, shape, and low energy electronic structure of carbon nanotubes. *Physical Review Letters*, *78*(10), 1932–1935.
- Kataura, H., Kumazawa, Y., Maniwa, Y., Umezū, I., Suzuki, S., Ohtsuka, Y., & Achiba, Y. (1999). Optical properties of single-wall carbon nanotubes. *Synthetic Metals*, *103*, 2555–2558.
- Khabashesku, V., Billups, W., & Margrave, J. (2002). Fluorination of single-wall carbon nanotubes and subsequent derivatization reactions. *Accounts of Chemical Research*, *35*, 1087–1095.
- Kim, Y., Choi, J., Chang, K., & Tománek, D. (2003). Defective fullerenes and nanotubes as molecular magnets: An ab initio study. *Physical Review B*, *68*, 125420.
- Klein, D. (1994). Graphitic polymer strips with edge states. *Chemical Physics Letters*, *217*, 261–265.
- Kleiner, A., & Eggert, S. (2001). Band gaps of primary metallic carbon nanotubes. *Physical Review B*, *63*(7), 073408.
- Kobayashi, K. (1993). Electronic-structure of a stepped graphite surface. *Physical Review B*, *48*, 1757–1760.
- Kobayashi, Y., Fukui, K., Enoki, T., Kusakabe, K., & Kaburagi, Y. (2005). Observation of zigzag and armchair edges of graphite using scanning tunneling microscopy and spectroscopy. *Physical Review B*, *71*, 193406.
- Kobayashi, Y., Fukui, K., Enoki, T., Kusakabe, K., & Kaburagi, Y. (2006). Edge state on hydrogen-terminated graphite edges investigated by scanning tunneling microscopy. *Physical Review B*, *73*, 125415.
- Krepel, D., & Hod, O. (2011). Lithium adsorption on armchair graphene nanoribbons. *Surface Science*, *605*, 1633–1642.
- Kroto, H., Heath, J., O'Brien, S., Curl, R., & Smalley, R. (1985). C-60 – Buckminsterfullerene. *Nature*, *318*(6042), 162–163.
- Kudin, K. N. (2008). Zigzag graphene nanoribbons with saturated edges. *ACS Nano*, *2*, 516–522.
- Kümmel, S., & Kronik, L. (2008). Orbital-dependent density functionals: Theory and applications. *Reviews of Modern Physics*, *80*, 3–60.
- Kusakabe, K., & Maruyama, M. (2003). Magnetic nanographite. *Physical Review B*, *67*, 092406.

- Lammert, P. E., Zhang, P., & Crespi, V. H. (2000). Gapping by squashing: Metal-insulator and insulator-metal transitions in collapsed carbon nanotubes. *Physical Review Letters*, *84*(11), 2453–2456.
- Lee, G., & Cho, K. (2009). Electronic structures of zigzag graphene nanoribbons with edge hydrogenation and oxidation. *Physical Review B*, *79*, 165440.
- Li, Z., Tang, Z., Liu, H., Wang, N., Chan, C., Saito, R., Okada, S., Li, G., Chen, J., Nagasawa, N., & Tsuda, S. (2001). Polarized absorption spectra of single-walled 4 angstrom carbon nanotubes aligned in channels of an AlPO 4–5 single crystal. *Physical Review Letters*, *87*, 127401.
- Li, J., Zhang, Y., & Zhang, M. (2002). The electronic structure and its theoretical simulation of carbon nanotube with finited length Part I: The frontier orbitals and its properties of short armchair nanotubes. *Chemical Physics Letters*, *364*, 328–337.
- Lieb, E. (1989). 2 theorems on the Hubbard-model. *Physical Review Letters*, *62*, 1201–1204.
- Liu, H. J., & Chan, C. T. (2002). Properties of 4 angstrom carbon nanotubes from first-principles calculations. *Physical Review B*, *66*, 115416.
- Liu, L., Jayanthi, C. S., Tang, M., Wu, S. Y., Tomblor, T. W., Zhou, C., Alexseyev, L., Kong, J., & Dai, H. (2000). Controllable reversibility of an sp^2 to sp^3 transition of a single wall nanotube under the manipulation of an afm tip: A nanoscale electromechanical switch? *Physical Review Letters*, *84*(21), 4950–4953.
- Liu, L., Jayanthi, C., Guo, H., & Wu, S. (2001). Broken symmetry, boundary conditions, and band-gap oscillations in finite single-wall carbon nanotubes. *Physical Review B*, *64*, 033414.
- Lu, J.-Q., Wu, J., Duan, W., Liu, F., Zhu, B.-F., & Gu, B.-L. (2003). Metal-to-semiconductor transition in squashed armchair carbon nanotubes. *Physical Review Letters*, *90*(15), 156601.
- Machón, M., Reich, S., Thomsen, C., Sánchez-Portal, D., & Ordejón, P. (2002). Ab initio calculations of the optical properties of 4-angstrom-diameter single-walled nanotubes. *Physical Review B*, *66*, 155410.
- Maiti, A. (2001). Application of carbon nanotubes as electromechanical sensors – Results from first-principles simulations. *Physica Status Solidi*, *226*, 87–93.
- Maiti, A. (2003). Carbon nanotubes – Bandgap engineering with strain. *Nature Materials*, *2*, 440–442.
- Maiti, A. (2009). *Integrated analytical systems, computational methods for sensor material selection*. New York: Springer.
- Maiti, A., Svizhenko, A., & Anantram, M. (2002). Electronic transport through carbon nanotubes: Effects of structural deformation and tube chirality. *Physical Review Letters*, *88*, 126805.
- Mananes, A., Duque, F., Ayuela, A., Lopez, M. J., & Alonso, J. A. (2008). Half-metallic finite zigzag single-walled carbon nanotubes from first principles. *Physical Review B*, *78*, 035432.
- Mehrez, H., Svizhenko, A., Anantram, M., Elstner, M., & Frauenheim, T. (2005). Analysis of band-gap formation in squashed armchair carbon nanotubes. *Physical Review B*, *71*, 155421.
- Mickelson, E., Huffman, C., Rinzler, A., Smalley, R., Hauge, R., & Margrave, J. (1998). Fluorination of single-wall carbon nanotubes. *Chemical Physics Letters*, *296*, 188–194.
- Milaninia, K. M., Baldo, M. A., Reina, A., & Kong, J. (2009). All graphene electromechanical switch fabricated by chemical vapor deposition. *Applied Physics Letters*, *95*, 183105.
- Minot, E., Yaish, Y., Sazonova, V., Park, J., Brink, M., & McEuen, P. (2003). Tuning carbon nanotube band gaps with strain. *Physical Review Letters*, *90*, 156401.
- Mintmire, J., Dunlap, B., & White, C. (1992). Are fullerene tubules metallic? *Physical Review Letters*, *68*, 631.
- Nagapriya, K. S., Berber, S., Cohen-Karni, T., Segev, L., Srur-Lavi, O., Tomanek, D., & Joselevich, E. (2008). Origin of torsion-induced conductance oscillations in carbon nanotubes. *Physical Review B*, *78*, 165417.
- Nakada, K., Fujita, M., Dresselhaus, G., & Dresselhaus, M. S. (1996). Edge state in graphene ribbons: Nanometer size effect and edge shape dependence. *Physical Review B*, *54*, 17954–17961.
- Nakamura, E., Tahara, K., Matsuo, Y., & Sawamura, M. (2003). Synthesis, structure, and aromaticity of a hoop-shaped cyclic benzenoid [10]cyclophenacene. *Journal of the American Chemical Society*, *125*, 2834–2835.

- Nardelli, M. B. (1999). Electronic transport in extended systems: Application to carbon nanotubes. *Physical Review B*, 60(11), 7828–7833.
- Nardelli, M. B., & Bernholc, J. (1999). Mechanical deformations and coherent transport in carbon nanotubes. *Physical Review B*, 60(24), R16338–R16341.
- Niimi, Y., Matsui, T., Kambara, H., Tagami, K., Tsukada, M., & Fukuyama, H. (2005). Scanning tunneling microscopy and spectroscopy studies of graphite edges. *Applied Surface Science*, 241, 43–48.
- Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Zhang, Y., Dubonos, S. V., Grigorieva, I. V., & Firsov, A. A. (2004). Electric field effect in atomically thin carbon films. *Science*, 306, 666–669.
- Okada, S., & Oshiyama, A. (2003). Nanometer-scale ferromagnet: Carbon nanotubes with finite length. *Journal of the Physical Society of Japan*, 72, 1510–1515.
- Onida, G., Reining, L., & Rubio, A. (2002). Electronic excitations: Density-functional versus many-body Green's-function approaches. *Reviews of Modern Physics*, 74, 601–659.
- Paier, J., Marsman, M., & Kresse, G. (2007). Why does the B3LYP hybrid functional fail for metals? *The Journal of Chemical Physics*, 127, 024103.
- Paulson, S., Falvo, M., Snider, N., Helsen, A., Hudson, T., Seeger, A., Taylor, R., Superfine, R., & Washburn, S. (1999). In situ resistance measurements of strained carbon nanotubes. *Applied Physics Letters*, 75, 2936–2938.
- Peng, S., & Cho, K. (2002). Nano electro mechanics of semiconducting carbon nanotube. *Journal of Applied Mechanics*, 69, 451.
- Peralta, J. E., Heyd, J., Scuseria, G. E., & Martin, R. L. (2006). Spin-orbit splittings and energy band gaps calculated with the Heyd-Scuseria-Ernzerhof screened hybrid functional. *Physical Review B*, 74, 073101.
- Perdew, J. P., Burke, K., & Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Physical Review Letters*, 77, 3865–3868.
- Perdew, J. P., Ernzerhof, M., & Burke, K. (1997). Rationale for mixing exact exchange with density functional approximations. *The Journal of Chemical Physics*, 105, 9982–9985.
- Pimenta, M., Gomes, A., Fantini, C., Cancado, L., Araujo, P., Maciel, I., Santos, A., Furtado, C., Peressinotto, V., Plentz, F., & Jorio, A. (2007). Optical studies of carbon nanotubes and nanographites. *Physica E*, 37, 88–92.
- Poot, M., & van der Zant, H. S. J. (2008). Nanomechanical properties of few-layer graphene membranes. *Applied Physics Letters*, 92, 063111.
- Prezzi, D., Varsano, D., Ruini, A., Marini, A., & Molinari, E. (2008). Optical properties of graphene nanoribbons: The role of many-body effects. *Physical Review B*, 77, 041477.
- Radovic, L. R., & Bockrath, B. (2005). On the chemical nature of graphene edges: Origin of stability and potential for magnetism in carbon materials. *Journal of the American Chemical Society*, 127, 5917–5927.
- Reich, S., Thomsen, C., & Ordejón, P. (2002). Systematic ab initio study of curvature effects in carbon nanotubes. *Physical Review B*, 65, 153405.
- Rigo, V. A., Martins, T. B., da Silva, A. J. R., Fazzio, A., & Miwa, R. H. (2009). Electronic, structural, and transport properties of Ni-doped graphene nanoribbons. *Physical Review B*, 79, 075435.
- Ritter, K. A., & Lyding, J. W. (2009). The influence of edge structure on the electronic properties of graphene quantum dots and nanoribbons. *Nature Materials*, 8, 235–242.
- Rocheffort, A., Salahub, D. R., & Avouris, P. (1998). The effect of structural distortions on the electronic structure of carbon nanotubes. *Chemical Physics Letters*, 297, 45–50.
- Rocheffort, A., Avouris, P., Lesage, F., & Salahub, D. R. (1999a). Electrical and mechanical properties of distorted carbon nanotubes. *Physical Review B*, 60(19), 13824–13830.
- Rocheffort, A., Salahub, D., & Avouris, P. (1999b). Effects of finite length on the electronic structure of carbon nanotubes. *Journal of Physical Chemistry B*, 103, 641–646.
- Rudberg, E., Salek, P., & Luo, Y. (2007). Nonlocal exchange interaction removes half-metallicity in graphene nanoribbons. *Nano Letters*, 7, 2211–2213.

- Rueckes, T., Kim, K., Joselevich, E., Tseng, G., Cheung, C., & Lieber, C. (2000). Carbon nanotube-based nonvolatile random access memory for molecular computing. *Science*, *289*, 94–97.
- Runge, E., & Gross, E. (1984). Density-functional theory for time-dependent systems. *Physical Review Letters*, *52*, 997–1000.
- Saito, R., Fujita, M., Dresselhaus, G., & Dresselhaus, M. (1992). Electronic structure of graphene tubules based on C 60. *Physical Review B*, *46*, 1804–1811.
- Saito, R., Dresselhaus, G., & Dresselhaus, M. S. (1998). *Physical properties of carbon nanotubes*. London: Imperial College Press.
- Sazonova, V., Yaish, Y., Ustunel, H., Roundy, D., Arias, T., & McEuen, P. (2004). A tunable carbon nanotube electromechanical oscillator. *Nature*, *431*, 284–287.
- Semet, V., Binh, V., Guillot, D., Teo, K., Chhowalla, M., Amaratunga, G., Milne, W., Legagneux, P., & Pribat, D. (2005). Reversible electromechanical characteristics of individual multiwall carbon nanotubes. *Applied Physics Letters*, *87*, 223103.
- Sevincli, H., Topsakal, M., Durgun, E., & Ciraci, S. (2008). Electronic and magnetic properties of 3d transition-metal atom adsorbed graphene and graphene nanoribbons. *Physical Review B*, *77*, 195434.
- Sfeir, M. Y., Beetz, T., Wang, F., Huang, L. M., Huang, X. M. H., Huang, M. Y., Hone, J., O'Brien, S., Misewich, J. A., Heinz, T. F., Wu, L. J., Zhu, Y. M., & Brus, L. E. (2006). Optical spectroscopy of individual single-walled carbon nanotubes of defined chiral structure. *Science*, *312*, 554–556.
- Shaver, J., Kono, J., Portugal, O., Krstic, V., Rikken, G. L. J. A., Miyauchi, Y., Maruyama, S., & Perebeinos, V. (2007). Magnetic brightening of carbon nanotube photoluminescence through symmetry breaking. *Nano Letters*, *7*, 1851–1855.
- Shemella, P., Zhang, Y., Mailman, M., Ajayan, P. M., & Nayak, S. K. (2007). Energy gaps in zero-dimensional graphene nanoribbons. *Applied Physics Letters*, *91*, 042101.
- Son, Y.-W., Cohen, M. L., & Louie, S. G. (2006a). Energy gaps in graphene nanoribbons. *Physical Review Letters*, *97*, 216803.
- Son, Y.-W., Cohen, M. L., & Louie, S. G. (2006b). Half-metallic graphene nanoribbons. *Nature*, *444*, 347–349.
- Spataru, C. D., Ismael-Beigi, S., Benedict, L. X., & Louie, S. G. (2004). Excitonic effects and optical spectra of single-walled carbon nanotubes. *Physical Review Letters*, *92*, 077402.
- Spataru, C. D., Ismael-Beigi, S., Capaz, R. B., & Louie, S. G. (2008). Quasiparticle and excitonic effects in the optical response of nanotubes and nanoribbons. *Topics in Applied Physics*, *111*, 195–227.
- Springborg, M., & Satpathy, S. (1994). Density-functional calculations of electronic and structural properties of small fullerene tubules. *Chemical Physics Letters*, *255*, 454–461.
- Stampfer, C., Jungen, A., Linderman, R., Oberfell, D., Roth, S., & Hierold, C. (2006). Nano-electromechanical displacement sensing based on single-walled carbon nanotubes. *Nano Letters*, *6*, 1449–1453.
- Stein, S., & Brown, R. (1987). Pi-electron properties of large condensed polyaromatic hydrocarbons. *Journal of the American Chemical Society*, *109*, 3721–3729.
- Sun, L., Li, Q., Ren, H., Su, H., Shi, Q. W., & Yang, J. (2008). Strain effect on electronic structures of graphene nanoribbons: A first-principles study. *Journal of Chemical Physics*, *129*, 074704.
- Svishchenko, A., Mehrez, H., Anantram, A. M. P., & Maiti, A. (2005). Sensing mechanical deformation in carbon nanotubes by electrical response: a computational study. *Proceedings of SPIE*, *5593*, 416–428.
- Tanaka, K., Yamashita, S., Yamabe, H., & Yamabe, T. (1987). Electronic-properties of one-dimensional graphite family. *Synthetic Metals*, *17*, 143–148.
- Tang, Z. K., Sun, H. D., Wang, J., Chen, J., & Li, G. (1998). Mono-sized single-wall carbon nanotubes formed in channels of AlPO 4–5 single crystal. *Applied Physics Letters*, *73*, 2287–2289.
- Tao, J., Perdew, J. P., Staroverov, V. N., & Scuseria, G. E. (2003). Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Physical Review Letters*, *91*, 146401.

- Telg, H., Maultzsch, J., Reich, S., Hennrich, F., & Thomsen, C. (2004). Chirality distribution and transition energies of carbon nanotubes. *Physical Review Letters*, *93*, 177401.
- Tombler, T., Zhou, C., Alexseyev, L., Kong, J., Dai, H., Lei, L., Jayanthi, C., Tang, M., & Wu, S. (2000). Reversible electromechanical characteristics of carbon nanotubes under local-probe manipulation. *Nature*, *405*, 769–772.
- Tsukada, M., Adachi, H., & Satoko, C. (1983). Theory of electronic-structure of oxide surfaces. *Progress in Surface Science*, *14*, 113–173.
- Uthaisar, C., & Barone, V. (2010). Edge effects on the characteristics of li diffusion in graphene. *Nano Letters*, *10*, 2838–2842.
- Uthaisar, C., Barone, V., & Peralta, J. E. (2009). Lithium adsorption on zigzag graphene nanoribbons. *Journal of Applied Physics*, *106*, 113715.
- Vanderbilt, D. (1990). Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Physical Review B*, *41*, 7892–7895.
- Vosko, S. H., Wilk, L., & Nusair, M. (1980). Accurate spin-dependent electron liquid correlation energies for local spin-density calculations – A critical analysis. *Canadian Journal of Physics*, *58*, 1200–1211.
- Wakabayashi, K., Fujita, M., Ajiki, H., & Sigrist, M. (1999). Electronic and magnetic properties of nanographite ribbons. *Physical Review B*, *59*, 8271–8282.
- Wang, F., Dukovic, G., Brus, L. E., & Heinz, T. F. (2005). The optical resonances in carbon nanotubes arise from excitons. *Science*, *308*, 838–841.
- Wang, X., Ouyang, Y., Li, X., Wang, H., Guo, J., & Dai, H. (2008). Room-temperature all-semiconducting sub-10-nm graphene nanoribbon field-effect transistors. *Physical Review Letters*, *100*, 206803.
- Wassmann, T., Seitsonen, A. P., Saitta, A. M., Lazzeri, M., & Mauri, F. (2008). Structure, stability, edge states, and aromaticity of graphene ribbons. *Physical Review Letters*, *101*, 096402.
- Weisman, R. B., & Bachilo, S. M. (2003). *Nano Letters*, *3*, 1235–1238.
- Wu, J., Zang, J., Larade, B., Guo, H., Gong, X., & Liu, F. (2004). Computational design of carbon nanotube electromechanical pressure sensors. *Physical Review B*, *69*, 153406.
- Yang, L., & Han, J. (2000). Electronic structure of deformed carbon nanotubes. *Physical Review Letters*, *85*(1), 154–157.
- Yang, L., Anantram, M. P., Han, J., & Lu, J. P. (1999). Band-gap change of carbon nanotubes: Effect of small uniaxial and torsional strain. *Physical Review B*, *60*(19), 13874–13878.
- Yang, L., Cohen, M. L., & Louie, S. G. (2007). Excitonic effects in the optical spectra of graphene nanoribbons. *Nano Letters*, *10*, 3112–3115.
- Yang, X., Dou, X., Rouhanipour, A., Zhi, L., Rader, H. J., & Müllen, K. (2008). Two-dimensional graphene nanoribbons. *Journal of the American Chemical Society*, *130*, 4216.
- Zhang, D.-B., James, R. D., & Dumitrica, T. (2009). Electromechanical characterization of carbon nanotubes in torsion via symmetry adapted tight-binding objective molecular dynamics. *Physical Review B*, *80*(11), 115418.
- Zhao, X., Liu, Y., Inoue, S., Suzuki, T., Jones, R. O., & Ando, Y. (2004). Smallest carbon nanotube Is 3 angstrom in diameter. *Physical Review Letters*, *92*, 125502.
- Zheng, H., & Duley, W. (2008). First-principles study of edge chemical modifications in graphene nanodots. *Physical Review B*, *78*, 045421.

Variation of the Surface to Bulk Contribution to Cluster Properties

30

Antonis N. Andriotis, Zacharias G. Fthenakis, and Madhu Menon

Contents

Introduction	1340
The Model	1342
Tight-Binding Molecular Dynamics Methodology	1342
Collinear Magnetic Effects	1344
Step 1: Inclusion of Randomness in the Direction of the Atomic Magnetic Moments ...	1344
Step 2: Inclusion of Spin-Orbit Interaction	1345
Step 3: Inclusion of Temperature Effects	1347
Computational Approach	1348
Results and Discussion	1349
Conclusion	1354
Bibliography	1354

Abstract

Recent computer simulations have indicated that there is a linear relationship between the melting and the Curie temperatures for Ni_n ($n \leq 201$) clusters. In this chapter, it is argued that this result is a consequence of the fact that the surface and the core (bulk) contributions to the cluster properties vary with the cluster size in an analogous way. The universal aspect of this result is also discussed. Among the many interesting consequences resulting from this relationship is the intriguing possibility of the coexistence of melting and magnetization. As demonstrated, these conclusions have as their origin the major contribution coming from the melting/magnetization ratio arising from surface effects and

A.N. Andriotis (✉) • Z.G. Fthenakis • M. Menon
Department of Physics and Astronomy and Center for Computational Sciences, University of
Kentucky, Lexington, KY, USA

Institute of Electronic Structure and Laser, FORTH, Heraklio, Crete, Greece
e-mail: andriot@iesl.forth.gr; fthenak@iesl.forth.gr; madhu@ccs.uky.edu

appear to overshadow all other contributions. As a result, this can be quantified with approximate methods which are suitable for describing any major surface contribution to a cluster property.

Introduction

As the cluster size increases, the cluster properties evolve toward their bulk counterparts. The understanding of this evolution is of fundamental importance not only from the perspective of basic science but also from the technological viewpoint. At a very approximate level, one can claim that the cluster properties can be described in terms of their surface and core (bulk) contributions and due to the fact that the surface to bulk ratio gets smaller as the cluster size increases. Consequently, it is natural to expect that the cluster properties will evolve to their corresponding bulk-phase ones for large cluster sizes.

As one particular example, we mention the melting temperature of large clusters. According to the proposed model, the following functional relationship for the variation of the melting temperature, $T_{\text{melt},N}^{\text{cl}}$, of a cluster as the number N of its atoms increases has been suggested:

$$T_{\text{melt},N}^{\text{cl}} = T_{\text{melt}}^{\text{bulk}} - \delta_{\text{melt}} N^{-1/3}, \quad (1)$$

where $T_{\text{melt}}^{\text{bulk}}$ is the melting temperature of the corresponding bulk phase and δ_{melt} is a constant that depends on N (Garcia-Rodeja et al. 1994; Gunes and Erkoc 2000; Lee et al. 2001; Nayak et al. 1998; Qi et al. 2001; Rey et al. 1993; Sun and Gong 1998). Correction terms in $N^{-2/3}$ and N^{-1} powers to the above expression have also been suggested (Doye and Calvo 2001). The term proportional to $N^{-1/3}$ in Eq. 1 reflects the surface to bulk contribution to the melting temperature.

Equation 1 was found to describe reasonably well the experimental findings in the large-size regime ($N > 500$). However, for clusters of smaller size, (especially for clusters with number of atoms $N \leq 50$), Eq. 1 does not ensure a quantitative description of the variation of $T_{\text{melt},N}^{\text{cl}}$ with the cluster size (see, e.g., Baletto and Ferrando 2005; Lee et al. 2001; Nayak et al. 1998; Qi et al. 2001). This is because for small clusters, (1) the surface-to-volume contribution to $T_{\text{melt},N}^{\text{cl}}$ is very large and (2) is very sensitive to the variations of the surface structure and the cluster geometry (symmetry), as both of these characteristics get altered as the cluster size changes.

Another property that has attracted much interest recently is the one pertaining to the evolution of the magnetic features of the magnetic clusters as the cluster size and temperature increase. This is mainly because of the potential applications of the magnetic grains in fabricating new materials for advanced magnetic storage devices and other applications (Bansman et al. 2005).

Recently, the results of our computer simulations led us to the conclusion that the Curie, $T_{C,N}^{\text{cl}}$, and the melting, $T_{\text{melt},N}^{\text{cl}}$, temperature of Ni_N clusters consisting of N atoms are linearly related over a large range of cluster sizes ($N \leq 201$), and this relationship is quantified by the following equation (Andriotis et al. 2007) :

$$T_{\text{melt},N}^{cl} = \alpha T_{C,N}^{cl} + \beta, \quad (2)$$

where α and β are constants. Least square fitting of our results leads to $\alpha = 0.5414$ and $\beta = 443.82^\circ\text{K}$. This relationship was suggested to be related to the ratio of the surface to bulk contributions to the cluster melting as well as to the average magnetic moment per cluster atom, $\bar{\mu}_{i,N}(T)$ $i = 1, \dots, N$. It was then claimed that this relationship could be justified within the mean field theory applied separately to the surface and core regions of the cluster (Andriotis et al. 2007).

A cursory thought seems to suggest that any type of direct relationship between $T_{\text{melt},N}^{cl}$ and $T_{C,N}^{cl}$ to be fortuitous since melting and magnetic order seem to reflect completely different aspects of the crystal potential. Therefore, Eq. 2 cannot be considered as one which reestablishes a valid physical relationship between these two cluster properties.

Furthermore, the variation of both $T_{\text{melt},N}^{cl}$ and $T_{C,N}^{cl}$ with the cluster size constitute separate distinct and complicated projects for both theory and experiment. This is because significant contributions to both of these physical quantities have their origin in surface as well as cluster-core (bulk) effects which, at first look, affect $T_{\text{melt},N}^{cl}$ and $T_{C,N}^{cl}$ differently. These contributions include: the effects of the cohesive energy, the shape (symmetry), the size, the surface to bulk ratio, the surface tension, the temperature of the cluster, etc.

In early reports, it was found that $T_{\text{melt},N}^{cl}$ is usually smaller than the corresponding bulk value. Of interest is the result applied for large enough spherical clusters of radius R :

$$T_{\text{melt},N}^{cl}/T_{\text{melt}}^{\text{bulk}} = 1 - k_m/R \quad (3)$$

where k_m is a material-dependent constant (see, e.g., Buffat and Borel 1976 and references therein). A similar expression describing the variation of the Curie temperature with the cluster size was also found within the phenomenological Landau–Ginsburg–Devonshire theory (Huang et al. 2000), i.e.,

$$T_{C,N}^{cl}/T_C^{\text{bulk}} = 1 - k_c/R \quad (4)$$

Use of Eqs. 3 and 4 gives the following values for α and β in Eq. 2:

$$\alpha = \frac{k_c}{k_m} \frac{T_C^{\text{bulk}}}{T_{\text{melt}}^{\text{bulk}}} \quad (5)$$

and

$$\beta = T_C^{\text{bulk}} - \alpha T_{\text{melt}}^{\text{bulk}} = \left(1 - \frac{k_c}{k_m}\right) T_C^{\text{bulk}}. \quad (6)$$

Diep and collaborators (1989) using Monte Carlo simulations studied the effect of the magnetic interactions on the melting temperature of a cluster. Although their

investigation was limited to very small clusters M_N , ($N \in [7, 17]$) of transition metals M, however, the conclusions they arrived at are very important. In particular, among others, they have found that the incorporation of the magnetic interactions leaves the cluster structure unchanged. However, magnetic interactions lead to sharper peaks in the specific heat (and, therefore, to more precise determination of melting and Curie temperatures), a slight increase in the melting temperature, and a slight cluster-volume contraction (magnetostriction). After examining more carefully the results of Diep et al. (1989) (included in their Fig. 12), we find that the relation of $T_{\text{melt},N}^{cl}$ and $T_{C,N}^{cl}$ is more or less linear with the exception of the data of the very small clusters, i.e., for $N \in [7, 8, 9]$. The linear relationship between $T_{\text{melt},N}^{cl}$ and $T_{C,N}^{cl}$ was also suggested recently on the basis of semiempirical and approximate phenomenological model descriptions (see Yang and Jiang 2005 and references therein).

All these descriptions support our findings (i.e., Eq. 2), which are based on a firm quantum mechanical model procedure as outlined in the following. In this work, we investigate the implications of such a relationship which seems to specify a universal aspect of the surface contribution to the cluster properties.

The Model

The investigation of clusters of medium and intermediate sizes consisting of transition metal atoms poses a severe challenge in terms of computer capacities and computer time. For this reason, approximate schemes have been employed with most pronounced being those based on empirical classical potentials. However, these models have limited applicability when there is a need to understand more about the electronic structure of these systems and follow its changes as the cluster size increases and approaches the bulk phase. It is thus necessary to use methods with firm ab initio footing while at the same time not sacrificing computational efficiency. One such method is based on the Tight-Binding approximation which we have adopted in our work.

In this section, we discuss briefly the implementation of this approach in order to model the temperature and magnetic features of transition metal clusters. We will firstly give a brief overview of our TB computational scheme that we used for systems at zero temperature. In the following sections, we will describe the generalization of our method enabling the inclusion of magnetic and temperature effects.

Tight-Binding Molecular Dynamics Methodology

The details of our Generalized Tight-Binding Molecular Dynamics (GTBMD) scheme can be found in Andriotis and Menon (1998) and Menon and Subbaswamy (1997). The GTBMD method makes explicit use of the nonorthogonality of the orbitals resulting in a transferable scheme that works well in the range all the way from a few atoms to the condensed solid. The scheme also includes d -electron

interactions enabling dynamic treatment of magnetic effects in transition metal systems. Here, we give a brief overview.

The total energy U is written in its general form as a sum of several terms (Andriotis and Menon 1998),

$$U = U_{\text{el}} + U_{\text{rep}} + U_0, \quad (7)$$

where U_{el} is the sum of the one-electron energies E_n for the occupied states:

$$U_{\text{el}} = \sum_n^{\text{occ}} E_n. \quad (8)$$

In the tight-binding scheme, E_n is obtained by solving the characteristic equation:

$$(\mathbf{H} - E_n \mathbf{S}) \mathbf{C}^n = 0, \quad (9)$$

where \mathbf{H} is the Hamiltonian of the system and \mathbf{S} the overlap matrix.

The Hellmann–Feynman theorem for obtaining the electronic part of the force is given by Menon and Subbaswamy (1997),

$$\frac{\partial E_n}{\partial x} = \frac{\mathbf{C}^{n\dagger} \left(\frac{\partial \mathbf{H}}{\partial x} - E_n \frac{\partial \mathbf{S}}{\partial x} \right) \mathbf{C}^n}{\mathbf{C}^{n\dagger} \mathbf{S} \mathbf{C}^n}. \quad (10)$$

The total energy expression also derives contributions from ion–ion repulsion interactions. This is approximated by a sum of pairwise repulsive terms and included in U_{rep} . This sum also contains the corrections arising from the double counting of electron–electron interactions in U_{el} (Andriotis and Menon 1998). U_0 is a constant that merely shifts the zero of energy. The contribution to the total force from U_{rep} is rather straightforward. One can then easily do molecular dynamics simulations by numerically solving Newton’s equation,

$$m \frac{d^2 x}{dt^2} = F_x = -\frac{\partial U}{\partial x} \quad (11)$$

to obtain x as a function of time.

Our TBMD scheme for a binary system consisting of elements A and B is based on a minimal set of five adjustable parameters for each pair (A,A), (B,B), and (A,B). These parameters are determined by fitting to experimental data for quantities such as the bond length, the vibrational frequency, and the binding energy of the dimers A_2 , B_2 , AB; the cohesive energy of the corresponding bulk states of the A, B, AB materials; and the energy level spacing of the lowest magnetic states of the dimer and trimer binary clusters consisting of atoms of A and B type. In the absence of experimental data, we fit to data of small clusters obtained using ab initio methods as described in the following subsection. It is apparent that only five parameters are required in the case of a single species system. The generalization to a system containing more than two kind of atoms is also plausible within this approach.

The fixed set of TB parameters are obtained from the universal scheme proposed by Harrison (1980) suitably scaled with respect to the interatomic distance (Andriotis and Menon 1998).

Collinear Magnetic Effects

In order to calculate the Curie temperature of a magnetic cluster, it is necessary to include non-collinear magnetic effects in our model description. These are introduced by extending our zero temperature (ZT) Tight-Binding Molecular Dynamics (TBMD) approach at the Hubbard-U level of approximation (Andriotis and Menon 1998) which we used to study magnetic clusters in the collinear magnetic approximation. According to this collinear model, an exchange-splitting parameter $s_0^{(i)}$ is introduced which is proportional to the intra-site Coulomb interaction U . This specifies the energy splitting between spin-up and spin-down electrons in the i th-atom in accordance with results obtained by ab initio methods. Thus, within this model, a site-diagonal spin-dependent Hamiltonian term $\mathbf{V}_{\text{spin}}^{(i)}$ is introduced which has the form:

$$\mathbf{V}_{\text{spin}}^{(i)} = \begin{pmatrix} s_0^{(i)} & 0 \\ 0 & -s_0^{(i)} \end{pmatrix} \quad (12)$$

In this model, it is assumed that all atomic magnetic moments (MMs) (of the cluster atoms) are collinear to the z-axis of a local xyz-system assigned to the i th cluster-atom.

The generalization of this model to include non-collinear effects is achieved in three steps. In the first step, we include the randomness in the directions of the atomic MMs. In the second, we include the spin-orbit interaction, and in the third, we include the temperature effects. For the sake of completeness, we briefly discuss this generalization in the following.

Step 1: Inclusion of Randomness in the Direction of the Atomic Magnetic Moments

In the first step, it is assumed that the deviation of the direction of the MM, μ_i , of the i th cluster-atom from the Z -axis of the global coordinate system XYZ is specified by the polar angles (θ_i, ϕ_i) defined with respect to this XYZ system. As a result, the potential $\mathbf{V}_{\text{spin}}^{(i)}$, originally defined in the local coordinate system xyz of the i th atom, is transformed to its expression $\mathbf{V}_{\text{spin}}^{(i),\text{global}}$ in the global system XYZ as follows (Anderson and Hasegawa 1955; Uhl et al. 1994):

$$\mathbf{V}_{\text{spin}}^{(i),\text{global}} = \Xi^\dagger(\theta_i, \phi_i) \mathbf{V}_{\text{spin}}^{(i)} \Xi(\theta_i, \phi_i), \quad (13)$$

where $\Xi(\theta_i, \phi_i)$ is the standard spin-1/2-rotation matrix :

$$\Xi(\theta_i, \phi_i) = \begin{pmatrix} e^{i\phi_i/2} \cos \theta_i/2 & e^{-i\phi_i/2} \sin \theta_i/2 \\ -e^{i\phi_i/2} \sin \theta_i/2 & e^{-i\phi_i/2} \cos \theta_i/2 \end{pmatrix} \quad (14)$$

It is assumed that the Z-axis of the global system can be arbitrarily chosen, and a usual choice is to take Z in alignment with the easy axis of the system.

Step 2: Inclusion of Spin-Orbit Interaction

In the second step, we introduce the Spin-Orbit (SO) interaction, $V_{SO}^{(i)}$, in the i th-atom within the \mathbf{L} - \mathbf{S} coupling scheme, i.e., $V_{SO}^{(i)} = -\lambda^{(i)} \mathbf{L}^{(i)} \cdot \mathbf{S}^{(i)}$ where, $\lambda^{(i)}$ is the spin-orbit coupling constant for the i th-atom, $\mathbf{L}^{(i)}$ its orbital angular momentum along the Z-axis, and $\mathbf{S}^{(i)}$ its total spin in the direction of μ_i .

Details of the implementation of the Spin-Orbit interaction within our TBMD method have been reported elsewhere (Andriotis and Menon 2004).

In the presence of a magnetic field, \mathbf{B} (assumed to be along the direction specified by the polar angles (θ_0, ϕ_0) with respect to XYZ-system), the atomic MMs of the cluster-atoms tend to become parallel to the direction of \mathbf{B} . The average projection of the MMs of the cluster-atoms, μ_{cl} , along the direction of \mathbf{B} (which is the experimentally measured quantity) is,

$$\mu_{cl} = \frac{1}{N_{cl}} \left| \sum_i^{N_{cl}} \mu_i \cos \gamma_i \right|, \quad (15)$$

where N_{cl} is the number of cluster-atoms and $\cos \gamma_i = \cos \theta_0 \cos \theta_i + \sin \theta_0 \sin \theta_i \cos(\phi_0 - \phi_i)$.

In a different formulation within the Hubbard-U model approximation to the e - e correlations, the spin-mixing interaction may be derived from a Coulomb-type Hamiltonian term of the form (Kato and Kokubo 1994; Ojeda et al. 1999): $\mathbf{V}_{\text{smix}} = -U \sum_{l\sigma} \rho_{l\sigma, l\bar{\sigma}} c_{l\bar{\sigma}}^\dagger c_{l\sigma}$, where $c_{l\sigma}^\dagger (c_{l\sigma})$ is the creation (annihilation) operator for an electron with spin σ at site l and $\rho_{l\bar{\sigma}, l\sigma}$ denote the electron density matrix elements, i.e., $\rho_{l\sigma, l\bar{\sigma}} = \langle c_{l\bar{\sigma}}^\dagger c_{l\sigma} \rangle$. It can be easily verified that the Hamiltonian term \mathbf{V}_{smix} is equivalent to that given by Eq. 13, i.e., $\mathbf{V}_{\text{spin}}^{(i), \text{global}}$.

Evaluation of the TB representation of the SO-interaction

In order to proceed with the evaluation of the TB matrix elements of the SO-interaction, we write the SO-term as

$$\mathbf{S} \cdot \mathbf{L} = L_x S_x + L_y S_y + L_z S_z \quad (16)$$

and compute the matrix elements with respect to the basis set. We are using a basis set of atomic orbitals $y_{ilm\sigma}(\mathbf{r})$, where the index i specifies the atom on which the atomic orbital (AO) is centered, l specifies the angular momentum, m is used to

count the various d -orbitals (i.e., d_{xy} , d_{xz} , d_{yz} , $d_{x^2-y^2}$, d_{z^2}), and σ denotes the spin. In terms of a linear superposition of these basis functions, the single electron wave functions take the form

$$\Psi_{i\sigma}(\mathbf{r}) = \sum_{l,m} C_{ilm\sigma} y_{ilm\sigma}(\mathbf{r}), \quad (17)$$

where $C_{ilm\sigma}$ denote the coefficients which are to be determined from the diagonalization of the Hamiltonian.

The spin operators of Eq. 16 refer to the global system and care has to be exercised as they act on the local spin states, the latter related with the former according to Eqs. 13 and 14. For example,

$$\langle \uparrow'_i | S_x | \downarrow'_i \rangle = \frac{1}{2} \cos \phi_i \sin \theta_i, \quad (18)$$

where the prime indicates the local functions and \uparrow , \downarrow indicate spin-up and spin-down states, respectively.

The matrix elements of the orbital angular momentum operators are obtained by operating on the angular part of the wave functions which, in our TBMD formalism are described by the Cubic harmonics. As an example, we write down the expression of the average value of the z -component of the orbital magnetic moment, L_z , in terms of its matrix elements:

$$\langle L_z \rangle = \sum_{i\sigma} \int d\mathbf{r} \Psi_{i\sigma}^*(\mathbf{r}) L_z \Psi_{i\sigma}(\mathbf{r}) \quad (19)$$

or

$$\langle L_z \rangle = \sum_{\sigma} \sum_i^{\text{occ}} \sum_{l,m} \sum_{l',m'} C_{ilm\sigma}^* C_{il'm'\sigma} \int d\mathbf{r} y_{ilm\sigma}^*(\mathbf{r}) L_z y_{il'm'\sigma}(\mathbf{r}) \quad (20)$$

Assuming orthogonality of AOs centered at different atoms, and orthonormal basis functions centered at one particular atom, we finally obtain the following expression for the d -orbital contribution to the orbital magnetic moment $\langle L_z^d \rangle$:

$$\langle L_z^d \rangle = \sum_{\sigma} \sum_i^{\text{occ}} \sum_{l=2} \sum_{mm'} \Delta_{mm'}^{il} C_{ilm\sigma}^* C_{il'm'\sigma} \quad (21)$$

where $\Delta_{mm'}^{il}$ are constants easily calculated by applying the relation:

$$L_z Y_{lm} = m Y_{lm} \quad (22)$$

where Y_{lm} are the spherical harmonics. For example, a straightforward calculation of the constants $\Delta_{mm'}^{il}$ shows that the only nonzero matrix elements of L_z are the following:

$$\langle d_{x^2-y^2} | L_z | d_{xy} \rangle = -2i \quad (23)$$

and

$$\langle yz | L_z | zx \rangle = -i \quad (24)$$

Combining the above, we construct the spin-dependent TB-representation of the SO-interaction term and add this to the other Hamiltonian terms.

Step 3: Inclusion of Temperature Effects

In the third step, our ZT-TBMD method has been extended by incorporating the *Nose*-bath (Nose 1984) and the Multiple Histogram approximations (Fanourgakis et al. 1997), so as to be applicable to cluster studies at finite temperatures in an efficient way (Andriotis et al. 2006, 2007; Fthenakis et al. 2003). This generalization allows one to calculate the caloric curve for the cluster and use this to study the effect of temperature on the structural, electronic, and magnetic properties of transition metal clusters and binary systems containing transition metal and semiconductor atoms. The method has been used to study the variation of structural and magnetic properties with temperature as well as to obtain the caloric curves of the Ni-clusters (Andriotis et al. 2006, 2007; Fthenakis et al. 2003).

Upon thermalization at temperature T , a cluster can be described by the canonical probability distribution function of total energy, $P_T(E)$, which specifies the probability that the system will be found in the energy interval $[E, E + \Delta E]$ at the specified temperature T . The distribution function corresponding to this temperature, within the canonical ensemble description, is (see Fanourgakis et al. 1997; Schmidt et al. 1997 and references therein):

$$P_T(E) = \frac{n_T(E)}{N_T} = \frac{[\Delta\Gamma(E)] e^{-E/k_B T}}{Z_T} \quad (25)$$

where $n_T(E)$ is the number of states in the energy interval $[E, E + \Delta E]$, N_T is the total number of accessible states, k_B is Boltzmann's constant, $\Delta\Gamma(E)$ the number of all the different states with energy in the interval $[E, E + \Delta E]$ (i.e., given by $Z_T = \sum_{E_i} \exp(-\beta E_i) = \sum_i \Delta\Gamma(E_i) \exp(-\beta E_i)$) and Z_T the partition function at temperature T .

A molecular dynamics (MD) simulation at a given temperature T provides numerical values for $n_T(E)$ at every accessible energy E . Having obtained these, we make use of the proposed Multiple Histogram Method (MHM) (Weerasinghe and Amar 1993), and obtain the partition functions Z_{T_j} for a finite set of temperatures T_j , $j = 1, \dots, M$ ($M \approx 200$) and the entropy terms $S(E_i) = k_B \ln[\Delta\Gamma(E_i)]$ (within an additive constant) for a much larger set of energy values E_i , $i = 1, \dots, N$ ($N \approx 6,000$) (Weerasinghe and Amar 1993).

Having obtained the quantities $n_T(E)$, $Z_T(E)$ and $S_T(E)$, we can then describe all the thermodynamic properties of the clusters and, in particular, the variation with temperature of their structural, electronic, and magnetic properties.

Computational Approach

The computation of the magnetic features of the clusters is performed within the above described non-collinear TBMD scheme as this allows for a full quantum MD relaxation of systems containing several hundred transition metal atoms while incorporating magnetic effects dynamically. More specifically, it includes: (1) e - e correlation effects at the Hubbard-U approximation (Andriotis and Menon 1998), (2) the spin-orbit interaction (in the $\mathbf{L} \cdot \mathbf{S}$ approximation), and (3) non-collinear magnetic effects (Andriotis and Menon 1998, 2004). Furthermore, the effect of temperature (Andriotis et al. 2006; Fthenakis et al. 2003) is included in the formalism while making use of the full s , p , d basis set and contains many unique features that make it ideally suited for the treatment of transition metal (TM) and semiconducting materials. For large-scale simulations we have developed a parallel algorithm that enables molecular dynamics simulations of systems containing atoms in excess of a thousand. This method is more accurate than the order- N methods that are being used at present to treat systems of these sizes. The successful application of our collinear TB scheme (see, e.g., Andriotis et al. 1998, 1999, 2000; Lathiotakis et al. 1996) guarantees similar success for the present generalization as well. Finally, we also mention that this approach has been suitably adapted in order for exclusive use in studying the transport properties (based on our computational codes as described in Andriotis and Menon 2001; Andriotis et al. 2002).

While the TBMD computational approach generalized in such a way is suitable for calculating the magnetic properties of the clusters (of a specific geometry), its use for relaxing the structure of particularly large clusters at nonzero temperatures has been found inefficient due to the extreme computational complexity. This is because the thermodynamic equilibration of a crystal requires sufficiently long MD relaxation time (of the order of 2–6 ns which is translated into 10–40 million MD steps with each step being of the order of a femtosecond).

In order to make our computations feasible, we firstly reach the thermodynamic equilibrium at each temperature using the classical Sutton–Chen interatomic potential (Sutton and Chen 1990) appropriately fitted to the TBMD results (Fthenakis et al. (unpublished)). It should be noted that for Ni, the classical Sutton–Chen potential (Sutton and Chen 1990) was found to give results closer in agreement with our TBMD method than any other classical potentials in use for Ni (Erkos 2001; Fthenakis et al. (unpublished)). While reaching the thermodynamic equilibrium, we apply our generalized TB formalism every 100 time-steps in order to calculate the MM of the cluster. In the MM calculation, the structure of the cluster is assumed frozen (as obtained within the classical potential approach at that particular time step), and the calculation of the MM of the cluster is repeated for a large number, $N_{\text{ran}}^{(i)}$, of atomic *spin configurations* taken randomly over the *ith-structural configuration* ($N_{\text{ran}}^{(i)} \approx 120 - 300$ for every i -time step in the present calculations).

In view of these results, Eq. 15 is generalized as follows:

$$\mu_{cl} = \frac{1}{KN_{cl}} \left| \sum_i^{N_{cl}} \sum_j^{N_{\text{ran}}^{(i)}} \mu_i^j \cos \gamma_i^j e^{-(E_i^j - E_0)/k_B T} \right|, \quad (26)$$

under the assumption that each spin configuration contributes to the magnetic state of the i th geometric configuration with probability $P_M(E_i^j) = \frac{1}{K} e^{-(E_i^j - E_0)/k_B T}$ where,

$$K = \sum_j N_{ran}^{(i)} e^{-(E_i^j - E_0)/k_B T}. \quad (27)$$

The index j in Eqs. 26 and 27 denotes quantities evaluated at the particular j th random atomic spin configuration of the i th structural configuration of the cluster. E_0 is taken to be the energy of the ferromagnetically aligned atomic-spin configuration. In this way, we take the average over the low-lying spin configurations of the cluster of a particular (frozen) geometric structure (i.e., as calculated at the specific time step). Finally, the thermodynamic average of μ_{cl} given by Eq. 26 over the various cluster geometric structures (i.e., time-steps) is obtained with the help of the probability $P_T(E)$ as given by Eq. 25.

Having obtained the temperature dependence of the average magnetic moment $\mu_{cl}(T)$ per cluster atom, we proceed with the calculation of the temperature dependence of the specific heat of each cluster by taking the derivative $d\{\mu_{cl}(T)\}^2/dT$ of the corresponding $\mu_{cl}(T)$ curves. Following Gerion et al. (2000), we obtain the Curie temperature $T_{C,N}^{cl}$ for each cluster by locating the maximum of the magnetic contribution to C_V . This is repeated for a number of clusters of various sizes.

Results and Discussion

In the present work, our focus is on the properties of the magnetic transition metal clusters and, in particular, on Ni clusters for which experimental data are available for comparison. Using the procedures discussed above, we have calculated the melting, $T_{melt,N}^{cl}$, and the Curie, $T_{C,N}^{cl}$, temperatures of the Ni_N clusters for $N \leq 201$. These results were discussed recently in Andriotis et al. (2007). The melting temperature $T_{melt,N}^{cl}$ has been derived on the basis of the Lindemann index using the caloric curve as obtained from the classical potential MD simulations. The Curie temperature $T_{C,N}^{cl}$ has been derived according to the quantum mechanical procedure discussed above. Both were found to increase with the cluster size tending to their corresponding bulk values as the size of the clusters increases, in good agreement with the existing experimental data (Andriotis et al. 2007).

The correlation between $T_{melt,N}^{cl}$ and $T_{C,N}^{cl}$ obtained from our results is shown in Fig. 1 by the black solid line. This demonstrates and supports the validity of Eq. 2. We discuss the universal aspect of this correlation and the conclusions that can be derived from in the following.

Firstly, it should be noted that the conclusions one can arrive at from the obtained relationship between melting and Curie temperatures of magnetic clusters depend crucially on the accuracy with which cluster melting temperatures are determined. This is a major issue as it has been extensively discussed in the literature (see, e.g., Baletto and Ferrando 2005; Qi et al. 2001; Sun and Gong 1998 and references

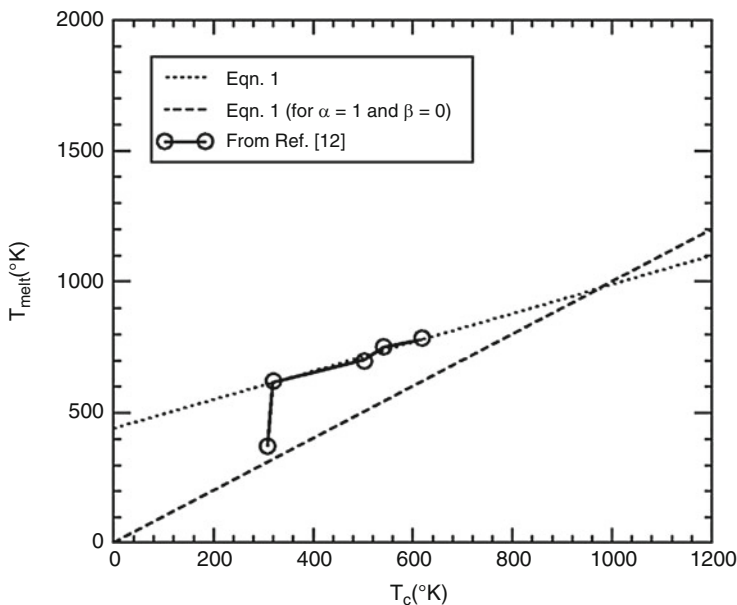


Fig. 1 Plot of $T_{\text{melt},N}^{\text{cl}}$ as a function of $T_{C,N}^{\text{cl}}$ for the Ni_N clusters studied in the present work, $N = 43, 80, 147, 177, 201$ (solid line). The dotted black line denoted Eq. 2 while and dashed black line describes Eq. 2 setting $\alpha = 1$ and $\beta = 0$. Points above (below) the dashed line have $T_{\text{melt}} > T_C$ ($T_{\text{melt}} < T_C$) respectively

therein). This is because surface melting of small particles occurs in a continuous manner over a broad temperature range in contradistinction to the melting of the solid-core (bulk-like) which occurs at a specific critical temperature (Garrigos et al. 1989). For this reason, it has been proposed that in order for $T_{\text{melt},N}^{\text{cl}}$ to be described quantitatively, the surface effects are usually treated separately from the core effects by introducing the *surface thickness*, t_0 , as a free parameter and calculate $T_{\text{melt},N}^{\text{cl}}$ by expressing firstly the heat of the cluster fusion in terms of t_0 and the cluster radius (see, e.g., Lai et al. 1996). A more commonly used approach for determining $T_{\text{melt},N}^{\text{cl}}$ is by employing the Lindemann's criterion, an approach followed in the present work (Fthenakis et al. 2003). Such a calculation is subject to the limitations and the accuracy of this method. In particular, the so derived melting temperatures depend strongly on the choice of the classical potential used (Andriotis et al. 2007; Fthenakis et al. 2003). This explains why the reported values for the melting temperature of Ni clusters cover a wide range. Nevertheless, the Sutton–Chen classical potential, employed in the present work for Ni, leads to accurate values for surface energies, vacancy energy, stacking fault energies, and bulk melting temperature in very good agreement with experiment (Qi et al. 2001).

The accuracy of Lindemann's criterion depends also on the steepness of its variation with temperature at phase transition and on the specification of its percentage increase which should be adequate to discriminate surface (partial) from

all-cluster melting. In the present case, we assign $T_{\text{melt},N}^{cl}$ to the temperature at which Lindemann's index starts increasing. This allows us to obtain the onset of cluster melting and to derive the melting temperature corresponding to a cluster phase in which unmolten parts with a possible magnetic order are still present.

Similarly, the determination of $T_{C,N}^{cl}$ depends crucially on the accurate location of the maximum of the heat capacity variation with temperature. This was demonstrated in our previous report (Andriotis et al. 2007) when discussing the $T_{\text{melt},N}^{cl}$ and $T_{C,N}^{cl}$ results of Ni_{43} and their deviation from the prediction of Eq. 2. Additionally, the surface energy contribution to the free energy of the cluster may be another reason that small cluster temperatures cannot be extrapolated to bulk phase values (Qi et al. 2001).

Following these clarifications, we next discuss some obvious and hidden consequences of Eq. 2. One of the first conclusion that can be deduced from these results is that the Curie temperature, $T_{C,N}^{cl}$, of a magnetic cluster has to follow a size-dependence relationship analogous to that of $T_{\text{melt},N}^{cl}$. That is, if $T_{\text{melt},N}^{cl}$ is given by Eq. 1, then $T_{C,N}^{cl}$ has to follow the following equation:

$$T_{C,N}^{cl} = T_C^{\text{bulk}} - \delta_C N^{-1/3}, \quad (28)$$

where δ_C is a constant that may have an N-dependence as δ_{melt} . In fact, assuming the validity of Eqs. 1 and 28 and taking the ratio between Eqs. 1 and 28 by parts, we obtain:

$$\frac{T_{C,N}^{cl} - T_C^{\text{bulk}}}{T_{\text{melt},N}^{cl} - T_{\text{melt}}^{\text{bulk}}} = \frac{\delta_C}{\delta_{\text{melt}}} = \delta_0, \quad (29)$$

where δ_0 is a constant which is expected to depend on N. It is apparent that Eq. 29 has exactly the form of our Eq. 2, suggesting that the N-dependence of the constant δ_0 is very weak.

It is worth noting that the numerical justification of Eq. 29 by our results as expressed by Eq. 2 cannot ensure that the trends described by Eqs. 1 and 28 and which are possibly valid for large clusters can be extrapolated to small clusters as well. The conclusion that comes from Eq. 2 is that whatever the functional relationship between the melting temperature of a cluster and its size (not necessarily limited to that of Eq. 1) is, the functional relationship followed by $T_{\text{melt},N}^{cl}$ should dictate the relationship between the Curie temperature with its size as well. This hypothesis is supported by the results of Diep and collaborators (Diep et al. 1989) who found that the incorporation of the magnetic interactions leaves the cluster structure unchanged, thereby justifying our computational procedure.

One may argue that the use of two noncomparable methods, i.e., that of a classical potential MD simulation for calculating the melting temperature and a quantum mechanical approximation for calculating the magnetic moments and the Curie temperature of a cluster, cannot lead to results that can be correlated. We addressed this issue by fitting the classical Sutton–Chen potential to the data for

small clusters in such a way that resulted in TBMD and fitted Sutton–Chen potential simulations giving the same structural properties for small Ni clusters. Obtaining similar structural results by both methods appears to confirm the validity of the classical potential MD simulations for our present purpose.

Furthermore, in order to resolve any reservations and ambiguities about the consistency of our conclusions derived from the use of mutually inconsistent methods in calculating the melting and Curie temperatures, it is demonstrated in the following that a calculation of the melting temperatures within our TBMD approximation is in excellent agreement with the results of the classical potential approximation used in the derivation of Eq. 2.

Following exactly the same procedure as the one we used to calculate the average magnetic moment per cluster atom (and from this the Curie temperature) (Andriotis et al. 2007), we calculate the average total energy, $\langle E_T \rangle$, of each cluster at its thermodynamic equilibrium at a series of temperatures T . An average over N_{ran} random spin configurations over the cluster atoms is taken at each k th time step for N_{cl} time steps (N_{ran} is taken approximately between 100 and 200). Finally, these spin-averaged values are averaged over time. That is,

$$\langle E_T \rangle = \frac{1}{N_{cl}} \sum_{k=1}^{N_{cl}} \frac{1}{N_{\text{ran}}} \frac{\sum_{i=1}^{N_{\text{ran}}} E_i^k e^{-(E_i^k - E_0)/k_B T}}{\sum_{i=1}^{N_{\text{ran}}} e^{-(E_i^k - E_0)/k_B T}}. \quad (30)$$

where E_0 is a reference energy (Andriotis et al. 2007). For completeness, it is recalled that

$$E_i^k = \sum_{j\sigma} \varepsilon_{j\sigma}^{i,k} + E_{\text{rep}}^{i,k} \quad (31)$$

where $\varepsilon_{j\sigma}^{i,k}$ denote the eigenvalues of the TB cluster Hamiltonian and $E_{\text{rep}}^{i,k}$, the sum of the repulsive interactions (Andriotis and Menon 1998) of the cluster at the i th random spin configuration and the k th time step. As in the case of the calculation of the Curie temperatures, the averaging process over time in Eq. 30 is performed while reaching the thermodynamic equilibrium every 100 time-steps.

In Figs. 2 and 3, we present our TB results for the variation with temperature of $\langle E_T \rangle$ for the Ni₁₄₇ and Ni₂₀₁ clusters. For Ni₁₄₇, it is observed that the onset of a phase change appears at ≈ 450 °K while the melting starts at ≈ 700 °K. For Ni₂₀₁, the melting takes place at ≈ 800 °K (taken to be the midpoint of the “parallel” shift of the two linear parts of the thermodynamic curve. The so obtained melting temperatures appear to be in excellent agreement with the results found using the classical potential method and the Lindemann criterion according to which $T_{\text{melt},N=147}^{cl} = 700$ °K and $T_{\text{melt},N=201}^{cl} = 780$ °K.

Further headway can be made if Eq. 2 is taken to be the zeroth-order approximation of a *piece-wise function of N* in analogy to similar findings (Gunes and Erkoç 2000; Qi et al. 2001) for the expression for $T_{\text{melt},N}^{cl}$ given by Eq. 1. In this view, the results of Diep et al. (1989) (see Fig. 12 of their work) in the extreme case of very small clusters ($N \in [7,17]$) can lend support to our results and conclusions.

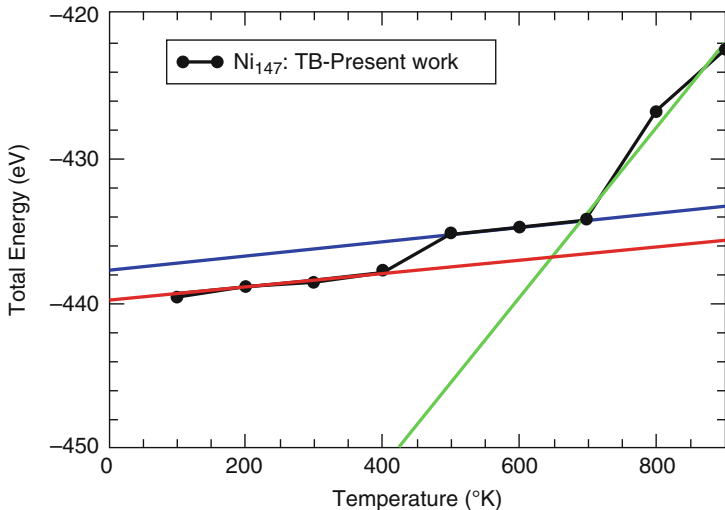


Fig. 2 Numerical results obtained within the TBMD method for the variation with temperature of the average total energy of the Ni₁₄₇ cluster. Straight lines are least square fits to portions of the numerical data

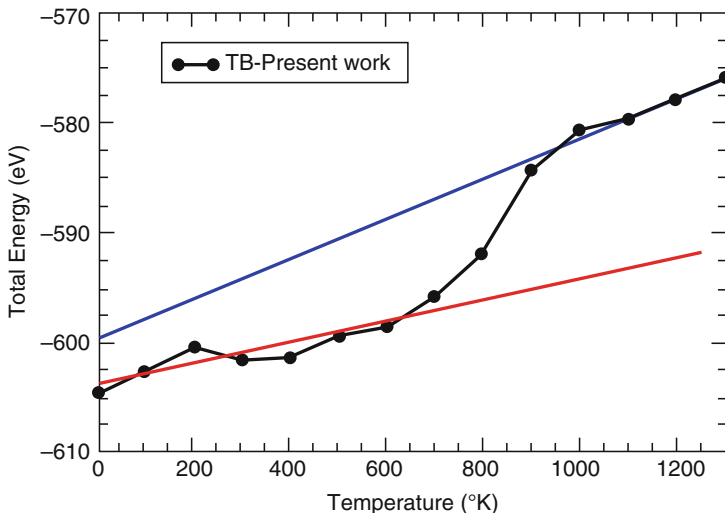


Fig. 3 The same as in Fig. 2 but for the Ni₂₀₁ cluster

The appearance of the nonzero constant term at the right-hand side of Eq. 2 indicates that there is a possibility for $T_{C,N}^{cl}$ to be greater than $T_{melt,N}^{cl}$. In particular, Eq. 2 predicts that $T_{C,N}^{cl}$ could be greater than $T_{melt,N}^{cl}$ if

$$T_C^{bulk} > T_{C,N}^{cl} > \frac{\beta}{1 - \alpha}. \tag{32}$$

However, according to our results, the above inequality does not hold for the Ni clusters since $\beta / (1 - \alpha) \approx 968^\circ \text{K}$, a value much greater than $T_C^{\text{bulk}} = 631^\circ \text{K}$. This is demonstrated in Fig. 1 with the indicated crossing of the dotted and dashed black lines with the former describing Eq. 2 and the latter describing the same equation taking $\alpha = 1$ and $\beta = 0$ (i.e., corresponding to the $T_{C,N}^{cl} = T_{\text{melt},N}^{cl}$ case). This incompatibility can be taken as an indication that the functional forms dictated by Eqs. 1 and 28 are not valid for the entire range of the cluster sizes.

The nonzero value of the constant β has another consequence; it predicts that the ratio $\frac{T_{\text{melt},N}^{cl}}{T_{C,N}^{cl}}$ depends on the cluster size and, in fact, increases as the cluster size decreases. If the variation of $T_{\text{melt},N}^{cl}$ is assumed as given by Eq. 1, the predictions of Eq. 2 lead to the conclusion that $T_{C,N}^{cl}$ decreases at a slower rate than $T_{\text{melt},N}^{cl}$ as the cluster size decreases and, therefore, the “melting temperatures” of *partially molten clusters* can be found to be lower than the Curie temperatures as Eq. 2 implies. However, such a conclusion has to be taken with care as the validity of Eq. 1 over the entire range of cluster sizes is not valid.

Conclusion

We have presented results for the variation with the cluster size of the melting and Curie temperatures of Ni_n , $n \leq 201$, clusters. Two complimentary methods were used, i.e., the classical MD employing the Sutton–Chen potential and the TBMD for obtaining the melting and the Curie temperatures of the clusters. We have demonstrated that by fitting the classical potential to the results of the TB description in the case of small clusters, we can achieve excellent agreement between the results of the two methods referring to the structural properties and the estimation of the melting temperatures of the clusters.

Our results demonstrate without any ambiguity that the variation of the cluster properties with the cluster size exhibits strong dependence on the ratio of the surface to bulk (core) contributions, the latter appearing to have the same functional dependence on the cluster size for both $T_{C,N}^{cl}$ and $T_{\text{melt},N}^{cl}$.

In view of the established dependence of the melting temperature of a cluster on its surface to volume contribution (as, for example, Eq. 1), our conclusion can be interpreted as an indication of a universal aspect of the surface to volume contribution to the cluster properties. This justifies previous findings based on approximate and semiempirical approximations.

Acknowledgments The present work is supported by grants from US-DOE (DE-FG02-00ER45817 and DE-FG02-07ER46375).

Bibliography

- Anderson, P. W., & Hasegawa, H. (1955). Considerations on double exchange. *Physical Review*, 100, 675.
- Andriotis, A. N., & Menon, M. (1998). Tight-binding molecular-dynamics study of ferromagnetic clusters. *Physical Review B*, 57, 10069.

- Andriotis, A. N., & Menon, M. (2001). Greens function embedding approach to quantum conductivity of single wall carbon nanotubes. *Journal of Chemical Physics*, *115*, 2737.
- Andriotis, A. N., & Menon, M. (2004). Orbital magnetism: Pros and cons for enhancing the cluster magnetism. *Physical Review Letters*, *93*, 026402.
- Andriotis, A. N., Menon, M., Froudakis, G. E., Fthenakis, Z., & Lowther, J. E. (1998). A tight-binding molecular dynamics study of ni(m)si(n) binary clusters. *Chemical Physics Letters*, *292*, 487.
- Andriotis, A. N., Menon, M., Froudakis, G. E., & Lowther, J. E. (1999). Tight-binding molecular dynamics study of transition metal carbide clusters. *Chemical Physics Letters*, *301*, 503.
- Andriotis, A. N., Menon, M., & Froudakis, G. E. (2000). Contrasting bonding behaviors of 3-d transition metal atoms with graphite and c60. *Physical Review B*, *62*, 9867.
- Andriotis, A. N., Menon, M., & Srivastava, D. (2002). Transfer matrix approach to quantum conductivity calculations in single wall carbon nanotubes. *Journal of Chemical Physics*, *117*, 2836.
- Andriotis, A. N., Fthenakis, Z., & Menon, M. (2006). Theoretical study of the effect of temperature on the magnetism of transition metal clusters. *Europhysics Letters*, *76*, 1088.
- Andriotis, A. N., Fthenakis, Z. G., & Menon, M. (2007). Correlated variation of melting and curie temperatures of nickel clusters. *Physical Review B*, *75*, 073413.
- Baletto, F., & Ferrando, R. (2005). Structural properties of nanoclusters: Energetic, thermodynamic and kinetic effects. *Reviews of Modern Physics*, *77*, 371.
- Bansman, J., Baker, S. H., Binns, C., Blackman, J. A., Bucher, J. P., Dorantes-Davila, J., Dupuis, V., Favre, L., Kechrakos, D., Kleibert, A., Meiwes-Broer, K. H., Pastor, G. M., Perez, A., Toulemonde, O., Trohidou, K. N., Tuailon, J., & Xie, Y. (2005). Magnetic and structural properties of isolated and assembled clusters. *Surface Science Reports*, *56*, 189.
- Buffat, P., & Borel, J. P. (1976). Size effect on the melting temperature of gold particles. *Physical Review A*, *13*, 2287.
- Diep, H. T., Sawada, S., & Sugano, S. (1989). Melting and magnetic ordering in transition-metal microclusters. *Physical Review B*, *39*, 9252.
- Doye, J. P. K., & Calvo, F. (2001). Entropic effects on the size dependence of cluster structures. *Physical Review Letters*, *86*, 3570.
- Erkos, S. (2001). In D. Stauffer (Ed.), *Annual reviews of computational physics* (Vol. IX). Singapore: World Scientific.
- Fanourgakis, G. S., Farantos, S. C., Parneix, P., & Brechignac, P. (1997). An effective transition state for a complex cluster isomerization process: Comparison between anharmonic and harmonic models for mg^+ar_{12} . *Journal of Chemical Physics*, *106*, 4954.
- Fthenakis, Z., Andriotis, A. N., & Menon, M. (2003a). Temperature evolution of structural and magnetic properties of transition metal clusters. *Journal of Chemical Physics*, *119*, 10911.
- Fthenakis, Z., Andriotis, A. N., & Menon, M. (2003b). Understanding the structure of metal encapsulated Si cages and nanotubes. *Journal of Chemical Physics*, *119*, 10911.
- Garcia-Rodeja, J., Rey, C., Gallego, L. J., & Alonso, J. A. (1994). Molecular-dynamics study of the structures, binding energies, and melting of clusters of fcc transition and noble metals using the voter and chen version of the embedded-atom model. *Physical Review B*, *49*, 8495.
- Garrigos, R., Cheyssac, P., & Kofman, R. (1989). Melting for lead particles of very small sizes: Influence of surface phenomena. *Zeitschrift für Physik D*, *12*, 497.
- Gerion, D., Hirt, A., Billas, I. M. L., Chatelain, A., & de Heer, W. A. (2000). Experimental specific heat of iron, cobalt, and nickel clusters studied in a molecular beam. *Physical Review B*, *62*, 7491.
- Gunes, B., & Erkoc, S. (2000). Melting and fragmentation of nickel nanoparticles: Molecular-dynamics simulations. *International Journal of Modern Physics*, *11*, 1567.
- Harrison, W. (1980). *Electronic structure and properties of solids*. San Francisco: W. H. Freeman.
- Huang, H., Sun, C. Q., & Hing, P. (2000). Surface bond contraction and its effect on the nanometric sized lead zirconate titanate. *Journal of Physics: Condensed Matter*, *12*, L127.
- Kato, M., & Kokubo, F. (1994). Partially antiferromagnetic state in the triangular hubbard model. *Physical Review B*, *49*, 8864.

- Lai, S. L., Guo, J. Y., Petrova, V., Ramanath, G., & Allen, L. H. (1996). Size-dependent melting properties of small tin particles: Nanocalorimetric measurements. *Physical Review Letters*, *77*, 99.
- Lathiotakis, N. N., Andriotis, A. N., Menon, M., & Connolly, J. (1996). Tight binding molecular dynamics study of Ni clusters. *Journal of Chemical Physics*, *104*, 992.
- Lee, Y. J., Lee, E. K., Kim, S., & Nieminen, R. M. (2001). Effect of potential energy distribution on the melting of clusters. *Physical Review Letters*, *86*, 999.
- Menon, M., & Subbaswamy, K. R. (1997). Nonorthogonal tight-binding molecular-dynamics scheme for silicon with improved transferability. *Physical Review B*, *55*, 9231.
- Nayak, S. K., Khanna, S. N., Rao, B. K., & Jena, P. (1998). Thermodynamics of small nickel clusters. *Journal of Physics: Condensed Matter*, *10*, 10853.
- Nose, S. (1984). A unified formulation of the constant temperature molecular dynamics methods. *Journal of Chemical Physics*, *81*, 511.
- Ojeda, M. A., Dorantes-Davila, J., & Pastor, G. (1999). Noncollinear cluster magnetism in the framework of the hubbard model. *Physical Review B*, *60*, 6121.
- Qi, Y., Cagin, T., Johnson, W. L., & Goddard, W. A. (2001). Melting and crystallization in Ni nanoclusters: The mesoscale regime. *Journal of Chemical Physics*, *115*, 385.
- Rey, C., Gallego, L. J., Garcia-Rogea, J., Alonso, J. A., & Iniguez, M. P. (1993). Molecular-dynamics study of the binding energy and melting of transition-metal clusters. *Physical Review B*, *48*, 8253.
- Schmidt, M., Kusche, R., Kronmuller, W., von Issendorff, B., & Haberland, H. (1997). Experimental determination of the melting point and heat capacity for a free cluster of 139 sodium atoms. *Physical Review Letters*, *79*, 99.
- Sun, D. Y., & Gong, X. G. (1998). Structural properties and glass transition in Al_n clusters. *Physical Review B*, *57*, 4730.
- Sutton, A. P., & Chen, J. (1990). Long-range Finnis-Sinclair potentials. *Philosophical Magazine Letters*, *61*, 139.
- Uhl, M., Sanrdatskii, L. M., & Kubler, J. (1994). Spin fluctuations in γ -Fe and in Fe_3Pt invar from local-density-functional calculations. *Physical Review B*, *50*, 291.
- Weerasinghe, S., & Amar, F. G. (1993). Absolute classical densities of states for very anharmonic systems and applications to the evaporation of rare gas clusters. *Journal of Chemical Physics*, *98*, 4967.
- Yang, C. C., & Jiang, Q. (2005). Size and interface effects on critical temperatures of ferromagnetic, ferroelectric and superconductive nanocrystals. *Acta Materialia*, *53*, 3305.

Michael Springborg

Contents

Introduction	1358
Basics	1362
Only Nuclei	1363
Methods	1363
Ni Clusters	1367
Thermodynamic Properties	1373
Bimetallic Clusters	1376
Clusters on Surfaces	1381
Only Electrons	1383
Both Nuclei and Electrons	1384
Methods	1384
Na and Au Clusters	1385
HAIO Clusters	1391
AB Semiconductor Clusters	1396
Metcars	1405
Conclusions	1409
Bibliography	1410

Abstract

Clusters contain more than just some few atoms but not so many that they can be considered as being infinite. By varying their size, their properties can often be varied in a more or less controllable way. Often, however, the precise relation between size and property is largely unknown: the sizes of the systems are below the thermodynamic limit so that simple scaling laws do not apply. Theoretical

M. Springborg (✉)

Physical and Theoretical Chemistry, University of Saarland, Saarbrücken, Germany

School of Materials Science and Engineering, Tianjin University, Tianjin, China

e-mail: m.springborg@mx.uni-saarland.de

studies of such systems can provide relevant information, although in many cases idealized systems have to be treated. The challenge of such calculations is the combination of the relatively large size of the systems together with an often unknown structure.

In this presentation, different theoretical methods for circumventing these problems shall be discussed. They shall be illustrated through applications on various types of clusters. These include isolated metal clusters with one or two types of atoms, metal clusters deposited on a surface, nanostructured HAIO, semiconductor nanoparticles, and metallocarbohedrenes. Special emphasis is put on the construction of descriptors that can be used in identifying general trends.

Introduction

Nanostructures are materials whose spatial extension in 1, 2, or 3 dimensions is roughly of the order of at most some 100 nm. Thereby material properties different from those of the macroscopic materials we know from our daily life emerge. According to the simplest description of the properties of some given material, we may apply scaling laws so that, e.g., the electrical resistance scales with the inverse cross section and linearly with the length of the system. With such scaling laws, it is indirectly assumed that the materials form some homogeneous continua and, accordingly, that the fact can be ignored that on the atomic length scale the materials are not at all homogeneous. However, since typical interatomic bond lengths are of the order of some tenths of a nm, a proper description of the material properties of nanostructures often cannot be based on the above assumption that the materials are homogeneous. Thus, for nanomaterials, one can no longer apply the abovementioned scaling laws, i.e., the systems of interest have sizes far from the thermodynamic limit and one has left the scaling regime so that “every single atom counts.” Indeed, the fact that the material properties depend, in some cases even strongly, on the size of the system for nanostructures is the reason for the large interest in such systems. Here, both the prospects related to possible practical applications of these materials as well as the wish to understand in detail the size–property relations are motivations behind this interest.

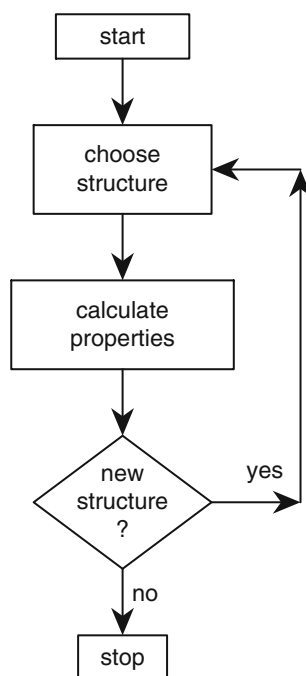
The present chapter is devoted to the results of our theoretical studies of clusters, i.e., nanostructures that are finite in all three spatial dimensions. Therefore, their number of surface atoms relative to their total number of atoms is large, and of the same reason, finite-size effects show up. In addition, they most often contain just one or, at most, some few types of atoms, with the possible existence of surfactants as an exception. The surfactants are radicals that are used in saturating dangling bonds at the surfaces and, thus, in stabilizing the materials. For theoretical studies, it is important to notice that these systems are fairly large, but not so large that they can be considered as being infinite and periodic and that, consequently, they may have a quite low symmetry so that in a theoretical calculation chemically identical atoms may have to be treated as being different since they have different surroundings.

In experimental studies, such clusters/colloids are often produced and studied in a solution or on the surface of some substrate. Moreover, as mentioned above, they are often stabilized through ligands. In total, this means that the systems often are interacting with other systems, this being the solvent, the substrate, or the ligands. This is contrasted by theoretical studies that often consider isolated nanoparticles in gas phase that, in addition, often are naked (i.e., without ligands). Another difference between experiment and theory is that the precise size of the experimentally studied systems often is only approximately known, whereas, per construction, theoretical studies consider well-defined systems, although for those one often assumes a certain structure that may or may not be realistic. Thus, experimental and theoretical studies are rarely competing but instead complementing each other.

When attempting to use electronic-structure methods in studying the properties of clusters, in particular the size dependence of those where one, accordingly, will consider a larger range of sizes, one very fast encounters some fundamental properties related to theoretical calculations. We shall here briefly outline those.

In a typical electronic-structure calculation, the number and types of the atoms for the system that shall be studied are known. Then (cf. Fig. 1), some initial structure is chosen and various properties for this structure are calculated. These properties are first of all the total energy but may also be others like the forces acting on the nuclei (i.e., the derivatives of the total energy with respect to the nuclear

Fig. 1 A flowchart for a typical electronic-structure calculation



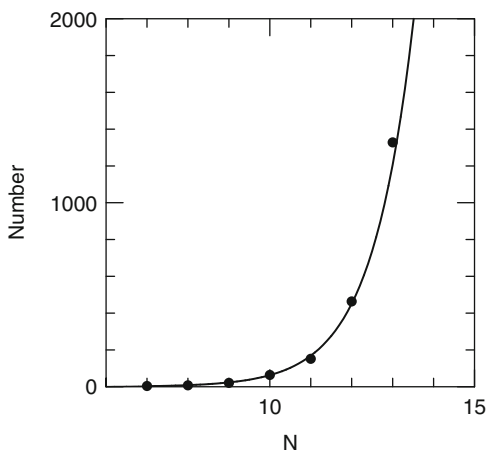
coordinates). Also available experimental information may be sought calculated. Subsequently, one may ask whether a new structure shall be studied. The reasons for doing so could be that the calculated forces are not approximately vanishing, that the calculated properties differ significantly from those obtained in experiment, or that one simply wants to explore further parts of the structure space.

For larger systems, this approach is related with two serious problems. First, if the system of interest contains N nuclei and M electrons, the computational costs for calculating the properties for one single structure scale as the size of the system to some power, i.e., like M^k or N^k , where k is some power from 2 and upward. For some of the most popular approaches, like density-functional and Hartree–Fock methods (see, e.g., Springborg 2000), k is typically 3, but for more advanced methods, k may be larger than 7. Thus, in this case, it should be obvious that for just intermediately large systems, the calculation of the properties for just one single geometry may become computationally very costly.

Independently of the scaling of the computational needs as a function of the size of the system, another complementary problem causes additional complications. This problem is related to the fact that the number of nonequivalent minima on the total-energy surface as a function of structure grows very fast with the size of the system. In fact, it has been shown (Wille and Vennik 1985) that the number of local total-energy minima grows faster than any polynomial in N or, alternatively expressed, that the determination of the global total-energy minimum is a so-called NP-hard problem.

A model system that is so simple that detailed studies can be performed is that of a cluster of identical atoms for which it is assumed that the total energy can be written as a sum of pair potentials, each one being a simple Lennard–Jones potential. This system was studied by Tsai and Jordan (1993), who found a rapidly increasing number of nonequivalent metastable structures as a function of size. In Fig. 2 their

Fig. 2 A schematic representation of the number of nonequivalent local total-energy minima for Lennard–Jones clusters (the *black circles*) together with a fit with an exponential (the *full curve*)



results have been fitted with an exponential, $a \cdot \exp(bN)$, and it is seen that the fit follows the calculated results fairly close. The fit gave $a = 0.00341$ and $b = 0.983$, which in turn means that for $N = 55$ the fit predicts that of the order of 10^{21} nonequivalent minima exist.

This result points directly to a central issue of this presentation, i.e., how can we determine the structure for a system with at least some 10s of atoms and for which no, or only very limited, information on the structure can be used.

The ground-state *structure* is obtained by minimizing the total energy. Employing the Born–Oppenheimer approximation, the nuclei are treated as classical and not quantum particles, and there is no variational principle that can be used in systematically getting arbitrarily close to the true ground-state structure, even when assuming that the total energy can be calculated accurately for any structure. Instead, most often one has to change the structure “by hand” until one feels confident that the structure of the lowest total energy has been identified. But it shall be stressed that there is absolutely no approach that with absolute certainty can guarantee that precisely that structure has been found. Trying to prove so would require complete knowledge about the total-energy surface as a function of all internal degrees of freedom. This information is beyond whatever will be available, even if the computers keep on becoming more and more powerful and the programs more and more efficient.

That this information will never become available can be seen by, e.g., considering the largest possible computer, i.e., the complete universe. Lloyd (2002) estimated that the complete universe, viewed as an enormous quantum computer that has been operating since the Big Bang, could have performed of the order of 10^{120} binary operations. Although this number is very large, it is also finite. Therefore, it means that considering, e.g., a system of 42 atoms, assuming that a total-energy calculation for this system and a given structure is just one single binary operation, and considering just 10 different values for each internal structural degree of freedom, exactly 10^{120} calculations are needed for obtaining this, relatively crude, total-energy hypersurface. Thus, the calculations would have required the use of the complete universe since the Big Bang!

In total, it is obvious that for any but the absolutely smallest systems it is not possible to explore anything but very limited parts of the total-energy hypersurface and, moreover, that it is never possible to be absolutely sure that the true global total-energy minimum has been found. Any total-energy minimum may provide a total energy that is close to that of the global minimum, although the structures of the two may be markedly different, simply due to the very large number of local total-energy minima. It is also obvious that any attempt to identify the global total-energy minimum has to be based on some kind of qualified search in the multidimensional structure space.

It is one of the purposes of this presentation to discuss some of the approaches that we are using in searching for the global total-energy minimum in the structure space. Another purpose is to present various descriptors/tools that we have derived in order to extract chemical or physical information from the calculated properties.

We emphasize that we shall concentrate on the results of our own studies. For more general reviews the reader is referred to, e.g., Baletto and Ferrando (2005), Ferrando et al. (2008), and Springborg (2009).

Basics

In the most general case, our goal is to calculate the properties of a system containing N nuclei and M electrons. We shall denote the nuclear positions $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N$ and those of the electrons $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_M$. Moreover, we use Hartree atomic units and set accordingly $m_e = |e| = 4\pi\epsilon_0 = \hbar \equiv 1$. The mass and charge of the k th nucleus is then M_k and Z_k , respectively. The combined coordinate \mathbf{x}_i denotes the position and spin coordinate of the i th electron. In the absence of external interactions and relativistic effects, the Hamilton operator for this system can then be written as a sum of five terms,

$$\hat{H} = \hat{H}_{kn} + \hat{H}_{ke} + \hat{H}_{nn} + \hat{H}_{ee} + \hat{H}_{en}, \quad (1)$$

with

$$\begin{aligned} \hat{H}_{kn} &= - \sum_{k=1}^N \frac{1}{2M_k} \nabla_{\mathbf{R}_k}^2 \\ \hat{H}_{ke} &= - \sum_{i=1}^M \frac{1}{2} \nabla_{\mathbf{r}_i}^2 \\ \hat{H}_{nn} &= \frac{1}{2} \sum_{k \neq l=1}^N \frac{Z_k Z_l}{|\mathbf{R}_k - \mathbf{R}_l|} \\ \hat{H}_{ee} &= \frac{1}{2} \sum_{i \neq j=1}^M \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \\ \hat{H}_{en} &= - \sum_{k=1}^N \sum_{i=1}^M \frac{Z_k}{|\mathbf{R}_k - \mathbf{r}_i|}, \end{aligned} \quad (2)$$

i.e., the kinetic energy operator for the nuclei, that for the electrons, and the three potential energy operators for the nucleus–nucleus, the electron–electron, and the electron–nucleus interactions, respectively.

The time-independent Schrödinger equation becomes then

$$\hat{H}\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) = E \cdot \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N). \quad (3)$$

As indicated above, we shall assume that the Born–Oppenheimer approximation is accurate. Then, \hat{H}_{kn} is ignored, and furthermore, the wave function Ψ that depends

functionally on both electronic and nuclear coordinates is written as a product of two functions,

$$\begin{aligned} \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \\ = \Psi_n(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \cdot \Psi_e(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N; \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M). \end{aligned} \quad (4)$$

Here, Ψ_n is the wave function for the nuclei (which for the present study is irrelevant) and Ψ_e is the electronic wave function. Ψ_e depends functionally on the electronic coordinates but also parametrically on the nuclear coordinates (implying that different structures have different electronic wave functions). Ψ_e is calculated from the electronic Schrödinger equation

$$\hat{H}_e \Psi_e = E_e \Psi_e \quad (5)$$

with

$$\hat{H}_e = \hat{H}_{ke} + \hat{H}_{ee} + \hat{H}_{en}. \quad (6)$$

Moreover, the total energy is

$$E = E_e + \hat{H}_{nn} = E(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N), \quad (7)$$

i.e., a function of the structure of the system.

Although the Born–Oppenheimer approximation does lead to some computational simplifications, solving the resulting equations is still computational demanding, so that when aiming at studying *many* structures for *larger* systems (as is the goal of the present work), it can easily become necessary to invoke additional approximations. Then, the type of approximation will depend on the system and scientific issue at hand, as we shall illustrate below.

Only Nuclei

Methods

Ultimately, the goal is to study the total energy, E , as a function of structure, i.e., $E(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N)$. The determination of the electronic orbitals, their energies, and the total electronic energy, E_e , provides thereby additional information that may be highly relevant for other purposes, but may also be considered as a complication that might be circumvented. The use of approximate expressions (i.e., force fields) of the form

$$\begin{aligned} E(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) \simeq \sum_i E_{t(i)}^{(1)} + \sum_{ij} E_{t(i),t(j)}^{(2)}(\mathbf{R}_i, \mathbf{R}_j) \\ + \sum_{ijk} E_{t(i),t(j),t(k)}^{(3)}(\mathbf{R}_i, \mathbf{R}_j, \mathbf{R}_k) + \dots \end{aligned} \quad (8)$$

represents one possibility. Here, $t(i)$ is the atom type of the i th atom, and the total energy has been split into atomic parts as well as 2-, 3-, ... body interactions. The various functions $E^{(i)}$ are then represented in some approximate, analytical, or numerical form.

Such approximations are very useful for the study of the structural properties of a larger number of larger systems, but may, of course, suffer from inaccuracies due to the approximate nature of the total-energy expression. The Lennard–Jones clusters discussed above represents one such expression, where only 1- and 2-body terms are included. Some further examples are discussed in Springborg (2006).

When applying such approaches to clusters, it is important to choose the method according to the chemical bonding of the system of interest. In metallic systems, the electrons are considered as being delocalized and packing effects are often the main driving force for the structure. On the other hand, for many semiconductors and insulators, the bonds between the atoms are partly covalent and partly ionic, so that directional bonding and electron transfers are important, i.e., a precise description of the electronic degrees of freedom is important. Yet other types of systems include, e.g., molecular crystals with weak interactions between the molecules, but here we are concerned with the first two types of systems.

One approach that we have been using for clusters of metal atoms is based on the embedded-atom method (EAM) of Daw, Baskes, and Foiles (Daw and Baskes 1983, 1984; Daw et al. 1993; Foiles et al. 1986). According to this method, the total energy of the system is written as a sum over atomic energies,

$$E_{\text{tot}} = \sum_{i=1}^N E_i \quad (9)$$

with

$$E_i = F_i(\rho_i^h) + \frac{1}{2} \sum_{i \neq j=1}^N \phi_{ij}(|\mathbf{R}_i - \mathbf{R}_j|) \quad (10)$$

being the energy of the i th atom.

In Eq. (10) the first term is the so-called embedding energy, which is obtained by considering each atom as an impurity embedded into a host provided by the rest of the atoms. The second term describes electron–electron interactions and is represented in terms of short-range pair potentials. The local density at site i is assumed being a superposition of atomic electron densities,

$$\rho_i^h = \sum_{j (\neq i)} \rho_j^a(|\mathbf{R}_i - \mathbf{R}_j|) \quad (11)$$

where $\rho_j^a(r)$ is the spherically averaged atomic electron density provided by atom j at the distance r .

With the EAM, the total-energy expression of Eq. (8) contains also contributions from P -body interactions with P being the number of nearest neighbors of the atoms of the system of interest. Other potentials that also include many-body interactions are the Gupta, Murrell–Mottram, Sutton–Chen, and Cleri–Rosato potentials among others. Since we have used the Gupta, Sutton–Chen, and Cleri–Rosato potentials in some studies, we shall here discuss these briefly.

With the Gupta potential (Gupta 1981), the total energy is written in terms of repulsive and attractive many-body terms,

$$E_{\text{tot}} = \sum_{i=1}^N [V^r(i) - V^m(i)] \quad (12)$$

where

$$V^r(i) = \sum_{j=1(\neq i)}^N A(a, b) \exp \left[-p(a, b) \left(\frac{|\mathbf{R}_i - \mathbf{R}_j|}{r_0(a, b)} - 1 \right) \right] \quad (13)$$

and

$$V^m(i) = \left\{ \sum_{j=1(\neq i)}^N \zeta^2(a, b) \exp \left[-2q(a, b) \left(\frac{|\mathbf{R}_i - \mathbf{R}_j|}{r_0(a, b)} - 1 \right) \right] \right\}^{\frac{1}{2}}. \quad (14)$$

The summations run over all (N) atoms of the system of interest. Furthermore, A , r_0 , ζ , p , and q are fitted to experimental values of the cohesive energy, lattice parameters, and elastic constants for the crystal at 0 K.

The Sutton–Chen potentials are based on the empirical N -body potentials that have been developed by Finnis and Sinclair (1984) for the description of cohesion in metals. With these potentials, the total internal energy is represented by a cohesive functional of pair interactions and a predominantly repulsive pair potential. The main difference of such a potential to a pair potential is that using only pair potentials, the calculated force exerted by one atom on another depends on the interatomic distance only, whereas in the present case it depends on all neighbors of both atoms. Finnis–Sinclair potentials are of relatively short range and extend only to third neighbors in fcc crystals. The long-range modification constitutes the Sutton–Chen potential. This potential has an extra $\frac{1}{r^6}$ van der Waals tail and, accordingly, the following form

$$E = \epsilon \left[\frac{1}{2} \sum_i \sum_{j \neq i} V(|\mathbf{R}_i - \mathbf{R}_j|) - c \sum_i \sqrt{\rho_i} \right] \quad (15)$$

with

$$V(|\mathbf{R}_i - \mathbf{R}_j|) = \left(\frac{a}{|\mathbf{R}_i - \mathbf{R}_j|} \right)^n$$

$$\rho_i = \sum_{j \neq i} \left(\frac{a}{|\mathbf{R}_i - \mathbf{R}_j|} \right)^m. \quad (16)$$

With the empirical many-body potential given by Cleri and Rosato (1993), the cohesive energy of the system is written as

$$E_c = \sum_i (E_R^i + E_B^i), \quad (17)$$

where

$$E_R^i = \sum_j A e^{-p \left(\frac{|\mathbf{R}_i - \mathbf{R}_j|}{r_0} - 1 \right)}$$

$$E_B^i = - \left\{ \sum_j \xi^2 e^{-2q \left(\frac{|\mathbf{R}_i - \mathbf{R}_j|}{r_0} - 1 \right)} \right\}^{\frac{1}{2}}. \quad (18)$$

As for the Gupta potential, the values of the parameters that enter the Sutton–Chen and Cleri–Rosato potentials are determined from experimental information.

In a series of studies, we have studied the properties of various metal clusters for which the total energy was approximated using one of those methods just mentioned. These include Ni (Grigoryan and Springborg 2001, 2003, 2004; Grigoryan et al. 2006), Al (Joswig and Springborg 2003), Ti (Joswig and Springborg 2007), Cu (Grigoryan et al. 2006), Au (Grigoryan et al. 2006; Alamanova et al. 2006), Ag (Alamanova et al. 2007), Ni–Cu (Hristova et al. 2008b), and K–Cs and Rb–Cs (Hristova et al. 2008a, 2009a) clusters.

In all cases, the total-energy methods were combined with unbiased global optimization methods. For the studies we shall report below in this section, we used the Aufbau/Abbau method that was developed earlier by us (Grigoryan and Springborg 2003, 2004) and which consists of the following steps:

1. We assume that we have obtained the structures of the global total-energy minima for the clusters with up to N atoms.
2. In addition, we consider clusters with $N + K$ atoms with $K \simeq 5 - 10$. For each of those sizes (i.e., N and $N + K$), we randomly generate and relax a large set of structures, from which the one of the lowest total energy is selected.
3. This leaves us with two “source” clusters with N and $N + K$ atoms. For the cluster with N atoms, we add at a random position a single atom and the structure

is relaxed. This is repeated many hundred times. In parallel, we consider the $N + K$ structures of $N + K - 1$ atoms that can be obtained by removing one of the atoms from the cluster with $N + K$ atoms and relaxing the structure. The structure of the lowest total energy is kept.

4. The procedure of step (3) is repeated $K - 1$ times until we have obtained two sequences of cluster structures with $N, N + 1, \dots, N + K - 1$, and $N + K$ atoms. From these two sets, the structures of the lowest total energies are chosen.
5. The steps (3) and (4) are repeated with the new sets of structures, until no lower total energies are found. Then it is assumed that the global total-energy minima for this size range have been identified, and we proceed to the next set of sizes.

By keeping track of not only the single energetically lowest isomer but more of those, we also obtain information about the energetic distribution of different isomers of the same cluster sizes.

In some studies, we used instead the basin-hopping method (Li and Scheraga 1987; Wales and Doey 1997) to determine the structures of total-energy minima. This method is closely related to Monte Carlo methods, but considers a transformed total-energy surface. Thus, for each structure, $\mathbf{R} \equiv (\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N)$, instead of considering the true total energy E at that structure, one considers the total energy \tilde{E} that the structure would have after having relaxed to its closest total-energy minimum,

$$\tilde{E}(\mathbf{R}) = \min\{E(\mathbf{R})\}, \quad (19)$$

where it is indicated that $\tilde{E}(\mathbf{R})$ is the lowest total energy that is obtained when starting from the structure \mathbf{R} but letting it locally relax. Ultimately, this leads to a transformed total-energy surface. The transformed energy surface \tilde{E} is much less structured than E , although the energies at the global and local total-energy minima are identical. The simpler shape of the transformed structure makes it significantly easier to identify the global total-energy minimum, without biasing the calculation in any way. In a practical calculation, instead of changing the structure according to the Monte Carlo steps directly, the structure is first locally relaxed using standard methods.

Ni Clusters

As a first example of isolated monatomic metal clusters, we shall in this subsection discuss some of our results for isolated Ni_N clusters (Grigoryan and Springborg 2001, 2003, 2004; Grigoryan et al. 2006).

The calculations outlined in the previous subsection give essentially nothing but a listing of the total energies and the nuclear coordinates for a set of clusters. Thus, in our study on the Ni_N clusters, we considered the four energetically lowest isomers up to $N = 150$ and obtained, accordingly, roughly 45,000 numbers. A large

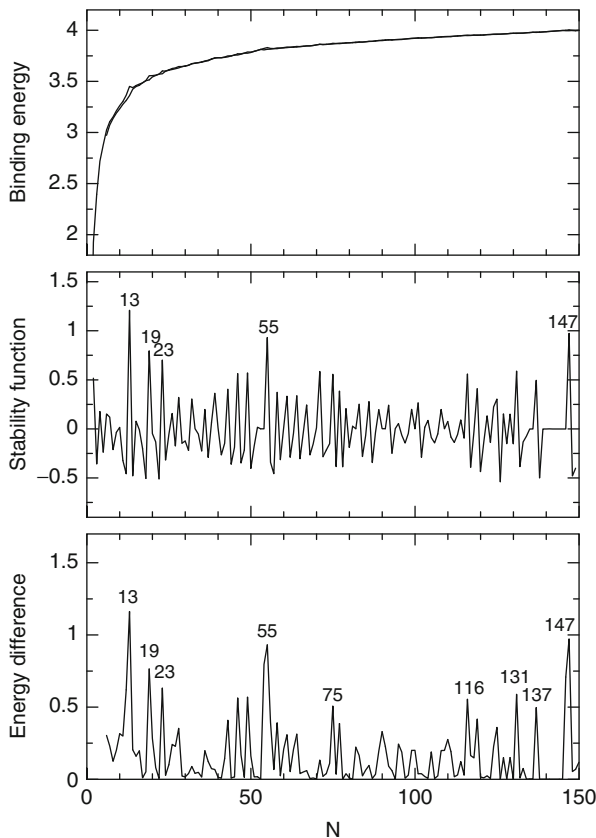


Fig. 3 The binding energy per atom for the energetically two lowest isomers (*upper part*), the stability function (*middle part*), and the total-energy difference for the two energetically lowest isomers (*lowest part*) for Ni_N clusters as functions of N . All energies are given in eV

challenge is, consequently, to extract physical and chemical information from this information.

In Fig. 3 we show in the upper panel the binding energy per atom. A careful inspection of the figure reveals that there are two curves very close in energy, since the results are given for the two energetically lowest isomers. Nevertheless, the curves are rather structureless which is a result of the fact that any cluster Ni_N cluster is stabler than any two noninteracting subparts Ni_{N-K} and Ni_K for all $K > 0$.

More information can be obtained from the so-called stability function

$$\Delta_2 E(N) = E_{\text{tot}}(N-1, 1) + E_{\text{tot}}(N+1, 1) - 2E_{\text{tot}}(N, 1) \quad (20)$$

with $E_{\text{tot}}(N, j)$ being the total energy for the energetically j th lowest isomer of the Ni_N cluster. $\Delta_2 E(N)$ possesses peaks for clusters that are particularly stable compared to those with one atom more or less. As seen in the middle panel in Fig. 3,

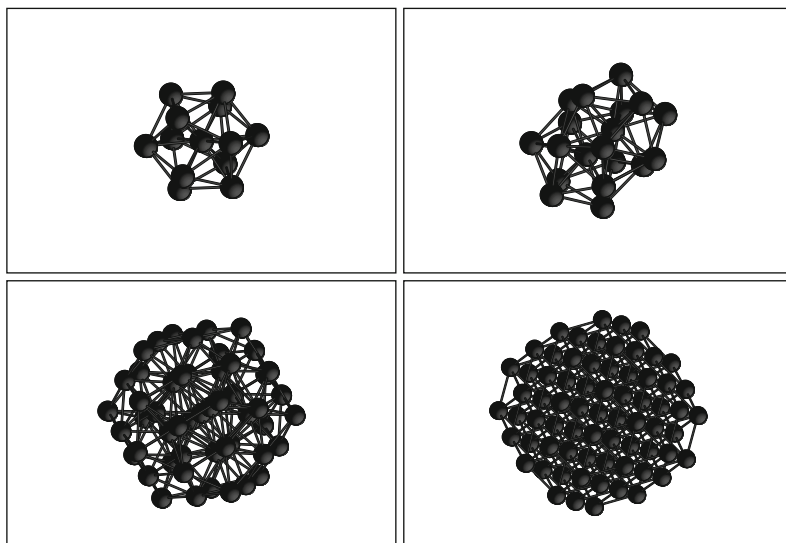


Fig. 4 The structures of the Ni_N clusters with (top, left) $N = 13$, (top, right) 19, (bottom, left) 55, and (bottom, right) 78

this is the case for clusters with $N = 13, 19, 23, 55$, and 147. These, so-called magic numbers, correspond to clusters with a particularly high-symmetric structure as can be seen from a more careful inspection of the structures. As examples we show in Fig. 4 the structures of the Ni_N clusters with $N = 13, 19, 55$, and 78.

Alternatively, one may compare the total energies for the energetically two lowest isomers of the same cluster size, i.e., $E_{\text{tot}}(N, 2) - E_{\text{tot}}(N, 1)$. This has peaks for clusters that are particularly stable compared to the next isomer of the same cluster size. As seen in Fig. 3, many of the magic-numbered clusters are also particularly stable according to this second criterion. This suggests that the characterization of certain clusters to be particularly stable is very robust.

In order to obtain information about the overall shape of the clusters, we consider the 3×3 matrix containing the elements

$$I_{st} = \frac{1}{u_l^2} \sum_{n=1}^N (R_{n,s} - R_{0,s})(R_{n,t} - R_{0,t}) \quad (21)$$

with $u_l = 1 \text{ \AA}$ being a length unit; s and t being x, y , or z , and

$$\mathbf{R}_0 = \frac{1}{N} \sum_{n=1}^N \mathbf{R}_n \quad (22)$$

being the center of the cluster; and R_n being the position of the n th atom. The three eigenvalues of this matrix, $I_{\alpha\alpha}$, can be used in separating the clusters into

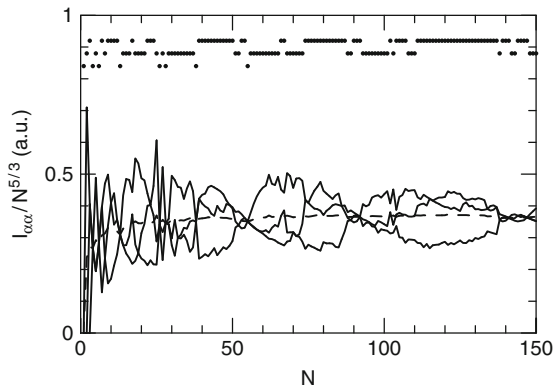


Fig. 5 The normalized eigenvalues of the matrix containing the elements of Eq. (21) as functions of N . The three *solid curves* give the eigenvalues and the *dashed curve* their average. In the *upper part* of the figure, a simple estimate of the overall shape of the clusters is given: either being spherical (marked with points in the lowest row), cigar shaped (points in the middle row), or lens shaped (points in the uppermost row)

being overall spherical (all eigenvalues are identical), more cigar-like shaped (one eigenvalue is large, the other two are small), or more lens shaped (two large and one small eigenvalue). Moreover, the average of the three eigenvalues, $\langle I_{\alpha\alpha} \rangle$, is a measure of the overall extension of the cluster. Finally, the sum of any two of those three eigenvalues gives a moment of inertia that is relevant for rotational properties of the clusters.

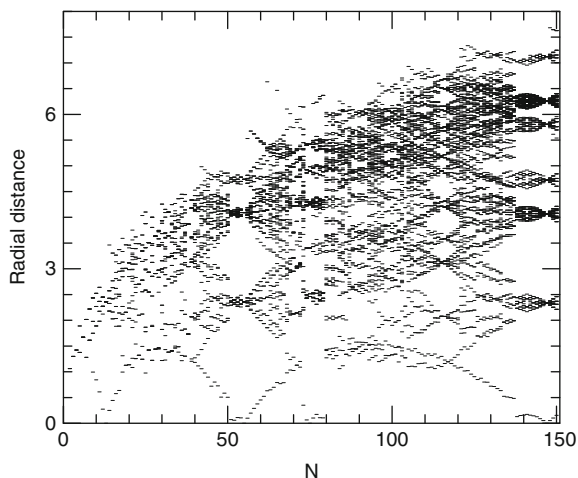
For a homogeneous, spherical cluster with N atoms, the eigenvalues are proportional to $N^{5/3}$. Therefore, in Fig. 5 we scale the eigenvalues by $N^{-5/3}$. This figure shows that the overall shape changes many times for clusters with N up to somewhat above 50. After that, the shape is more constant, with some few exceptions that in some cases are due to the fact that the three eigenvalues all are very close so that only small atomic displacements would change the shape classification.

In order to obtain further information on the structure of the clusters, we introduce the concept of radial distances. For each atom, its radial distance is defined as

$$r_n = |\mathbf{R}_n - \mathbf{R}_0|. \quad (23)$$

By plotting various properties as functions of the radial distances, we can obtain further insight into the cluster properties. At first we show the radial distances themselves; cf. Fig. 6. This figure shows that the overall size of the cluster (using the largest radial distance as a measure) in general increases with size of the cluster, although not monotonically. Moreover, for certain size ranges, like slightly above 50 and 140 atoms, the radial distances tend to form different groups. This implies that those clusters have an onion-like structure with different, essentially concentric, shells of atoms. For other sizes, the clusters tend to have a much more irregular structure.

Fig. 6 The radial distances (in Å) as a function of cluster size, i.e., each small line represents (at least) one atom with that radial distance

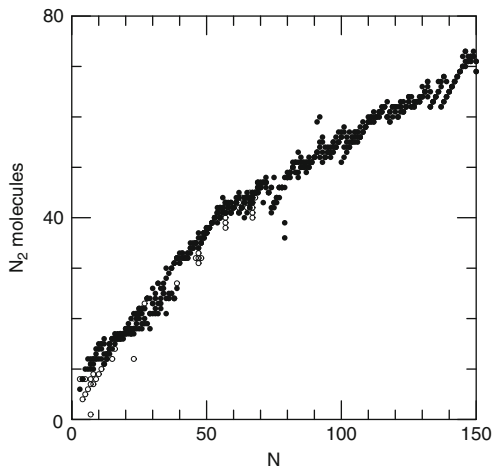


Riley and coworkers (Parks et al. 1997, 1998, 2001) have presented experimental results of so-called chemical probe experiments on Ni_N clusters. By measuring the mass of the clusters before and after exposure to (in this case) an N_2 atmosphere, the number of N_2 molecules that are adsorbed on the surface of the cluster can be determined, which in turn gives information on the surface of the cluster. Thereby, Riley et al. make use of some simple empirical rules for the adsorption: Nickel atoms with a coordination of 4 or less will bind two molecules; for a coordination between 5 and 8, they will bind one molecule; for a coordination of 9, there may be one molecule bonded; and for larger coordination, there will be no molecule bonded. Using the same rules, we calculated the number of N_2 molecules that could be adsorbed on the structures we have determined for the four energetically lowest isomers and show the results, in comparison with experiment, in Fig. 7. It is here remarkable that there is a very good agreement between theory and experiment and, moreover, that the overall shape of the size dependence of this number appears to change slope for clusters with some 60–70 atoms. Unfortunately, the experimental studies do not allow for analyzing this prediction.

For the study of growth processes as well as for analyzing the size for which the clusters obtain structures that resemble small parts of the macroscopic crystalline material, we have found the concept of similarity functions useful. In order to compare two structures consisting of N_1 and N_2 atoms, we shall, for each, calculate either the interatomic distances or the radial distances, subsequently sort these, and, finally, define a so-called similarity function through

$$S = \left\{ 1 + \left[\frac{1}{M} \sum_{i=1}^M (d_i - d'_i)^2 \right]^{1/2} / u_l \right\}^{-1}. \quad (24)$$

Fig. 7 The (*open circles*) experimentally determined number of N_2 binding sites for Ni_N clusters as function of N in comparison with theoretical calculated numbers (*closed circles*)



Here, $u_l = 1 \text{ \AA}$. Moreover $\{d_i\}$ and $\{d'_i\}$ are the sorted radial or interatomic distances for the two structures, and M is their total number [being either N or $N(N - 1)/2$ with N the smallest of N_1 and N_2]. The two structures are very similar (different) if S approaches 1 (0). We add that our experience has shown that qualitatively the same results are obtained independently of whether the radial or the interatomic distances are used. Another result of our experience is that S below roughly 0.7–0.8 usually indicates that the structures are quite different.

At first, we shall use the similarity functions in studying whether the structure with N atoms can be considered similar to the one with $N - 1$ atoms plus an extra atom. To this end, we consider each of the N structures that can be obtained by removing one of the atoms from the N atom cluster and keeping the positions of the remaining $N - 1$ atoms. Then, we compare this $N - 1$ atom fragment with the cluster with $N - 1$ atoms, using the similarity function. Ultimately, the largest of these N values for S is used. Figure 8a shows the results when comparing the energetically lowest isomers of the two clusters with $N - 1$ and N atoms. Also for this property we see many dips in the function for cluster sizes up to around 50 atoms after which S gets much more regular. This effect becomes even more pronounced when comparing the Ni_N cluster of the lowest energy with the four energetically lowest isomers with $N - 1$ atoms. Then [cf. Fig. 8b], S approaches 1 for almost all $N > 50$, suggesting that more different isomers are active in the growth processes: the most stable Ni_N cluster has a structure that may be derived from that of an energetically higher structure of the Ni_{N-1} cluster.

Finally, we show in Fig. 8 also the similarity function that is obtained when comparing the cluster structures with fragments of the crystal or with an icosahedron. In this case, the similarity function is based on the radial distances, and for the crystal we have considered three different spherical fragments with different definitions of the center. It is immediately seen that for some ranges of N (e.g., N around 75), the structures resemble fragments of the fcc crystal, but also that for even larger sizes

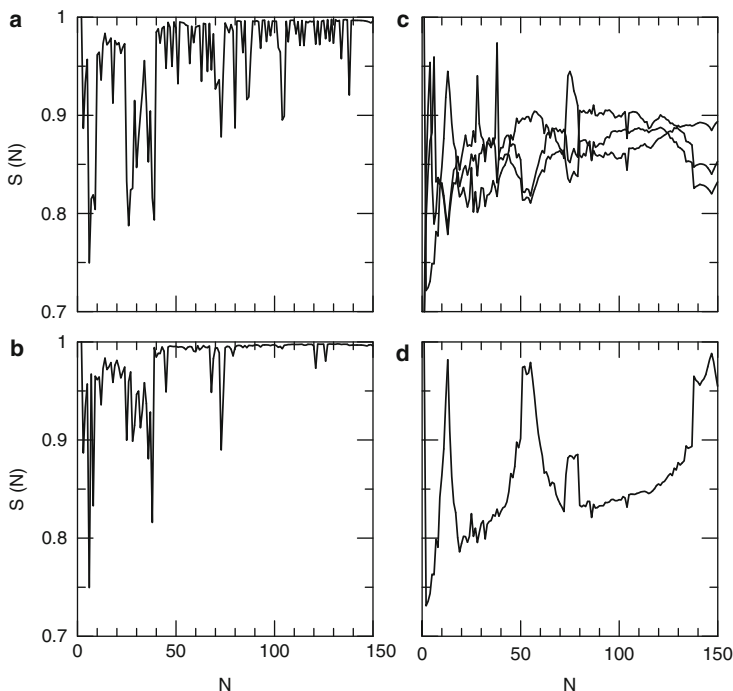


Fig. 8 The similarity functions for comparing the structure of the Ni_N cluster with (a) the energetically lowest cluster with $N - 1$ atoms, (b) the four energetically lowest clusters with $N - 1$ atoms, (c) a fragment of the fcc crystal, and (d) an icosahedron. For details, see the text

(around 145 atoms) as well as for $N \simeq 55$, there is considerably larger structural similarity with an icosahedron. Thus, even clusters with around 150 atoms are too small to approach the macroscopic crystalline limit.

Thermodynamic Properties

From the optimized structures as discussed in the preceding subsection, we may proceed to calculate thermodynamic properties as a function of temperature for at least not too elevated temperatures. Using a harmonic approximation, we expand the total energy as a function of nuclear coordinates up to second order,

$$\begin{aligned}
 E(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N) &\simeq E(\mathbf{R}_1^e, \mathbf{R}_2^e, \dots, \mathbf{R}_N^e) \\
 &+ \frac{1}{2} \sum_{k_1, k_2=1}^N \sum_{\alpha_1, \alpha_2=x, y, z} \frac{\partial^2 E}{\partial R_{k_1, \alpha_1} \partial R_{k_2, \alpha_2}} (R_{k_1, \alpha_1} - R_{k_1, \alpha_1}^e) (R_{k_2, \alpha_2} - R_{k_2, \alpha_2}^e),
 \end{aligned}
 \tag{25}$$

where the upper indices e mark the values at the total-energy minimum. This approximation can subsequently be used in calculating the N_v vibrational frequencies of the cluster of interest, whereby $N_v = 3N - 6$ except for linear systems where $i N_v = 3N - 5$. Subsequently, we can calculate the vibrational contributions to the thermodynamic properties, i.e.,

$$\begin{aligned}
 Z &= \sum_n e^{-E_n/k_B T} = e^{-E_0/k_B T} \prod_{i=1}^{N_v} [2 \sinh(\alpha_i/2)]^{-1} \\
 F &= -k_B T \ln Z = E_0 + k_B T \sum_{i=1}^{N_v} [\alpha_i/2 + \ln(1 - e^{-\alpha_i})] \\
 S &= -\partial F/\partial T = k_B \sum_{i=1}^{N_v} [\alpha_i/(e^{\alpha_i} - 1) - \ln(1 - e^{-\alpha_i})] \\
 E &= F + TS = (E_0 + \sum_{i=1}^{N_v} \hbar\omega_i/2) + \sum_{i=1}^{N_v} \hbar\omega_i/(e^{\alpha_i} - 1) \\
 C &= \partial E/\partial(k_B T) = \sum_{i=1}^{N_v} [\alpha_i^2 e^{\alpha_i}/(e^{\alpha_i} - 1)^2]. \tag{26}
 \end{aligned}$$

In these equations, Z , F , S , E , and C are the vibrational partition function, Helmholtz free enthalpy, entropy, internal energy, and heat capacity, respectively. Moreover, k_B is the Boltzmann constant and E_0 is the optimized total energy of the cluster at $T = 0$. Different $E_n = E_0 + \sum_{i=1}^{N_v} \hbar\omega_i(n_i + 1/2)$ with $n = \{n_1, n_2, \dots, n_{N_v}\}$ are the different vibrational energy levels of the cluster. Moreover, $\alpha_i = \hbar\omega_i/k_B T$. The formulas above include the zero-point contributions. Finally, the heat capacity is given in units of k_B . At sufficiently elevated temperatures, $k_B T \gg \hbar\omega_i$, C approaches the classical value of N_v , whereas for very low temperatures, $k_B T \ll \hbar\omega_i$, C approaches $(\hbar\omega_1/k_B T)^2 \exp(-\hbar\omega_1/k_B T)$ which differs markedly from the bulk T^3 (Debye) temperature dependence. Thus, with decreasing temperature, the vibrational heat capacity in clusters approaches zero faster than in bulk.

We used this approach for pure nickel and pure copper clusters (Grigoryan and Springborg 2010). From the obtained vibrational frequencies, we observed an interest aspect: for almost all cluster sizes, the highest vibrational frequency was higher than that of the corresponding crystalline system with the most important exception being for clusters for which the structure resembles a fragment of the crystal.

Figure 9 shows the resulting average heat capacity per vibrational mode, $C_{VM} = C/N_v$ for different temperatures. For the lowest temperatures, we see a very strong dependence on cluster size. Thus, Ni clusters with, for instance, 33, 37, and, less pronounced, 83 and 86 atoms and Cu clusters with $N = 16, 30, 44, 74, 95$, and 111 (clusters with either low-symmetrical structures with incomplete geometrical shells

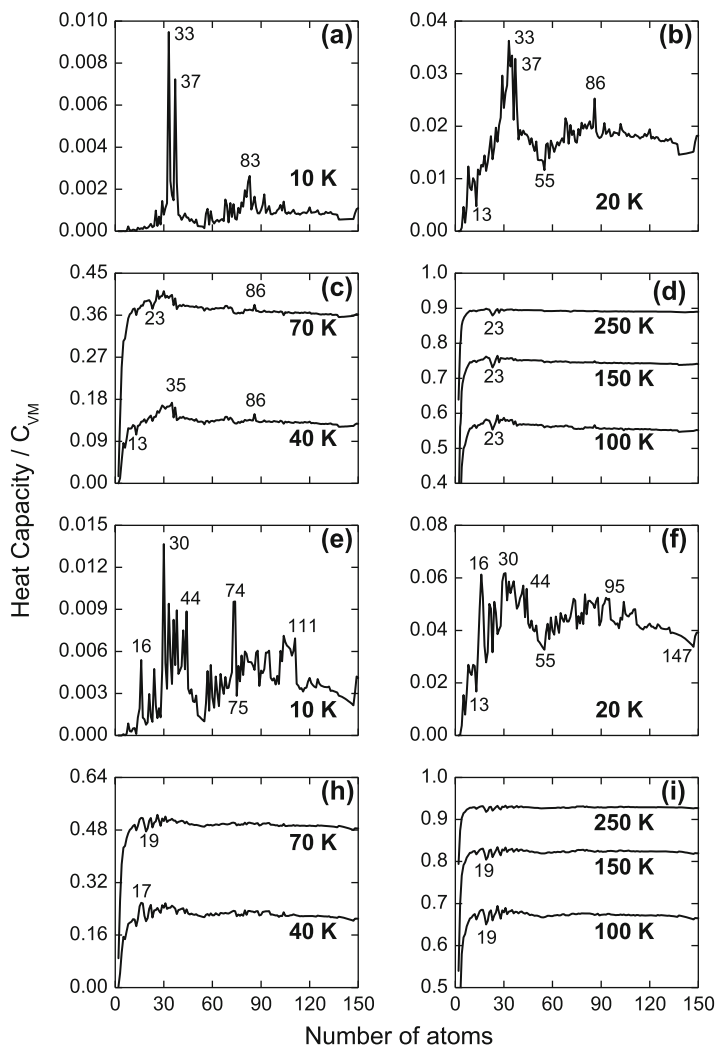


Fig. 9 Heat capacity of Ni [(a)–(d)] and Cu [(e)–(f)] clusters as a function of cluster size for different temperatures: (a) and (e) $T = 10$ K, (b) and (f) $T = 20$ K, (c) $T = 40$ and 70 K, and (d, i) $T = 100, 150,$ and 250 K

or decahedral/fcc structures) have large values for the vibrational heat capacities, whereas the particularly stable, i.e., magic-numbered, icosahedral clusters with $N = 13, 23, 55,$ and 147 show the most pronounced minima at low temperatures. As Fig. 9 suggests, the size dependence of C_{VM} becomes less pronounced with increasing temperature with clusters with less than roughly 30 atoms forming an exception. On the average, the heat capacities for copper clusters show stronger oscillations and larger values than those for Ni clusters.

Since we in the structure optimization calculations determined not only the energetically lowest structure but also the second and third lowest ones, we can use the vibrational spectra of those to study, for instance, a solid–solid transition temperature within a two-minima model. Two minima, “1” and “2,” with $E_0^{(1)} < E_0^{(2)}$ can represent two competing structural types as a decahedron and an icosahedron or a truncated octahedron and an icosahedron. According to the superposition approximation (Wales 2003), at the transition temperature, $T_{ss}(1 \leftrightarrow 2)$, the two probabilities P_1 and P_2 to find cluster structures “1” and “2” are identical. The corresponding equation for the transition temperature of the structural change from “1” to “2” is given by $P_1[T_{ss}(1 \leftrightarrow 2)] = P_2[T_{ss}(1 \leftrightarrow 2)]$ or $n_1 Z_1[T_{ss}(1 \leftrightarrow 2)] = n_2 Z_2[T_{ss}(1 \leftrightarrow 2)]$. Here, n_j is a degeneracy factor that takes into account the identical contributions from all the $2N!/o_j$ permutation–inversion isomers of structure j with o_j being the order of the point group, which is, for example, 114 for an icosahedron (I_h), 48 for a fcc structure (O_h), and 20 for a decahedron (D_{5h}).

Using this expression and the lowest-energy structures of Ni and Cu clusters together with their vibrational spectra, we have studied the structural changes from fcc to icosahedral structure in 38-atomic Ni and Cu clusters as well as from decahedral (Ni_{79}) to fcc (Ni_{79}) and from fcc (Cu_{79}) to decahedral (Cu_{79}) structures. As transition temperatures, we found 482 K for Ni_{38} , 587 K for Cu_{38} , and 87 K for Cu_{79} . Furthermore, we found that, despite the small total-energy difference of 0.04 eV between the two lowest isomers of Ni_{79} , the structural transition between them will not take place up to temperatures of at least 800 K.

Bimetallic Clusters

For binary clusters (i.e., clusters with two types of atoms), the additional existence of so-called homotops (Lloyd et al. 2004; Jellinek and Krissinel 1996) increases the computational demands enormously. Homotops for, e.g., $A_p B_q$ clusters, are defined as clusters with the same size, composition, and geometric arrangement but differing in the way in which A- and B-type atoms are arranged. Their number for the $A_p B_q$ cluster equals $\frac{(p+q)!}{p!q!}$.

As a first example, we shall here discuss the results of our study on bimetallic $\text{Ni}_n \text{Cu}_m$ clusters (Hristova et al. 2008a). In this case we used the basin-hopping method for optimizing the structure in combination with the EAM for the calculation of the total energy of a given structure. The EAM (as many other methods that do not include electronic/orbital degrees of freedom explicitly) has a preference for closed-packed structures so that in many cases the EAM method predicts very similar structures for pure clusters with just one element. Therefore, often only one or two types of structures were found for different $\text{Ni}_n \text{Cu}_m$ clusters with the same total number of atoms, $N = n + m$. The less obvious issue is accordingly how the two types of atoms are distributed within the cluster and if more than one overall structure exists, which structure is found for which stoichiometry.

For the $\text{Ni}_n \text{Cu}_m$ clusters, we found only three different types of scenarios. Either all clusters for a given $N = n + m$ have the same structure, which is the case

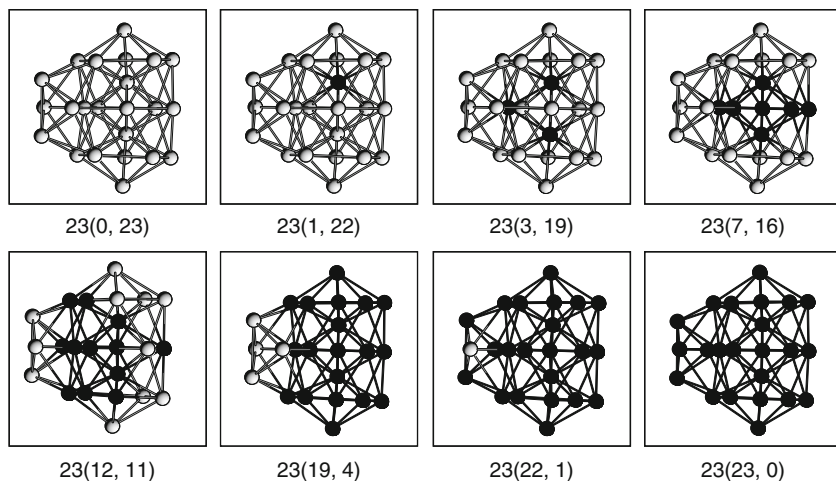


Fig. 10 The structures of Ni_nCu_m clusters for a fixed value of $N = n + m = 23$ and different values of (n, m) . The dark atoms mark the Ni atoms

for $N = 13$ and 23 , for instance. Figure 10 shows the results for $N = 23$. For $N = 38$, on the other hand, the two pure clusters have the same structure, whereas for $n = 5 - 25$, another structure is found as shown in Fig. 11. Finally, for $N = 15$ the two pure clusters have different structures, and the transition from the structure of the pure Cu cluster to that of the pure Ni cluster takes place for just 2 Ni atoms (cf. Fig. 12). In all cases, the Ni atoms tend to occupy inner positions, whereas the Cu atoms more often are found on the surface. This difference can be understood when comparing the cohesive energy of the two metals: the larger value for Ni suggests that this type of atom has a stronger preference for high coordination than Cu has.

In a very exhaustive study (Molayem et al. 2011a, b), we considered all Ni_mAg_n and Cu_mAg_n nanoalloys with $n + m$ up to 60. We shall here briefly discuss some few of the key results.

For the monometallic clusters like the Ni_n clusters of the preceding subsections, we identified particularly stable clusters (i.e., magic clusters) by comparing the total energy of the Ni_n cluster with those of the Ni_{n-1} and Ni_{n+1} clusters. For the nanoalloys, it is less obvious how a similar definition can be used in identifying particularly stable stoichiometries. Thus, we may, e.g., compare the total energy of the Ni_mAg_n system with those of the $\text{Ni}_{m-1}\text{Ag}_n$ and $\text{Ni}_{m+1}\text{Ag}_n$ systems, or with those of the $\text{Ni}_m\text{Ag}_{n-1}$ and $\text{Ni}_m\text{Ag}_{n+1}$ systems, or with those of the $\text{Ni}_{m-1}\text{Ag}_{n+1}$ and $\text{Ni}_{m+1}\text{Ag}_{n-1}$ systems, or with those of the $\text{Ni}_{m-1}\text{Ag}_{n-1}$ and $\text{Ni}_{m+1}\text{Ag}_{n+1}$ systems to mention just some possibilities.

The results for different definitions of the stability function related to the discussion above demonstrated, both for Ni_mAg_n and for Cu_mAg_n clusters, that with very few exceptions the different definitions would identify different clusters as

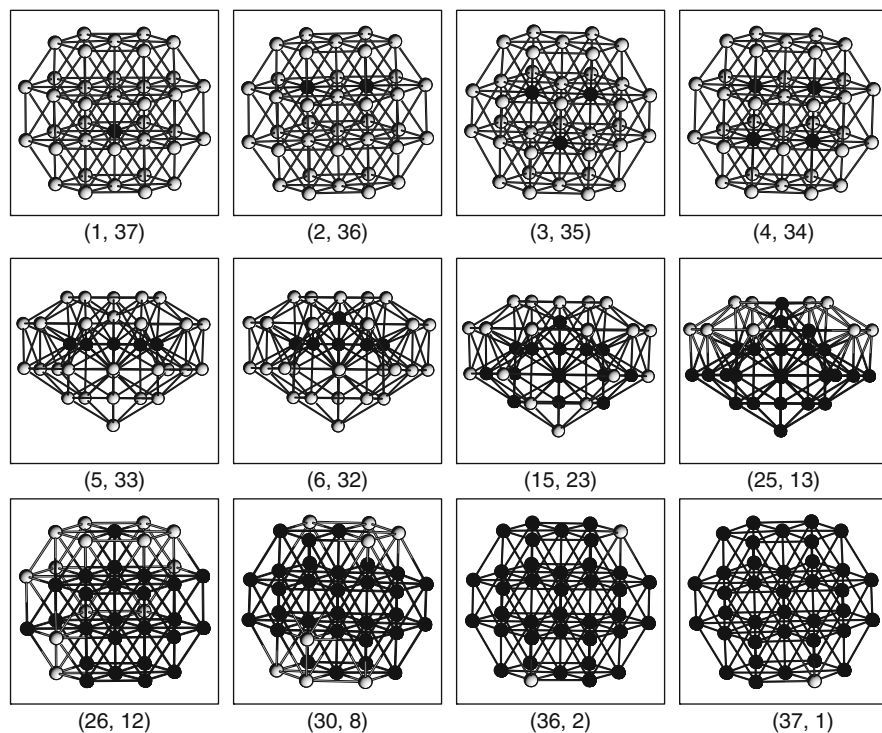


Fig. 11 The structures of Ni_nCu_m clusters for a fixed value of $N = n + m = 38$ and different values of (n, m) . The dark atoms mark the Ni atoms

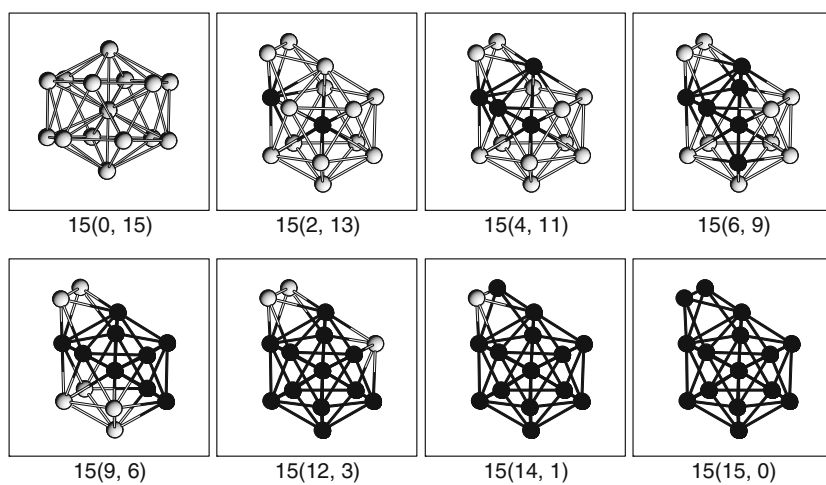


Fig. 12 The structures of Ni_nCu_m clusters for a fixed value of $N = n + m = 15$ and different values of (n, m) . The dark atoms mark the Ni atoms

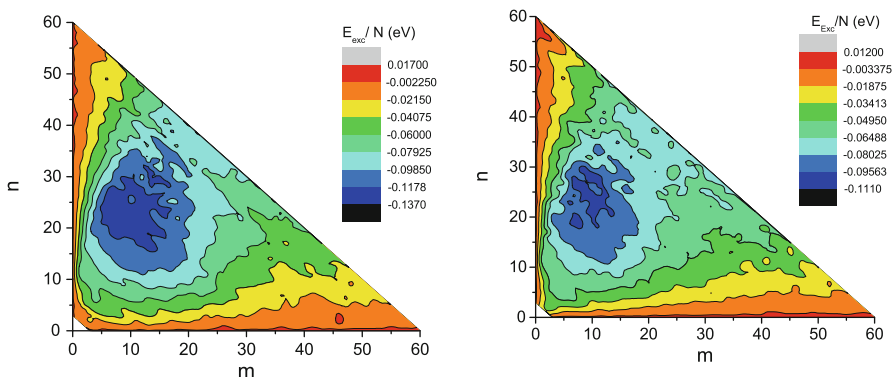


Fig. 13 The excess energy per atom for (left part) Ni_mAg_n and (right part) Cu_mAg_n clusters as a function of (m, n) for $N = m + n$ from 2 to 60

being particularly stable. Alternatively, we may consider an excess energy defined through

$$E_{\text{exc}} = E(m, n) - m \frac{E(m+n, 0)}{m+n} - n \frac{E(0, m+n)}{m+n} \quad (27)$$

Here, $E(m, n)$ is the energy for the Ni_mAg_n or Cu_mAg_n clusters. However, also this definition suffered from being biased: if the pure clusters with $m + n$ atoms are particularly stable (unstable), the excess energy would be particularly high (low). Thus, we have to conclude that it is hardly possible to identify uniquely particularly stable stoichiometries.

In Fig. 13 we show E_{exc}/N with $N = m + n$ for the two types of clusters of this study. As can be seen in the figure, E_{exc}/N is negative for almost all sizes and stoichiometries, which indicates that mixing is favorable. It is moreover interesting to notice that there is a certain size range, i.e., $m \simeq 10$ and $n \simeq 22$ for the Ni_mAg_n clusters and m between $m \simeq 6$ – 15 and n between $n \simeq 17$ – 30 for the Cu_mAg_n clusters, for which E_{exc}/N is particularly negative, suggesting that these clusters are the most stable ones. Another possible explanation, i.e., that the corresponding pure Ag, Cu, and Ni clusters are particularly unstable, does not hold here. A comparison between the two systems in Fig. 13 demonstrates that the energetic properties of the Ni_mAg_n or Cu_mAg_n clusters differ somewhat. This is much less the case for the structural properties, and therefore, we shall discuss the latter for only the Ni_mAg_n clusters.

The fact that mixing in general is preferred can be studied further via structural information. Thus, mixing vs. segregation of the nanoalloys can be studied through the bond order parameter σ (Pittaway et al. 2009). For an A_mB_n nanoalloy, it is defined as

$$\sigma = \frac{N_{A-A} + N_{B-B} - N_{A-B}}{N_{A-A} + N_{B-B} + N_{A-B}} \quad (28)$$

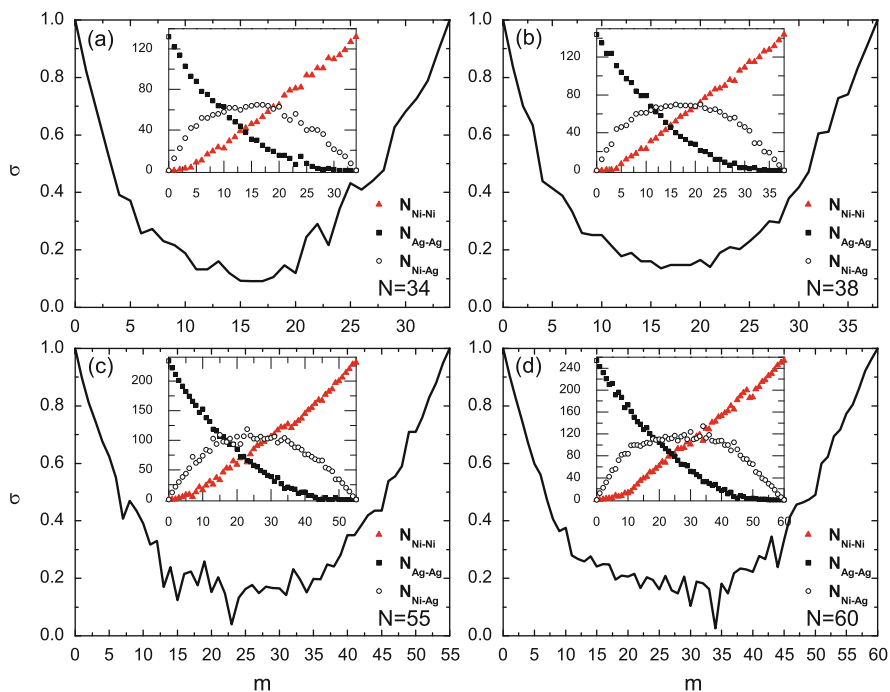


Fig. 14 Bond order parameter as a function of composition (number of Ni atoms, m) for the global minima of four sizes of interest ($N = 34, 38, 55,$ and 60). The inserts show the number of the three possible types of bonds versus m . *Solid triangles* and *squares* refer to the numbers of Ni–Ni and Ag–Ag bonds, respectively, whereas *open circles* are for Ni–Ag bonds

with N_{i-j} ($i, j = A, B$) being the number of nearest-neighbor bonds between atoms of type i and j . σ is positive for segregated, almost zero for disorderly mixed, and negative for mixed and onion-like phases of nanoalloys.

Figure 14 shows the bond order parameter versus number of Ni atoms for all the compositions of four selected sizes, $N = 34, 38, 55,$ and 60 . In addition, in each case, we also show the corresponding number of Ni–Ni, Ag–Ag, and Ni–Ag bonds. Having only positive values of σ implies some degree of segregation, and as we shall argue below, this segregation is mainly due to the formation of core–shell-like structures. For the clusters with comparable numbers of Ni and Ag atoms, there is a relatively large number of Ni–Ag bonds so that σ obtains lower values. Thus, the lowest value of σ for $N = 34$ is found for the (16, 18) and (17, 17) clusters. On the other hand, for $N = 38, 55,$ and 60 , the lowest values of the bond order parameter are found for more asymmetric clusters, i.e., $(m, n) = (16, 22), (23, 32),$ and $(34, 26)$, respectively.

A possible spatial separation of the Ni and Ag atoms can be (partly) identified through the radial distances of Eq. (23) when considering these individually for the two types of atoms. We show in Fig. 15 the ratio of average radial distances of Ni and

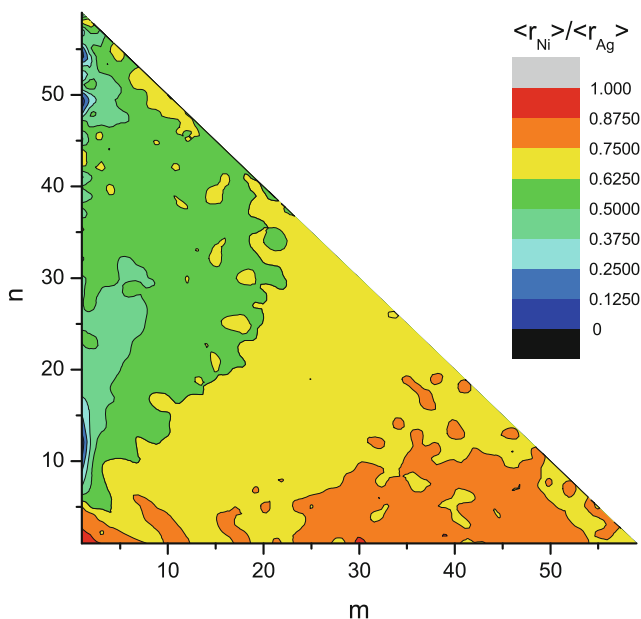


Fig. 15 The ratio of the average radial distances of the Ni atoms to that of the Ag atoms in $Ni_m Ag_n$ clusters as a function of (m, n) for $N = m + n$ from 2 to 60

Ag atoms for all of the considered clusters. A segregation of the Ag atoms toward sites with larger distances from the center of the clusters is clearly recognized as the ratio of distances is mostly less than 1 and is only for Ni-rich clusters close to 1. This suggests a partial formation of core-shell structures, although the segregation is not complete.

Clusters on Surfaces

As mentioned in the introduction, section “Introduction,” experimental studies of clusters often consider clusters on some surface. Both clusters that have been deposited on the surfaces and clusters that are grown directly on the surfaces have been studied. Since theoretical studies often consider clusters in the gas phase, a highly relevant question is whether the latter have anything to do with clusters on a surface. In some few studies, we have addressed this issue.

Using molecular-dynamics simulations and the EAM, we studied the deposition of the particularly stable Ni_{13} and Cu_{13} clusters (cf. section “Ni Clusters”) on the (111) surfaces of crystalline Ni and Cu (Kasabova et al. 2007) as well as more different Cu clusters deposited on the Cu(111) surface (Alamanova et al. 2008). The deposition energies, 0–1 eV/atom, imply according to experimental classifications that the depositions are to be considered as being soft. Nevertheless, the depositions

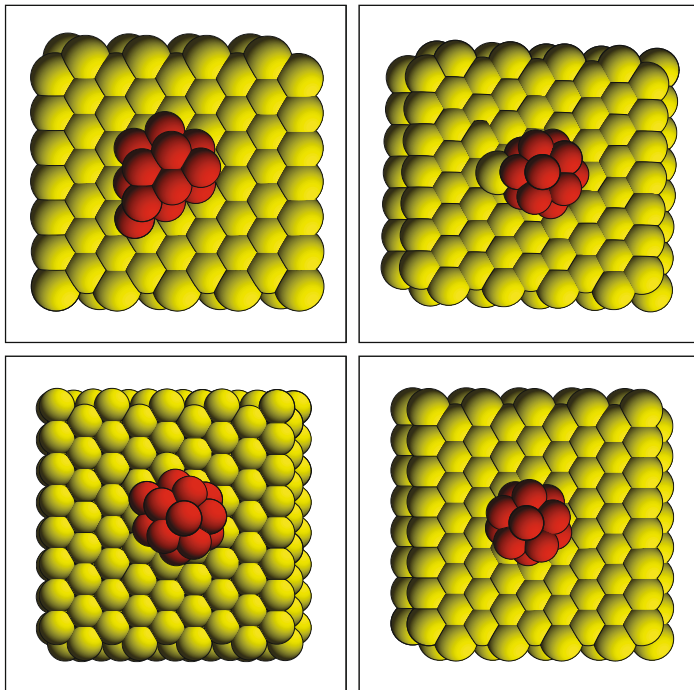


Fig. 16 Top view of the final structures obtained when depositing A_{13} clusters on the B(111) surface with 0 eV/atom deposition energy. A and B are (top, left) Ni and Cu, (top, right) Cu and Ni, (bottom, left) Cu and Cu, and (bottom, right) Ni and Ni. The dark atoms are those of the clusters. The deposition energy equals in all cases 0 eV/atom

even at the smallest possible deposition energies (cf. Fig. 16) led to clusters that were highly deformed compared to the gas-phase structures. To some extent, Ni clusters deposited on the Cu surface were partly intact, which may be understood through the higher cohesive energy and bulk modulus of Ni compared with Cu. Thus, in this respect, Ni and Cu can be classified as being hard and soft, respectively.

Alternatively, one may suggest that when clusters and surfaces are made of the same metal, the cluster atoms will lead to an epitaxial growth of the substrate. To quantify this suggestion, we introduced a so-called index of epitaxy (Alamanova et al. 2008), I , through

$$q = \sum_i^N |\mathbf{R}_i - \mathbf{R}_c|^2$$

$$I = \frac{1}{1 + q/u_i^2}, \quad (29)$$

where $|\mathbf{R}_i - \mathbf{R}_c|$ is the distance between the positions of the i th atom and the closest-lying fictitious atom in the infinite, periodic crystal that is obtained by continuing

the structure of the substrate periodically. $I = 1$ if perfect epitaxy is obtained. Our calculations indicated, however, very little epitaxy.

In another study, we consider the structures that result when a cluster is grown directly on the substrate (Hristova et al. 2009b). We considered Ag_N clusters grown on Ag(111) and Ni(111) surfaces. Since the (111) surfaces are closed packed and since the EAM (as mentioned above) has a strong (and, for the present systems, realistic) preference for structures with high-coordinated atoms, the growth of Ag clusters on the Ag(111) surface led to planar structures that moreover were roughly spherical. Thereby, the highest coordination for the cluster atoms could be obtained, and since substrate and cluster are of the same material, there is no strain involved.

This is different for Ag clusters on the Ni(111) surface. In this case, the lattice mismatch between Ag and Ni leads to some strain in the clusters grown epitaxially on the surface, and accordingly, the Ag clusters tend to get a somewhat elongated structure.

Only Electrons

The approaches we have discussed in the preceding section were based on the rationale that for metal clusters, there is no directional bonding, and instead, the electrons are delocalized over the complete system. Thereby, particularly stable clusters (i.e., magic numbers) occur for clusters for which the atoms can form a particularly compact, closed-packed geometrical structure. The electronic effects were at most indirectly included so that, e.g., stability due to electronic shell effects is not taken into account.

This is, on the other hand, the main effect that is included in the jellium model. Within the jellium model, it is assumed that the system of interest forms a compact, closed-packed structure and that the valence electrons are delocalized over the complete system. Due to the delocalization of the electrons, they do hardly feel the precise arrangement of the nuclei and core electrons so that the effects of those can be approximated as being those of a homogeneous charge density with the same average density as for the true system. The valence electrons move then in the electrostatic potential created by this jellium.

The jellium model has been used intensively for the studies of clusters of metals; see, e.g., de Heer (1993) and Brack (1993). In fact, one of the first papers in more recent times that treated the special properties of clusters was a combined experimental and theoretical study on Na_N clusters (Knight et al. 1984). Sodium can be considered the prototype of metal clusters for which the unidirectional bonding due to the $3s$ valence electrons suggests that the arguments above provide a good starting point for rationalizing the properties of Na clusters. Thus, packing effects that can be modeled accurately with a simple potential that does not include electronic orbitals will be responsible for the (essentially spherical) structures, and the particularly stable structures can be identified with a jellium model for delocalized electrons. In their study, Knight et al. assumed that the jellium was spherical and calculated, subsequently, the orbitals for the electrons moving in

this jellium and under the influence of each other. Comparing with experimental mass abundance spectra, they could explain the experimental observation of certain cluster sizes that showed up particularly often as related to these clusters' high stability. The latter was, in turn, due to the closing of shells of electronic orbitals, equivalent to what explains the inertness of the inert gases He, Ne, etc. in the periodic table.

In some more recent studies, we have used the spherical jellium model to analyze the properties of Cs_N clusters (Springborg 1999) and of a C_{60} molecule coated with alkali metals (Springborg et al. 1996). We shall here, however, not discuss these studies further but below address the accuracy of this jellium model.

Both Nuclei and Electrons

Methods

When including both electrons and nuclei in the calculations, the calculations may become so demanding that the number of structures and/or the size of the systems has to be limited. It is then often very useful to augment calculations based on less accurate methods with more accurate ones on selected structures in order to access the accuracy of the approximate methods. In some case we have done so, too, using density-functional methods (see, e.g., Springborg 2000 for a description of such methods).

As a compromise between accuracy and computational demands, we have often used the so-called DFTB (density-functional tight-binding) method of Seifert and coworkers (Seifert and Schmidt 1992; Porezag et al. 1995; Seifert et al. 1996). The DFTB method is based on the density-functional theory (DFT) of Hohenberg and Kohn (1964) in the formulation of Kohn and Sham (1965). For more details, we refer the reader to the original works or to Springborg (2000). Within the DFTB approach, the total energy relative to the noninteracting atoms is given as

$$E_{\text{tot}} \simeq \sum_i^{\text{occ}} \epsilon_i - \sum_j \sum_m \epsilon_{jm} + \frac{1}{2} \sum_{j \neq k} U_{jk}(|\mathbf{R}_j - \mathbf{R}_k|), \quad (30)$$

where ϵ_i is the energy of the i th orbital for the system of interest and ϵ_{jm} is the energy of the j th orbital for the isolated m th atom. Moreover, U_{jk} is a short-range pair potential between atoms j and k that is adjusted so that results from parameter-free density-functional calculations on two-atomic systems as a function of the interatomic distance are accurately reproduced. Finally, only the valence electrons are explicitly included in the calculations, whereas the other orbitals are treated within a frozen-core approximation.

The elements of the Hamilton and overlap matrices, i.e., $\langle \chi_{m_1 n_1} | \hat{H} | \chi_{m_2 n_2} \rangle$ and $\langle \chi_{m_1 n_1} | \chi_{m_2 n_2} \rangle$ with χ_{mn} being the n th atomic orbital of the m th atom, are obtained

from calculations on diatomic molecules. The Hamilton operator contains the kinetic energy operator as well as the potential. The latter is approximated as a superposition of the potentials of the isolated atoms,

$$V(\mathbf{r}) = \sum_m V_m(|\mathbf{r} - \mathbf{R}_m|), \quad (31)$$

and we assume that the matrix element $\langle \chi_{m_1 n_1} | V_m | \chi_{m_2 n_2} \rangle$ vanishes unless at least one of the atoms m_1 and m_2 equals m .

Thus, within this approach, all information that enters the calculations are extracted from the properties of diatomic molecules. These can, in turn, be determined from accurate, density-functional calculations. In order to access the accuracy of the approach, we perform additional calculations on the infinite, periodic systems. If these provide accurate results, we assume that also for the intermediate, finite clusters, the results are accurate.

For an unbiased structure determination, we combine the DFTB method with genetic algorithms. The genetic algorithms are based on the principles of natural evolution and are, therefore, also called evolutionary algorithms (Holland 1975; Goldberg 1989) and have been found to provide an efficient tool for global geometry optimizations. One version of the genetic algorithms that we are using is as follows.

A population of P initial structures is chosen randomly (these clusters are called parents), and each structure is relaxed to its nearest total-energy minimum. By cutting each parent randomly into two parts, a next set of P structures is obtained by interchanging (“mating”) these two parts and allowing the resulting “children” to relax, too. Comparing the energies of the $2P$ clusters of both sets, those P with the lowest total energies are chosen to form the set of parents for the next generation. This procedure is repeated for many hundred generations until the lowest total energy is unchanged for a large number of generations.

Alternatively, this “two-parent” method may be replaced by a “single-parent” method where the two parts of one cluster simply are interchanged. The advantage of this latter approach is that for systems with more than one type of atoms, stoichiometry is automatically conserved. Moreover, in some cases, we found that the latter is faster than the former.

Na and Au Clusters

As discussed in section “Only Electrons,” Na clusters may be considered as being the prototype for which the jellium model provides a good starting point. Therefore, it seems to be highly appropriate to start our discussion with considering this system but applying other theoretical approaches. Accordingly, we have studied the properties of Na_N clusters both using the EAM method of section “Only Nuclei” and the DFTB method. In both cases, the total-energy methods were combined with unbiased structure optimization methods (Tevekelyska et al. 2010).

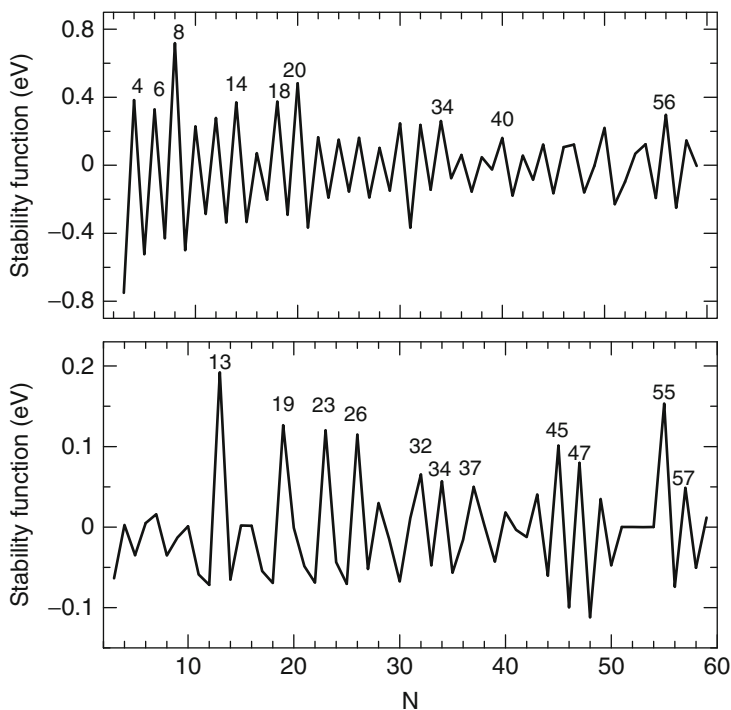


Fig. 17 Stability function (in eV) as function of the size of the clusters from the DFTB (*upper part*) and the EAM (*lower part*) calculations

In Fig. 17 we show the resulting stability function and in Fig. 18 the radial distances of the atoms. The stability functions show only a marginal similarity, and only the DFTB calculations find particularly stably clusters for sizes that agree with those of the mass abundance experiments and the jellium calculations (2, 8, 18, 20, 34, 40, 58, ...) (Knight et al. 1984). On the other hand, the radial distances indicate that the structures found in the EAM calculations are somewhat more symmetric than are those of the DFTB calculations. In fact, based on a shape analysis (not shown), we find that none of the DFTB structures can be considered as being roughly spherical, in marked contrast to the inherent assumptions behind the jellium calculations.

Although gold often is considered as being a normal metal, it has remarkable bonding properties that manifest themselves clearly in nanosystems (see, e.g., Springborg 2009). Thus, gold is often found in situations with low coordinations. This is attributed to effects related to the electronic orbitals (although the precise role of various effects like *sd* hybridization, relativistic effects, the range of the atomic potential, etc. has not yet been clarified unambiguously) that are absent in the EAM. Therefore, that EAM calculations for gold clusters (Alamanova et al. 2006) led to too compact structures may not surprise.

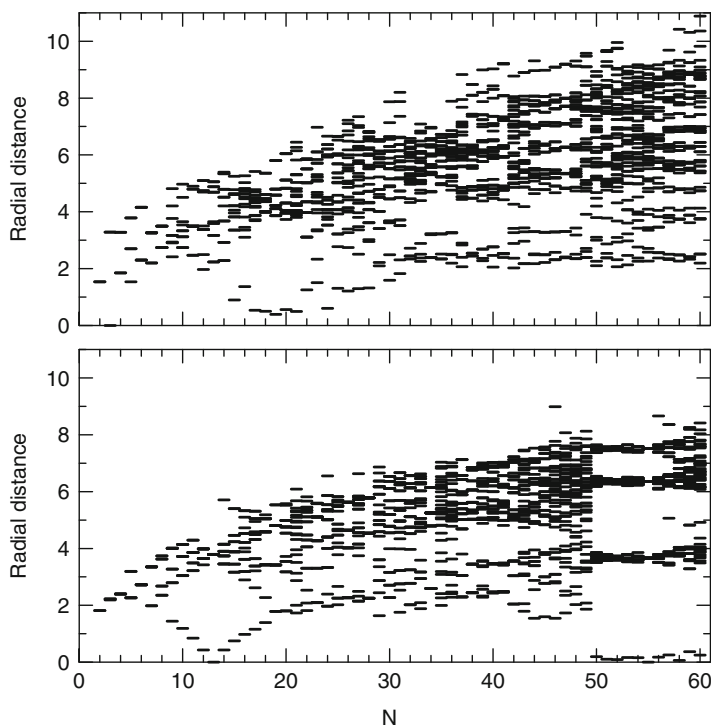


Fig. 18 Radial distances (in Å) for sodium clusters obtained with DFTB (*lower part*) and EAM (*upper part*) as a function of the number of atoms N . In each panel, a *small horizontal line* for a given N marks that at least one atom has that distance to the center of the cluster

Instead, we used the DFTB method in combination with the genetic algorithms to study the properties of Au_N clusters with N up to 58 (Dong and Springborg 2007). The main findings are reproduced in Fig. 19. The total energy per atom is seen to fall steeply until around 20 atoms after which it depends much weaker on the cluster size. It is known that small, neutral Au_N clusters are planar up to N somewhat above 10 (see, e.g., the discussion in Dong and Springborg 2007), although this is not found in the present calculations. On the other hand, as seen in the figure, the difference in the total energy between the planar and our nonplanar structures is small. This difference can be taken as an estimate of the accuracy of the present DFTB approach. Moreover, also the experimentally observed tetragonal Au_{20} structure is not found in our calculations, although it has a total energy that is very close to that of our optimized structure. On the other hand, as all other electronic-structure studies on Au_{13} , our calculations do not find the icosahedron to be the most stable structure.

From the radial distances of the Au_N clusters, we see that they all are low symmetric. This can also be seen when comparing the (high-symmetric) structures of Ni_N clusters with those of Au_N clusters with the same number of atoms, i.e.,

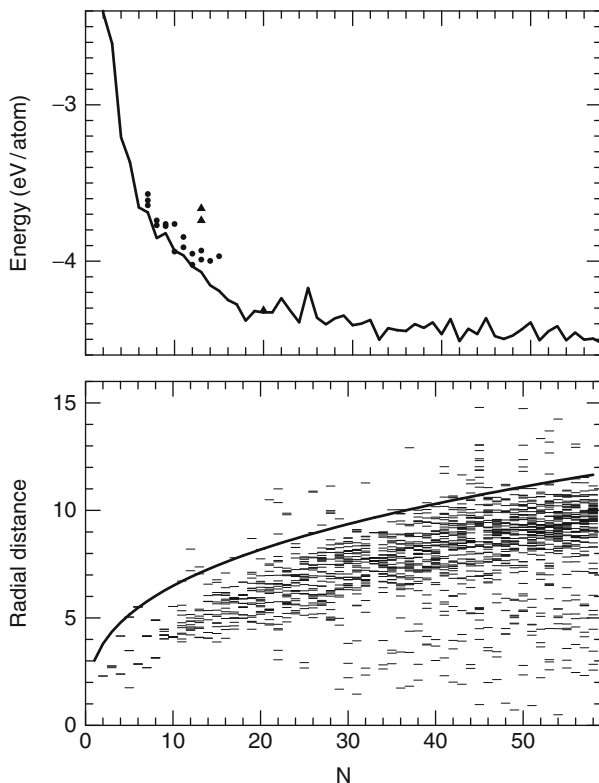


Fig. 19 The upper part shows the variation in the total energy per atom (relative to that of the isolated atom) for the optimized Au_N clusters (*solid curve*) together with those of planar structures (*dark circles*) and those of icosahedral Au_{13} and tetragonal Au_{20} clusters (*dark triangles*). The *lower part* shows the radial distances as a function of cluster size, i.e., each small line represents (at least) one atom with that radial distance. The curve shows the radius of the spherical jellium with a density as in the crystal

when comparing Figs. 4 and 20. Moreover, for $10 \leq N \leq 20$, no atom has a small value for its radial distance, implying that these structures are shell-like (see also Fig. 20). Indeed the finding of such “golden cages” is in agreement with other studies (Bulusu et al. 2006). Finally, we add that when analyzing the coordination of the individual atoms, we found that no atom has a coordination larger than 11, i.e., the value of 12 of crystalline fcc gold is not found even for the fairly large clusters of this study. This may be considered a further confirmation of the ability of gold atoms to exist in low-coordinated situations.

In a more recent study (Dong et al. 2013b), we augmented this study on the gold clusters with an analysis of the vibrational properties and the size and temperature dependence of the heat capacity. We used thereby an approach as given in Eqs. (25) and (26) with the only difference being in the way the total energy as a function

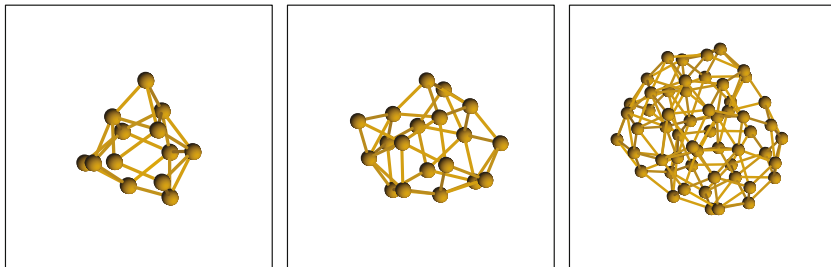


Fig. 20 The structures of the Au_N clusters with (from left to right) $N = 13, 19,$ and 55

of structure is calculated. For the second-order derivatives of the total energy with respect to structural coordinates, it turned out to be most accurate to use a finite-difference approximation

$$\begin{aligned}
 f_{ij} &= \frac{\partial}{\partial q_i} \frac{\partial E}{\partial q_j} = \frac{\partial}{\partial q_j} \frac{\partial E}{\partial q_i} = \frac{1}{2} \left(\frac{\partial}{\partial q_i} \frac{\partial E}{\partial q_j} + \frac{\partial}{\partial q_j} \frac{\partial E}{\partial q_i} \right) = \frac{-1}{2} \left(\frac{\partial F_i}{\partial q_j} + \frac{\partial F_j}{\partial q_i} \right) \\
 &\simeq \frac{-1}{4\Delta s} [F_i(q_j + \Delta s) - F_i(q_j - \Delta s) + F_j(q_i + \Delta s) - F_j(q_i - \Delta s)]
 \end{aligned}
 \tag{32}$$

where Δs is a small but finite change in the coordinates, and $F_m(q_n \pm \Delta s)$ denotes the m th force component for the structure where the n th coordinate has been changed by $\pm \Delta s$. From test calculations, we found that $\Delta s = 0.01$ a.u. results in numerically stable results. It turned out to be useful to introduce a characteristic temperature for each vibration through

$$\alpha_i = \frac{\hbar\omega_i}{k_B T} \equiv \frac{T_i}{T}.
 \tag{33}$$

For each mode, its contribution to the vibrational heat capacity is a monotonously increasing function of temperature, and the temperature $T \simeq 2.35T_i$ is the one at which the contribution of the given mode changes most rapidly as a function of temperature. At $T \simeq 2.35T_i$ the contribution of the i th mode equals roughly 64 % of its maximal contribution (at $T \rightarrow \infty$).

In Fig. 21 we show the heat capacity per atom of the Au_N clusters for temperatures of 70, 298.25, 500, 700, and 1200 K. We observe a strong size dependence of the vibrational heat capacity per atom in particular for the smallest clusters and lowest temperatures. That this is so can be understood from Eq. (26). Thus, for the highest temperatures, the vibrational heat capacity per atom approaches $(3 - \frac{6}{N})k_B$ for nonlinear clusters, which is smooth and monotonically increasing as a function of N and converges to $3k_B$ for $N \rightarrow \infty$. On the other hand, for a low temperature, T , only those clusters for which there exist one or more vibrational modes with characteristic temperatures $T_i < T$ will have significant heat capacities.

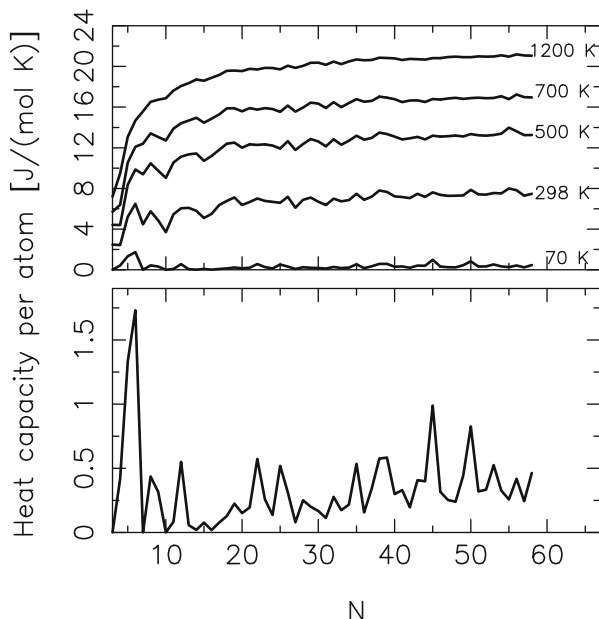
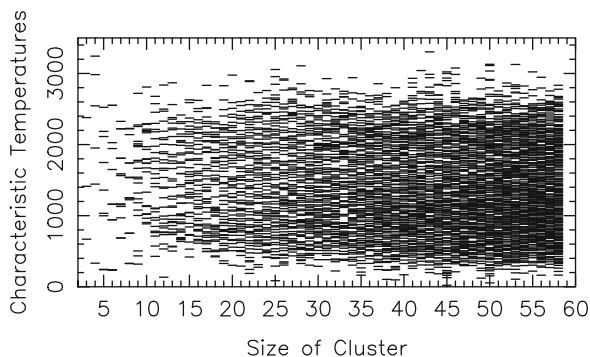


Fig. 21 The vibrational contribution to the heat capacity per atom for different temperatures (given in the *upper panel*) as function of the size of the clusters. The *lower panel* shows an expanded version of the results for 70 K

An interesting issue is whether a correlation between stability and heat capacity exists. One may speculate that clusters that are particularly stable (unstable) also are particularly rigid (soft), i.e., have particularly high (low) vibrational frequencies. That would imply that the heat capacity then has a minimum (maximum) for those clusters. However, by comparing the heat capacity per atom with the stability function (not shown), it is clear that there is no correlation between stability and heat capacity. Thus, the largest values for the heat capacity are found for $N = 5$ and 6, for which, however, the stability function does not show any pronounced minima. On the other hand, the most stable clusters are found for $N = 42$ and 51, for which the heat capacity has average values.

By plotting the characteristic temperatures as a function of cluster size (this is done in Fig. 22), we can identify those clusters that have a particularly large heat capacity at low temperatures as those that have modes with particularly low characteristic temperatures. As seen in the figure, this is the case for $N = 5, 12, 22, 25, 32, 35, 38, 39, 40, 45, 50, 51,$ and 53. The reason that some of the vibrational modes for certain cluster sizes have particularly low characteristic frequencies is often due to, e.g., wagging of some low-coordinated atoms. This is, for instance, seen for planar clusters.

Fig. 22 The characteristic temperatures for the different vibrational modes as function of the size, N , of the clusters. For a given N , each *line* marks that at least one mode has that characteristic temperature



HAIO Clusters

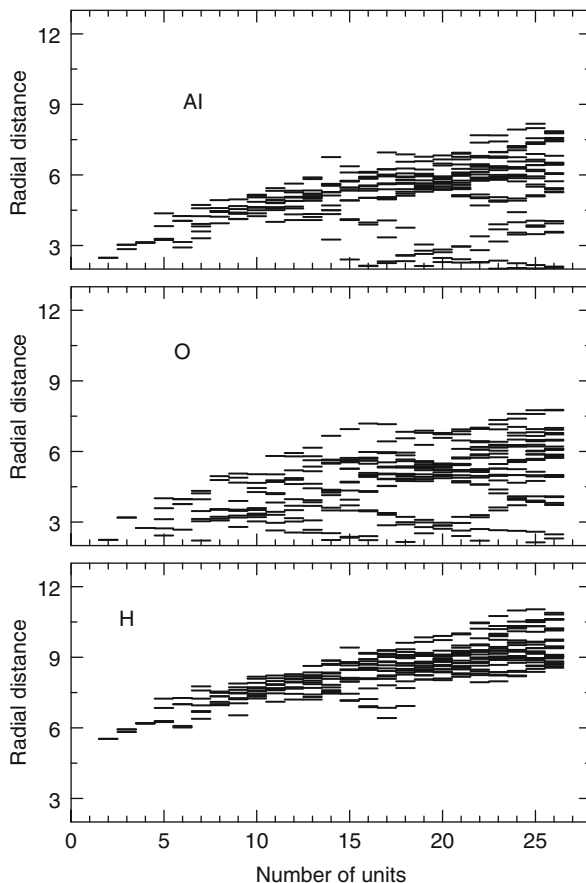
HAIO is a nanostructured material that can be used as substrate for deposition of other materials. However, little is known about the material itself, except that it is stoichiometric (i.e., it contains 1:1:1 mole fractions of H, Al, and O). From nuclear resonance and vibrational spectroscopy experiment, some information can be obtained about the short-range arrangements of the atoms.

In a collaboration with members of one of the groups that synthesize this material, we studied its properties using different theoretical methods (Dong et al. 2005). Genetic algorithms together with the DFTB method were used in an unbiased determination of isolated $(\text{HAIO})_n$ clusters with $n \leq 26$. Moreover, since the material is macroscopic and contains interacting nanostructures, we also studied the interactions between two of those optimized clusters using the DFTB approach. Finally, the predictions of the calculations were supported by parameter-free density-functional calculations on selected structures for clusters with $n \leq 6$.

At first, we show in Fig. 23 the radial distances for each atom type as a function of n . It is evident from this figure that the clusters consist of a core of aluminum and oxygen atoms covered by a shell of hydrogen atoms. A more careful inspection of the structures (not shown) reveals that for the Al–O core, only heteroatomic (i.e., Al–O) and not homoatomic (Al–Al or O–O) bonds exist.

The results of the parameter-free density-functional calculations are in support for these structural findings. We considered in total 50 isomers of $(\text{HAIO})_n$ with $n = 1, 2, 3, 4,$ and 6 . When analyzing the structures of those, we found that none possessed O–O or H–H nearest neighbors and that it was relatively easy to identify Al–H, Al–O, and O–H bonds as being pairs with an interatomic distance smaller than 2.2, 2.5, and 1.5 Å, respectively. The Al–Al interatomic distances, on the other hand, showed a large spread, and it was not possible readily to identify a cutoff distance below which the Al atoms could be considered as being bonded. In the subsequent analysis, we therefore considered two extreme values, 2.5 and 4.1 Å.

Fig. 23 The radial distances (in a.u.) for Al, O, and H atoms, separately, as a function of the size of the cluster n for $(\text{HAIO})_n$ clusters. In each panel, a *small horizontal line* shows that at least one atom of the corresponding type has that distance to the center of the cluster for a given value of n



Next we studied the total energy per unit as a function of number of A–B bonds per unit, with A and B being H, Al, and O. The results are shown in Fig. 24. The results are very scattered, but it is possible to identify certain trends. First, the total energy decreases as the number of Al–O bonds increases. Second, a similar, but much weaker, trend can be identified for the number of Al–Al bonds (here we have used 4.1 Å as our cutoff value, but 2.5 Å gives very similar results). Third, there is a clear preference for structures with one Al–H bond per unit. These observations can be quantified by approximating the total energy of the various $(\text{HAIO})_n$ isomers as

$$E_{\text{tot}} \simeq n \cdot E_0 + n_{\text{AlAl}} \cdot E_{\text{AlAl}} + n_{\text{AlO}} \cdot E_{\text{AlO}} + n_{\text{AlH}} \cdot E_{\text{AlH}} + n_{\text{OH}} \cdot E_{\text{OH}}. \quad (34)$$

A least-squares fit gave $E_{\text{AlAl}} = -0.175 \text{ eV}$ (-0.122 eV), $E_{\text{AlO}} = -1.000 \text{ eV}$ (-1.170 eV), $E_{\text{AlH}} = 0.750 \text{ eV}$ (0.678 eV), and $E_{\text{OH}} = 1.119 \text{ eV}$ (0.798 eV), when using 4.1 Å (2.5 Å) as cutoff distance for Al–Al bonds. These numbers show that Al–O bonds are strongly preferred and that when choosing between adding H to either Al or O, it is energetically preferred to create Al–H bonds.

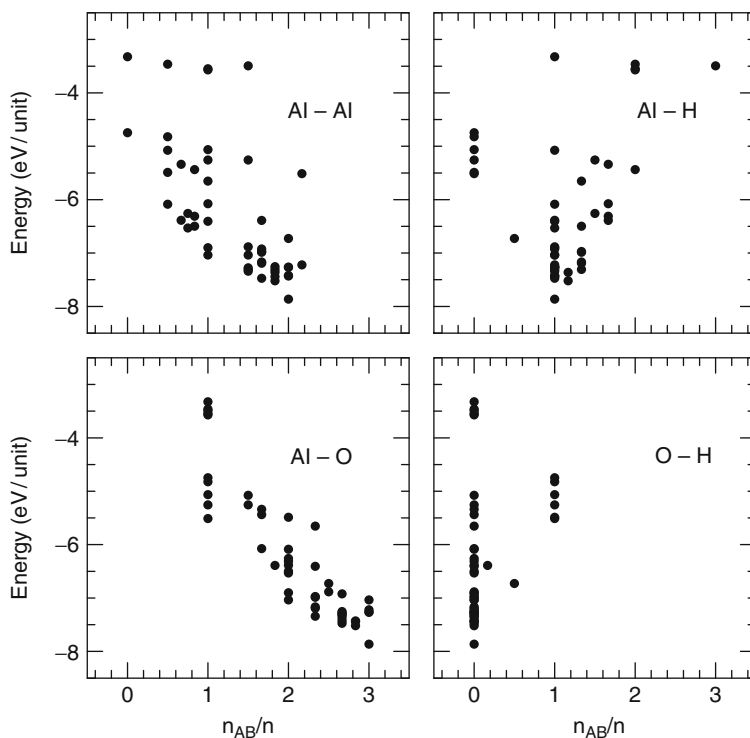


Fig. 24 Variation in the total energy per unit as function of the number of A–B bonds per unit. The results are from the parameter-free density-functional calculations for the $(\text{HAIO})_n$ clusters with $n = 1, 2, 3, 4, 6$, and the meaning of A and B is given in the panels

Our theoretical findings that the H atoms prefer to stay outside an AIO core, which for larger clusters becomes increasingly difficult when requiring that the material is stoichiometric, suggest that stable structures of HAIO may occur for layers of HAIO. In order to study this proposal further, we considered theoretically extended HAIO systems consisting of either one or two layers of HAIO. We add that these calculations ignore interlayer interactions that can be very important and that, therefore, may modify our conclusions significantly, when included. For a single layer of HAIO, one can imagine two highly symmetric cases, i.e., one where all H atoms are on the one side of the layer and one where every second H atom is above and every second H atom below the AIO layer. For the case that all H atoms are on the same side of the layer, the Al and O atoms form a layer with bond lengths of $3.33 \text{ a.u.} = 1.76 \text{ \AA}$ and $3.34 \text{ a.u.} = 1.77 \text{ \AA}$. Moreover, the hydrogen atoms are sitting on the outside of the layer bonded to the Al atoms with Al–H bond lengths of $3.21 \text{ a.u.} = 1.70 \text{ \AA}$. The Al–O–Al bond angles are $85^\circ\text{--}86^\circ$ and $147^\circ\text{--}152^\circ$. Remarkably different things occur when the H atoms are sitting alternating on the two sides of the single AIO layer. Then we found that the layer split into several small parts all with the same kind of structure, i.e., the H atoms are binding to the Al atoms with

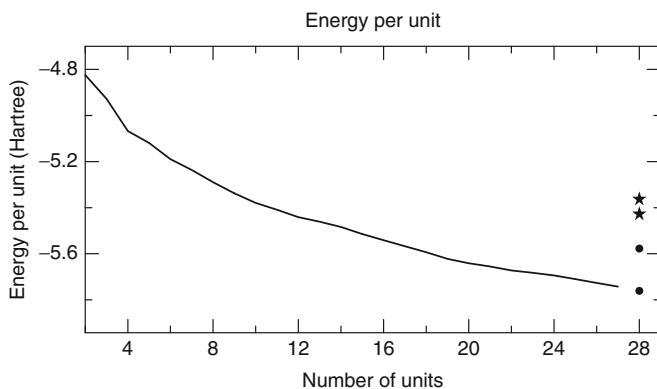


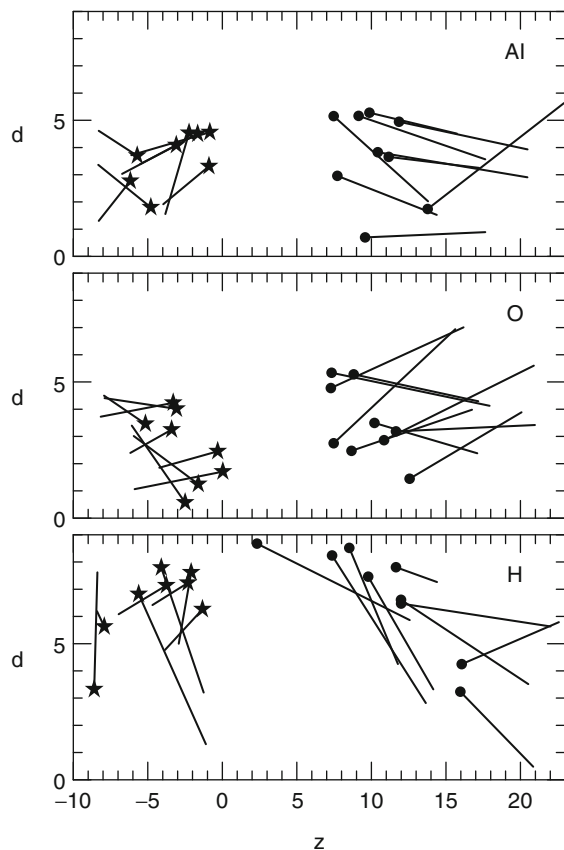
Fig. 25 Variation in energy per unit for isolated clusters and for the layers of HAIO. The results for the layers are shown to the right with the one-layer results marked with *stars* (here, the lowest total energy is for the case that the H atoms are on the same side, whereas they are alternating on the two different sides in the other case) and the two-layer results marked with *circles* (here, the lowest total energy is for the case that Al–O bonds form the bonds between the layers). The energy is given in Hartrees (1 Hartree \simeq 27.21 eV)

Al–H bond lengths around 3.08 a.u. = 1.63 Å, whereas the Al–O bond lengths are around 3.40 a.u. = 1.80 Å. For the case of two layers of HAIO, we studied two cases, i.e., either the Al atoms of one layer were placed on top of the O atoms of the other layer, or they were placed on top of the Al atoms of the other layer. It turned out that the first situation was much more stable than the second one, which may not surprise, and in the second case, we find that the system breaks into small parts.

Finally, in Fig. 25 we show the energy per unit for the finite clusters in comparison with that for the layers. It is remarkable that the two-layer structure is not significantly more stable than the finite clusters. We believe that the systems prefer to have H atoms on some surface. However, since the surface area scales as $n^{2/3}$, the available area per H atom scales like $n^{-1/3}$, meaning that above a certain critical size, the finite (HAIO) $_n$ clusters will be less stable simply due to too little space for the H atoms on the surface. This effect is not found for the layers that per construction are infinite. But it suggests that there is a competition between clusters of HAIO and layers of HAIO, which may explain why different synthesis routes lead to different materials. It is not possible to determine directly the critical size of the finite (HAIO) $_n$ clusters above which they become unstable, but our results suggest that this size may be comparable to the largest clusters of the present study.

As mentioned above, HAIO is a nanostructured material, so that the individual clusters may interact with each other. In this context it is an interesting issue whether the H atoms will try to remain outside a central part, as we have observed for the individual, isolated clusters. We decided, therefore, to study the interaction between two clusters by putting two of the previously optimized clusters together. This was done as follows. We placed two clusters of n_1 and n_2 units so close to each other that they would interact. The initial structures were those of the isolated (HAIO) $_{n_1}$

Fig. 26 A graphical illustration of the structural relaxations of bringing the $n_1 = 8$ (left part) and $n_2 = 9$ (right part) HAIO clusters together. Shown are the relaxations in a cylindrical coordinate system with z and d being the position along the cylindrical axis and distance from it, respectively, both in a.u. The three panels show the displacements of the Al, O, and H atoms, individually. For details about the presentation, see the text



and $(\text{HAIO})_{n_2}$ clusters, and we considered very many relative orientations of the two clusters, out of which we chose the one that led to the lowest total energy after structural relaxation. For the case that the two clusters were brought so close that they interact, before the combined cluster was relaxed, it was found that the hydrogen atoms are placed between the two cores of AIO, but after the combined cluster was relaxed, the hydrogen atoms are only sitting on the surface of the combined cluster. This supports our consensus that clusters with hydrogen sitting on the surface of the clusters are most stable.

This is illustrated in Fig. 26 that has been obtained as follows. We consider the case of $n_1 = 8$ and $n_2 = 9$ units before and after relaxation. In each case (i.e., before and after relaxation), we calculate the center of the two parts according to Eq. (22). The line joining these two centers defines the z axis in a cylindrical coordinate system with $z = 0$ being the midpoint between the two centers. Subsequently, we superpose the two coordinate systems in one figure and show the initial and final values of z and the distance to the z axis (denoted d) for each atom separately by joining these points with a straight line. Finally, we depict these lines for each type of atom individually. The stars mark the final positions of the $n_1 = 8$ system, whereas

the closed circles mark the final positions of the $n_2 = 9$ system. In particular the H atoms tend to increase d upon relaxation, i.e., to move away from the region between the two clusters. On the other hand, first of all the O atoms but also to a lesser extent the Al atoms are seeking to fill out the space between the two clusters when they are combined.

AB Semiconductor Clusters

Semiconductors are interesting materials first of all because of their electronic properties. By modifying the material, it is partly possible to vary these properties in a controllable way, and further possibilities are obtained by considering semiconductor nanoparticles. When the spatial extension of the clusters becomes comparable with that of an exciton, the electronic and optical properties of the semiconductor will depend markedly on the cluster size.

An important class of semiconductors is the II–VI and III–V semiconductors. For those an additional, relevant (and interesting) issue is that the macroscopic, crystalline material may exist in (at least) two structures that are energetically very close, i.e., the zinc blende and the wurtzite structure (Yeh et al. 1992a, b). Accordingly, it is likely that for the finite-sized clusters, the energetic ordering of those two structures may change. To access this issue, we have studied finite clusters of a larger range of II–VI and III–V semiconductor nanoparticles (Joswig et al. 2000, 2002, 2003; Roy and Springborg 2003, 2005; Sarkar and Springborg 2003; Sarkar et al. 2005; Goswami et al. 2006; Asaduzzaman and Springborg 2006; ur Rehman et al. 2011). We have used the DFTB method that includes electronic degrees of freedom and not attempted to optimize the structure completely. Instead, starting from the infinite, periodic crystals, we have cut out a spherical part that subsequently was allowed to relax to the structure of the closest local total-energy minimum. In most cases we considered stoichiometric clusters that were obtained by placing the center of the abovementioned sphere on the midpoint of a nearest-neighbor bond.

In Fig. 27 we show representative examples for the relative total energy of clusters derived from the two different crystal structures. It is clear that the relative total energy depends critically both on size and on material of the clusters. This is in agreement with experimental observations for nanoparticles of the same material but different sizes (see, e.g., Weller and Eychmüller 1996). Moreover, although the total energy does show an overall decreasing behavior as a function of size of the cluster, it is also structured.

There is an interesting correlation between the variation of the total energy and the energy gap between the highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbital. In Fig. 28 we illustrate this for a single example, i.e., CdS clusters. Similar results are obtained for the other systems. The figure demonstrates that high stability coincides with a high energy gap, which may be interpreted in terms of the so-called hard and soft acids and bases (HSAB) principle (Pearson 1997). Moreover, a more detailed analysis of the frontier orbitals gives that for the less stable structures, i.e., the ones with a small HOMO–LUMO gap,

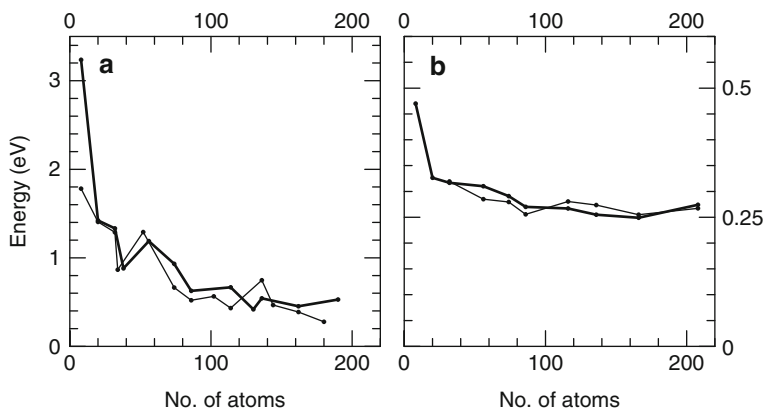


Fig. 27 The variation in total energy per atom pair for stoichiometric (a) CdS and (b) CdSe clusters for (*thick curves*) zincblende- and (*thin curves*) wurtzite-derived clusters

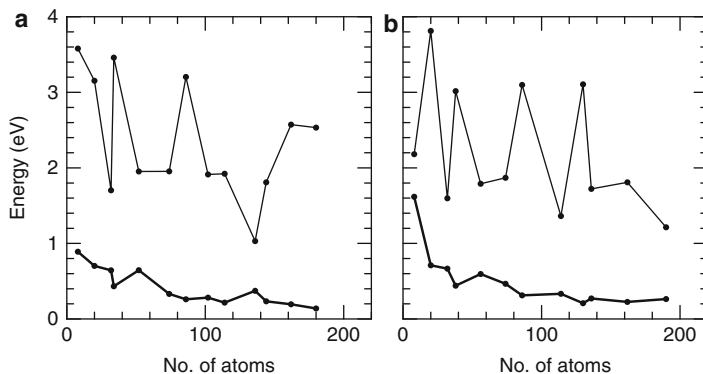


Fig. 28 Variation in the total energy per atom pair (*thick curves*) and band gap (*thin curves*) for CdS clusters with (a) zinc blende- and (b) wurtzite-derived structures as a function of the total number of atoms in the cluster

and at least one of the HOMO and LUMO orbitals is localized to the surface region. Because of these surface states, less stable clusters are particularly reactive.

From these results, we obtain also an example of how strongly and irregularly the properties can depend on the size of the clusters. For Cd_nS_n clusters, we find a HOMO–LUMO gap of around 1.6 eV for $n = 16$, which increases to 3.2 eV for $n = 17$ (Joswig et al. 2003). When analyzing the structures of those two, this difference can be explained: for $n = 16$ the cluster contains a number of single-coordinated atoms, making it less stable, whereas no such atoms are found for $n = 17$.

Finally, the fact that at least one of the frontier orbitals is localized to the surface is in agreement with experimental studies of the recombination of holes and electrons in such clusters (Lifshitz et al. 1998).

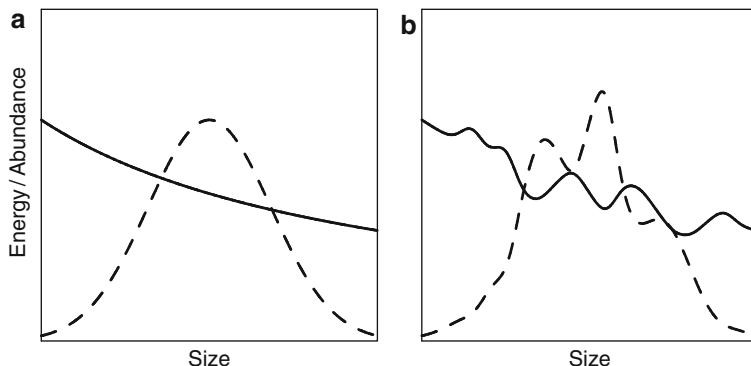


Fig. 29 Schematic presentation of the variation in the total energy per atom (*solid curve*) as a function of size of the cluster together with the size distribution (*dashed curve*) (a) as it usually is assumed together with (b) a modification as it may be in reality. The *solid curve* may also represent the optical absorption threshold as a function of cluster size

The results above have two important implications. First, in experimental studies, it is often assumed (cf. Fig. 29) that one obtains a more or less narrow size distribution that is very regular as a function of size. Moreover, the size of the clusters is often estimated through the energy of the optical transition for which it is assumed that is a monotonous function of size. However, our results give that the total energy is structured as a function of size, suggesting that the size distribution also will be structured, and also that the energy for optical transitions is less regular as a function of cluster size, although a general decrease as a function of increasing cluster size can be identified.

Second, since the stability of the clusters is strongly dependent on the presence/absence of low-coordinated atoms on the surface, it should be possible to stabilize the clusters through properly chosen surfactants or by modifying the surface in other ways (Frenzel et al. 2005). For our stoichiometric AB clusters, the outermost atoms are formed to equal parts by A and by B atoms. Considering nonstoichiometric clusters constructed, e.g., as a spherical cutout of the crystal with the center at an atomic position, the surface will contain exclusively one type of atoms. In one study, on InP clusters (Roy and Springborg 2003), we considered this case and did indeed observe a strong dependence of the properties on the surface, although the properties did not only depend on the type of the outermost atoms. A further study (Roy and Springborg 2005) demonstrated that surfactants could modify the electronic properties of the clusters, leading in most, but not all, cases to a larger HOMO–LUMO gap.

Alternatively, one may obtain more stable structures by considering the so-called core/shell nanoparticles, which are systems containing a core of one material coated by a shell of another. We studied such systems with the two materials being either CdS and CdSe or Si and Ge (Sarkar et al. 2005; Asaduzzaman and Springborg 2006) and found, besides a saturation of dangling bonds on the core part of the

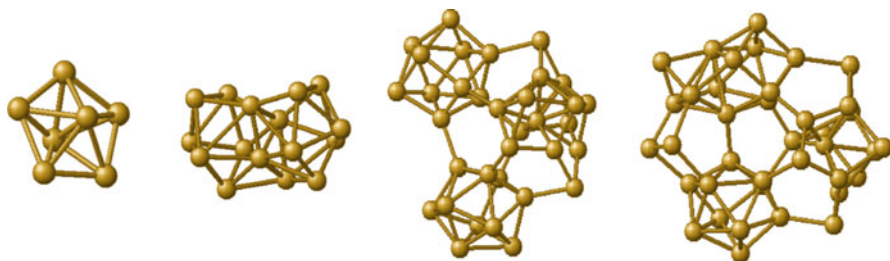


Fig. 30 The optimized structures of the Si_n clusters for $n = 7, 16, 34,$ and 38 (from left to right)

nanoparticle, that new properties may emerge from such systems. This includes that in some few cases, the HOMO and LUMO are localized to the two different materials, making a charge separation upon optical excitation possible. Furthermore, the more stable systems were obtained when the shell material was the one with the lowest surface energy.

The results that we have discussed here were all obtained by assuming that the structure of the clusters resembles that of a finite part of the crystals. This may be the case for larger clusters and clusters protected through ligands on the surface, but for smaller ones, deviations from this may occur. To study whether this assumption is realistic, we have used the genetic algorithm approach in obtaining the stablest structures for Si_n , Ge_n , and Si_nGe_n clusters with n up to 44 for the pure clusters and up to 22 for the mixed ones (ur Rehman et al. 2011).

According to the results, there is a tendency for the Ge clusters to be more open and have lower-coordinated atoms than what is found for the other two systems. For the larger clusters, particularly stable sizes are found for $n = 16, 34, 28,$ and 42 for Si_n clusters; $n = 16, 26, 39,$ and 43 for Ge_n clusters; and $n = 7, 12, 14,$ and 18 for Si_nGe_n clusters. The structures of some of those Si_n clusters are shown in Fig. 30. It is clear that the structures have a low symmetry. Moreover, the larger structures consist of compact parts, each often containing around roughly 10 atoms but sometimes being smaller. These parts are less strongly bonded with each other. This finding is consistent with the dissociation pattern where we found that it is energetically cheapest to dissociate a cluster into two parts that are of comparable size. This is in fact contrasted by the results for the Ni_n clusters discussed earlier in this review and where we found that the Ni_n clusters tend to split off just 1–3 atoms even for clusters with up to 150 atoms in total. Furthermore, for the smaller Si, Ge, or mixed clusters, the overall structure is that of an elongated, cigar-shaped, oblate object, whereas the larger clusters are more spherical.

A more careful analysis of the results reveals that the structures for the different types of clusters are not at all similar. Thus, Si atoms tend to be higher coordinated than what is found for the Ge atoms. In order to study this in more detail, we first identify an interatomic distance, d^0 , which separates bonded pairs of atoms (their mutual distance is less than d^0) from nonbonded pairs. For this purpose, we define a pair correlation function

$$g(r) = \sum_n \sum_{i \neq j} f(r - r_{nij}). \quad (35)$$

Here, the n summation is over all clusters of a given type, whereas the i and j summations are over all atoms of a given clusters. r_{nij} is the distance between the i th and j th atom of the n th cluster. Finally, $f(s)$ is a narrow, normalized Gaussian for which we chose a full width at half maximum of 0.01 a.u. The resulting function (not shown) for each of the three types of interatomic distances (i.e., Si–Si, Ge–Ge, and Si–Ge) has a fairly narrow maximum for small r . For the Si_n clusters, we found for larger r essentially only a broad feature, whereas for the Ge containing clusters, a second maximum for a slightly larger r was observed. The first maximum corresponds to the typical nearest-neighbor distance in the three different types of clusters, i.e., 5.03 a.u. for Si_n clusters, 6.73 a.u. for Ge_n clusters, and 5.02 a.u. for Si_nGe_n clusters. Above the first maximum, there is a range for which $g(r)$ becomes very close or identical to 0. We shall use the midpoint of this range in defining d^0 which, thus, becomes

$$\begin{aligned} d_{\text{Si}}^0 &= 5.35 \text{ a.u.} \\ d_{\text{Ge}}^0 &= 5.85 \text{ a.u.} \\ d_{\text{SiGe}}^0 &= 5.85 \text{ a.u.} \end{aligned} \quad (36)$$

However, only for the Ge_n clusters, $g(d^0) = 0$.

Using these values, we can next calculate the relative occurrence of different coordinations for the atoms of the clusters, i.e., $o(n_{\text{AB}})$. Thereby, n_{AB} is the number of B atoms around an atom of type A. The results are shown in Fig. 31 where we have used three different values for d^0 , i.e., those of Eq. (36) as well as those multiplied by either 0.95 or by 1.05. Some smaller differences from these different values are observed, but the main trends remain. Hence, even for the pure clusters, Ge atoms tend to be lower coordinated (hardly above fourfold coordinated) than Si atoms. This difference between Si and Ge can also explain why the Si_n clusters are more compact than the Ge_n clusters. Also for the Si–Ge clusters, there are only few Ge–Ge bonds, whereas most often heteroatomic Si–Ge bonds are found, but also some Si–Si bonds exist.

The fact that the structures are fairly open although still consisting of compact subparts makes it difficult to quantify structural similarity using the approaches we have discussed earlier in this presentation. Nevertheless, in a first study (ur Rehman et al. 2011), we used the interatomic distances in discussing structural similarity between different clusters. First we compare two clusters with the same number of atoms, N . For each of those, we calculate all $\frac{N(N-1)}{2}$ interatomic distances that subsequently are sorted. Then, we calculate

$$q = \left[\frac{2}{N(N-1)} \sum_{i=1}^{N(N-1)/2} \left(\frac{d_{\text{A},i}}{d_{\text{A}}^0} - \frac{d_{\text{B},i}}{d_{\text{B}}^0} \right)^2 \right]^{1/2}, \quad (37)$$

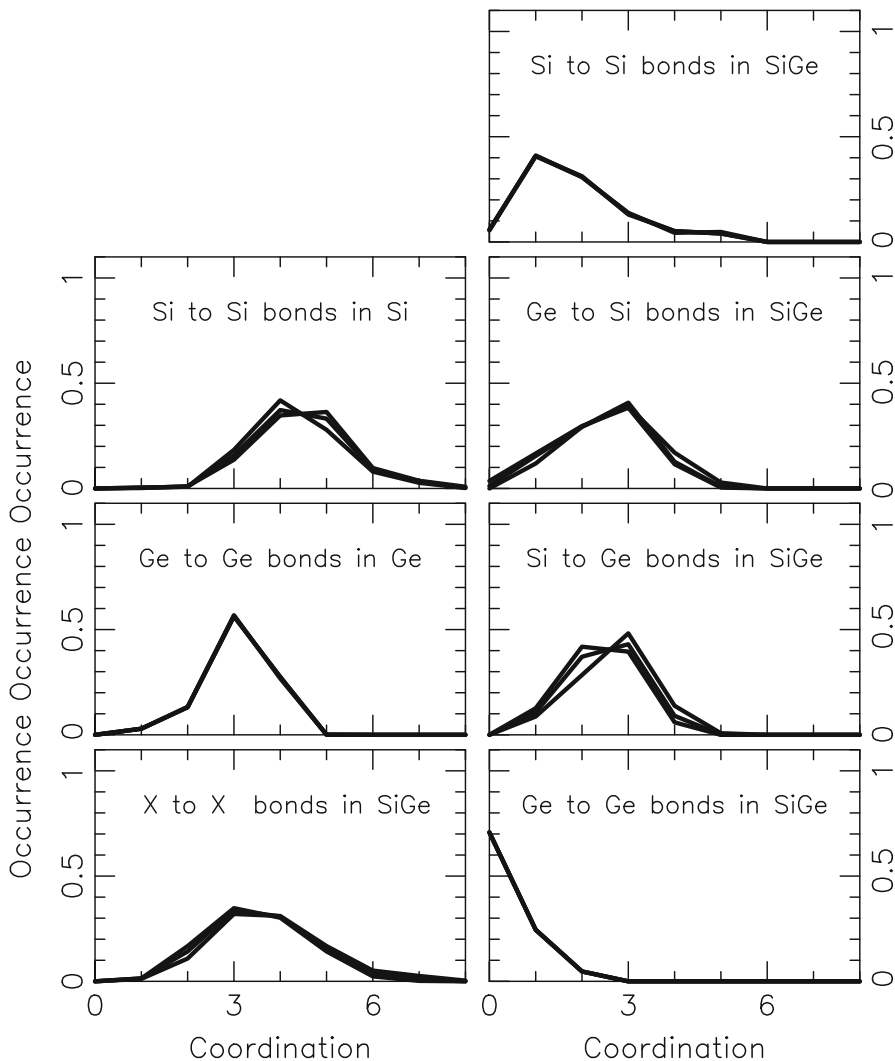


Fig. 31 The function $o(n_{AB})$ that describes how many B neighbors an A atom has in the different types of structures. In the panels, the notation “B to A in C” represents the B and A atoms in C clusters. The three different curves in each panel are obtained by using three slightly different values for determining whether two atoms are bonded

where A and B equal Si, Ge, or SiGe, and the interatomic distances have been scaled by the lengths of Eq. (36). From this, we define a similarity function according to

$$S = \frac{1}{1 + q}, \tag{38}$$

that approaches 1 (0) if the two structures are structurally very similar (different). For the Si–Ge clusters, we do not distinguish between atom types. Alternatively, we may replace the $\frac{N(N-1)}{2}$ interatomic distances with the N radial distances or the N reciprocal radial distances.

In a more recent work for quantifying structural similarity (Dong et al. 2013a), we were motivated by what you do in your mind when examining whether two systems are similar. Thus, at first each of the two structures is scaled, and subsequently one of the two is rotated and translated so that it maximally overlap with the other structure.

In describing our new approach, we will let B be the larger system with M atoms and A be the smaller one with N atoms. Also in the special case that $M = N$, we will use this notation. In detail, our approach consists then of the following steps:

1. For each of the two systems, we use scaled relative coordinates, i.e.,

$$\begin{aligned} \mathbf{s}_{A,i} &= \mathbf{r}_{A,i}, & i &= 1, 2, \dots, N \\ \mathbf{s}_{B,i} &= \mathbf{r}_{B,i}, & i &= 1, 2, \dots, M. \end{aligned} \quad (39)$$

2. Subsequently, we apply an initial rotation and translation to system A. By repeating this and the following steps for different initial rotations and translations, we hope to be able to find the optimal way of placing structure A on structure B.

$$\mathbf{s}_{A,i} \rightarrow \mathbf{s}_{A,i}' = \underline{\underline{R}}_0 \mathbf{s}_{A,i} + \mathbf{t}_0. \quad (40)$$

$\underline{\underline{R}}_0$ is the rotational matrix expressed in terms of a set of Euler angles and \mathbf{t}_0 is the translation vector.

3. For each atom, i , of system A, we identify that atom in system B, $n_{BA}(i)$, that is closest to the atom i of system A, i.e., the partner of atom i . $n_{BA}(i)$ is that atom of system B that makes

$$q_i = |\mathbf{s}_{A,i}' - \mathbf{s}_{B,n_{BA}(i)}|^2 \quad (41)$$

as small as possible. It may happen that two (or more) atoms of system A, i_1 and i_2 , have the same partner, i.e., $n_{BA}(i_1) = n_{BA}(i_2)$. In that case, $n_{BA}(i_1)$ ($n_{BA}(i_2)$) is kept if the corresponding value of q_{i_1} is smaller (larger) than the equivalent value of q_{i_2} , and for the i_2 th (i_1 th) atom of system A, another partner of system B is found. Thus, no two atoms of system A will have the same partners.

4. We apply a rotation and a translation to the complete system A. This changes the positions of the atoms in system A according to

$$\mathbf{s}_{A,i}' \rightarrow \mathbf{s}_{A,i}'' = \underline{\underline{R}} \mathbf{s}_{A,i}' + \mathbf{t}. \quad (42)$$

Here, $\underline{\underline{R}}$ is a rotation matrix, described through three Euler angles α , β , and γ . Moreover, \mathbf{t} is a translation containing the three parameters x , y , and z . Thus, in total, 6 parameters, α , β , γ , x , y , and z , define the displacement of system A.

5. These 6 parameters are optimized so that

$$Q(\alpha, \beta, \gamma, x, y, z) = \frac{1}{N} \sum_{i=1}^N |\mathbf{s}_{A,i}'' - \mathbf{s}_{B,n_{BA}(i)}|^2 \quad (43)$$

is minimized.

6. It cannot be excluded that through the transformation of Eq. (42), the atoms of system A will be closer to other atoms than its original partners. Therefore, we repeat step 3 and identify a new set of partners for the atoms of system A. For this new set, we calculate Q of Eq. (43) for the special case that $\alpha = \beta = \gamma = 0$ and $\mathbf{t} = \mathbf{0}$ (i.e., the rotation and translation are not optimized). If this new value of Q is lower than the one we have optimized in the previous step, we go back to step 4 for this new set of partners.
7. Otherwise, we may return to step 2 for a new set of initial rotations and translations.
8. Finally, in order to check whether system A is the mirror image of system B, the whole procedure is repeated by replacing, e.g., all z coordinates of the atoms of system A by their negative values (or, in the special case, that they all vanish, then the x or y values).

From the smallest value of Q_R , we define finally the similarity function as

$$S = \frac{1}{1 + Q_R^{1/2}}. \quad (44)$$

This new approach and the earlier one, given through Eqs. (37) and (38), were applied for comparing Si_n with Ge_n , Si_{2n} with Si_nGe_n , Ge_{2n} with Si_nGe_n , Si_n with Si_{n-1} , Ge_n with Ge_{n-1} , Si_nGe_n with $\text{Si}_{n-1}\text{Ge}_{n-1}$, and Si_n , Ge_n , and Si_nGe_n with the diamond crystal structure. For the diamond crystal structure, we constructed a cubic cutout of the crystal containing $5 \times 5 \times 5$ units of each 8 atoms. These 8 atoms were placed at $(0, 0, 0)$, $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, $(\frac{1}{2}, \frac{1}{2}, 0)$, $(\frac{3}{4}, \frac{3}{4}, \frac{1}{4})$, $(0, \frac{1}{2}, \frac{1}{2})$, $(\frac{1}{4}, \frac{3}{4}, \frac{3}{4})$, $(\frac{3}{4}, \frac{1}{4}, \frac{3}{4})$, and $(\frac{1}{2}, 0, \frac{1}{2})$, and the three lattice vectors are $(0, 0, 1)$, $(0, 1, 0)$, and $(1, 0, 0)$.

We considered two different ways of choosing the scaling parameters of Eq. (39). In one set of calculations, we set these for all Si clusters equal to 5.35 a.u., whereas we for all Ge and all SiGe clusters used a value of 5.85 a.u. Alternatively, we used the average value for the nearest- and next-nearest-neighbor distance of the crystalline material. For the mixed clusters, we chose the average of the values for the pure clusters.

Figure 32 shows the similarity function as calculated in our previous study together with the similarity functions based on either the radial distances, the inverse radial distances, the interatomic distances, and the inverse interatomic distances. In each case, $S \simeq 1$ suggests that the two structures that are being compared are very similar. Thus, any drop in a similarity function in the left panels in Fig. 32 suggests that a structural change has taken place at that size. If S remains low, each cluster has a structure different from that of the cluster of the neighboring, lower size, whereas

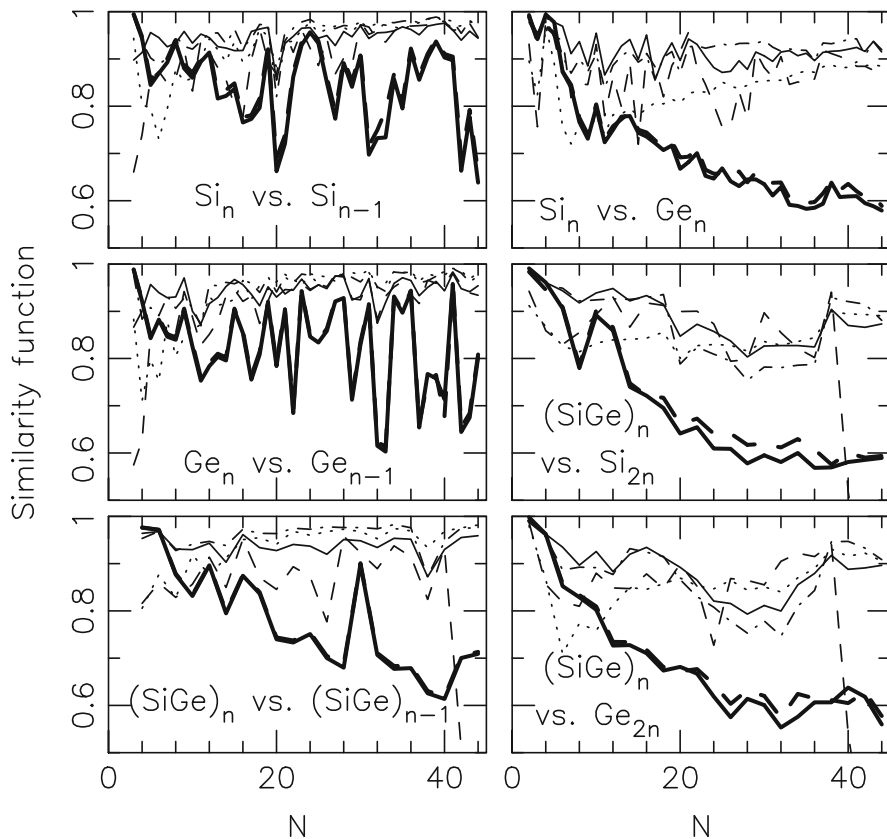


Fig. 32 The *thick curves* show the similarity functions of our new approach, whereas the *thin curves* show those of our earlier approach. N is the total number of atoms in the clusters, i.e., n for the pure Si_n and Ge_n clusters and $2n$ for the Si_nGe_n clusters. With the new approach, the *solid curves* mark the results based on setting the scaling parameters for Si_n , Ge_n , and Si_nGe_n equal to 5.35, 5.85, and 5.85 a.u., respectively, whereas they were set equal to 5.851, 6.095, and 5.973 a.u., respectively, for the results of the *dashed curves*. With the old approach, only the former set of scaling parameters was used. Moreover, then the *solid curves* show the results based on the radial distances, the *dashed curves* the results based on the inverse radial distances, the *dash-dotted curves* those based on the interatomic distances, and the *dotted curves* those based on the inverse interatomic distances. In the last case, q of Eq. (37) has been multiplied by 10

if S after a jump to a low value becomes close to 1, then a new structural motif has emerged that is kept for more cluster sizes.

In Fig. 32 we first notice that the two sets of results with the new approach for different scaling parameters lie almost on top of each other, implying that this approach is not very sensitive to the value of these scaling parameters, as long as these are chosen reasonably. Second, we see that S based on the inverse interatomic

distances possesses very little structure even after multiplying q of Eq. (37) by 10 (so that S in this case covers essentially the same range as in the other cases). Accordingly, this choice does not provide very much information. Third, for S based on the inverse radial distances, we see very low values for certain cluster sizes, which is due to the fact that one atom happens to sit very close to the center of the cluster.

Figure 32 shows also that the structures of neighboring sizes of the same types of systems (i.e., Si clusters, Ge clusters, or SiGe clusters) are much more similar than are the structures of the same sizes but different types of systems. In fact, the latter appear to become increasingly different with increasing size. These findings match actually what a simple inspection of the structures intuitively gives.

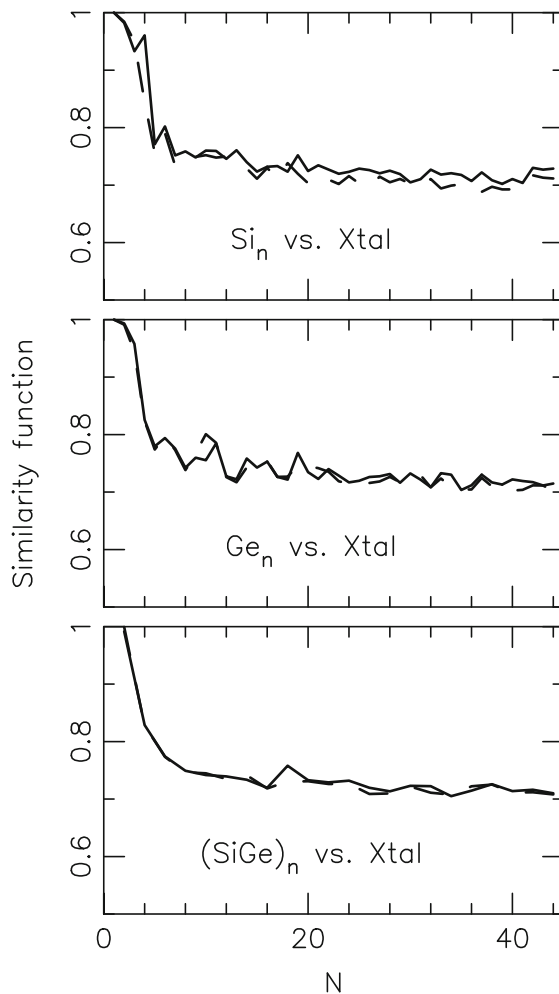
Next, we compare the structures of the clusters with the diamond crystal structure. The results of this comparison are shown in Fig. 33. It is interesting to notice that also in this case the results are independent of the precise values of the scaling parameters. Thus, although the comparison of two larger, identical cutouts of a crystal structure with slightly different scaling parameters will yield a similarity function below 1, the relative low values in Fig. 33 even for the largest clusters are not a result of inadequately chosen scaling parameters but indeed imply that the structures of the clusters are quite different from fragments of the diamond structure.

Metcars

With the last example of this presentation, we shall discuss the systems that initially led us to start the theoretical work on unbiased structure optimization for nanoparticles.

In the beginning of the 1990s, the group around Castleman in Pennsylvania produced metal-containing carbon clusters M_mC_n . In mass abundance spectra with M being Ti, they observed a peak for $(m, n) = (8, 12)$ (Guo et al. 1992), suggesting that this cluster was particularly stable. They proposed a cage-like structure of cubic T_h symmetry with 12 five-membered rings, each containing two metal and three carbon atoms. Thus, the topology is as that of the smallest possible fullerene molecule C_{20} but containing metal atoms. An alternative description of the structure is to consider it as formed by a cube of the eight metal atoms. The 12 carbon atoms form six C_2 units that are placed symmetrically above the sides of the cube with the C–C bonds parallel to the edges. Later, theoretical studies – first of all by Dance (1992) – suggested that the structure instead had tetragonal T_d symmetry. Compared with the T_h symmetry, the C_2 dimers are for the T_d symmetry lying parallel to the diagonals of the six sides of the cube of the eight metal atoms, and in addition, the cube becomes somewhat distorted. First recently, theoretical and experimental studies agree that the tetragonal structure is the correct one (see, e.g., Rohmer et al. 2000). In the meantime it has been found that also for

Fig. 33 The similarity function for the comparison of the structures of the Si_n , Ge_n , and Si_nGe_n clusters with a fragment of the diamond crystal. N is the total number of atoms in the clusters, i.e., n for the pure Si_n and Ge_n clusters and $2n$ for the Si_nGe_n clusters. The *solid curves* mark the results based on setting the scaling parameters for Si_n , Ge_n , and Si_nGe_n equal to 5.35, 5.85, and 5.85 a.u., respectively, whereas they were set equal to 5.851, 6.095, and 5.973 a.u., respectively, for the results of the *dashed curves*



other metals (e.g., Hf, Zr, V), those so-called metallocarbohedrenes or metcars can be formed, and even metcars with more different types of metals have been produced. However, whether they occur in significant amounts depends strongly on the experimental conditions, suggesting that kinetic effects play a significant role in their production.

In order to study the properties of these Ti-based clusters further, the group around Ganteför in Konstanz, Germany, decided to study a whole class of Ti_mC_n clusters with different values of (m, n) centered around $(m, n) = (8, 12)$ (Blessing 1998). Using two different sources, Ti and C atoms were produced that subsequently were allowed to form the clusters mentioned above. After ionization, a magnetic field could be used to mass-separate the clusters, and finally, photoelectron spectra

could be recorded for these mass-selected clusters. Accompanying theoretical studies would then be very useful as a support for the interpretation of the experimental results.

As one of the clusters of the experiments, let us choose Ti_6C_{10} , i.e., $(m, n) = (6, 10)$, whose structure is absolutely unknown. One suggestion could be to assume that it has a structure that can be formed by removing two carbon and two titanium atoms from the larger $(m, n) = (8, 12)$ cluster, but the first question is then: which four atoms shall be removed? Starting with the two proposed structures (i.e., those of T_h and T_d symmetry) for the $(m, n) = (8, 12)$ cluster mentioned above, we tried this strategy but obtained a new structure in each attempt. The simplest approach would then be to choose the structure of the lowest total energy. However, it has to be remembered that the experiment is performed in a completely different way: instead of starting with the larger $(m, n) = (8, 12)$ cluster and from that removing some atoms, the $(m, n) = (6, 10)$ cluster is formed from essentially isolated atoms, and therefore, it is not at all given that its structure has any resemblance to that of the $(m, n) = (8, 12)$ cluster.

Therefore, in order to obtain unbiased information on the structures of the different Ti_mC_n clusters, we started applying the genetic algorithms that were discussed earlier in this section together with the DFTB method and applied the approach to the metcars (Joswig et al. 2001; Joswig and Springborg 2008). In Fig. 34 some of the results are summarized. This figure shows the total energies for two series of Ti_mC_n metcars, i.e., with $m = 7, 8$ and $n = 10\text{--}14$. We have defined an energy gain, for clusters with same m , as the difference in the total energy of $\text{Ti}_m\text{C}_{n-1}$ and Ti_mC_n . For $m = 8$, experiments found the Ti_8C_{12} metcar to be a particularly stable cluster. However, our calculations do not find a specially high energy gain for this cluster, although they do for the Ti_7C_{12} cluster. Also the HOMO–LUMO gap does not indicate a high stability for the Ti_8C_{12} metcar. Actually, it has the smallest gap in the $m = 8$ series [cf. Fig. 34f]. Furthermore, there is no evidence that the number of bonds is related to the stability of the clusters, as also indicated by the results in Fig. 34. In total, the structural properties give no hints to the experimentally observed special stability of the Ti_8C_{12} metcar.

Instead, one may suggest that effects (for instance, kinetic effects) that are beyond the calculational approach are important. An experimental support for this suggestion can be found in the work of van Heijnsbergen et al. (1999). These authors found that the mass abundance spectra depend critically on experimental parameters so that in one experiment the Ti_8C_{12} metcar showed a large abundance, whereas in another it hardly occurred and a $\text{Ti}_{14}\text{C}_{13}$ cluster (which has a structure similar to a part of the rocksalt crystal; cf. Joswig and Springborg 2008) was found in large abundance.

Finally, we mention that exactly the Ti_6C_{10} metcar that we discussed above was found in our calculation to be particularly stable and of special high symmetry; see Fig. 35. However, so far it has not been discussed in detail in connection with experimental studies.

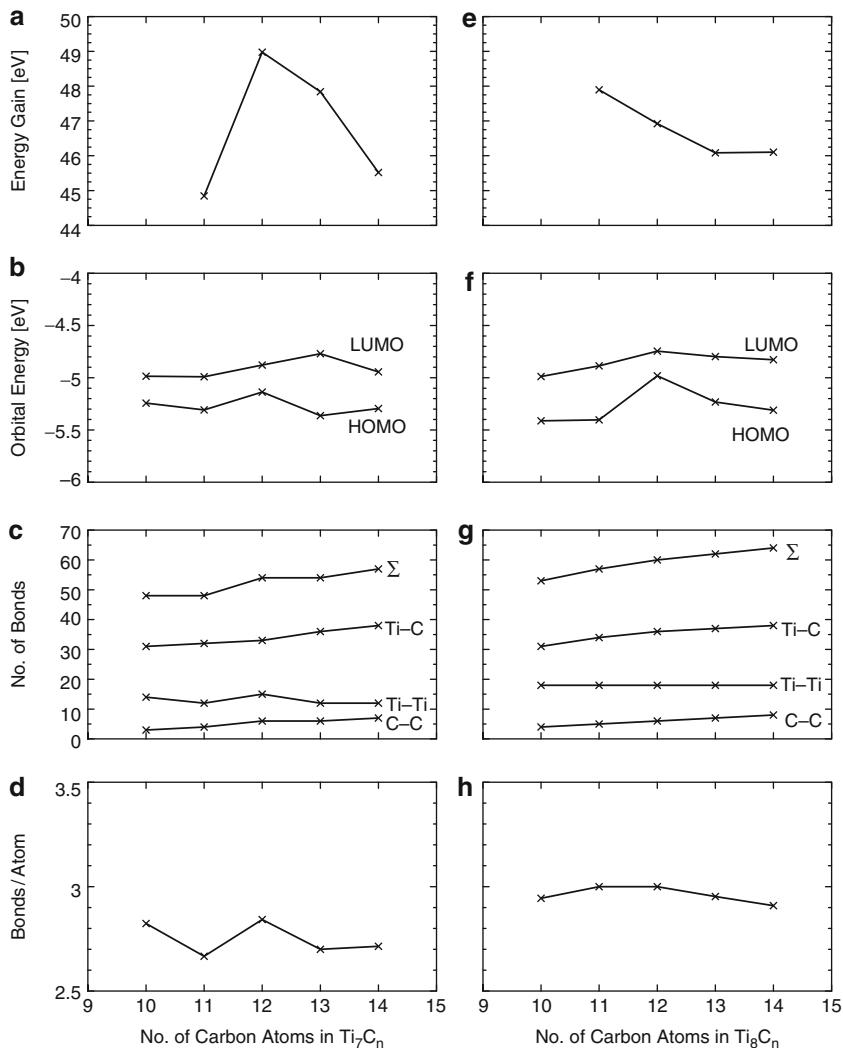
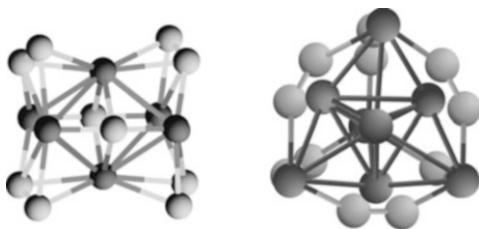


Fig. 34 (a), (e) Energy gain; (b), (f) HOMO and LUMO orbital energies; (c), (g) number of C-C, Ti-Ti, and Ti-C bonds and their total number; and (d), (h) total number of bonds divided by total number of atoms for (left) Ti_7C_n and (right) Ti_8C_n as functions of n

Fig. 35 Optimized structure of (right) Ti_6C_{10} and (left) Ti_8C_{12} . Light (dark) spheres represent C (Ti) atoms



Conclusions

In this overview we have presented various, typical results from our theoretical studies on structural and electronic properties of clusters and colloids. We have put the main emphasis on presenting the principles behind our approaches and less on the discussion of the different systems, separately. It shall be added that there are many other, related studies by others, but that we have here chosen to focus on our own work.

The major challenge related to studies as those we have presented here is the complexity of the calculations due to the fairly large size combined with an often low symmetry. In addition, when aiming at extracting trends, as we have been, one needs to study not only one cluster size of a given material but a whole range, adding to the computational demands.

Moreover, the fact that there is no standard method that can be used in identifying unambiguously the structure of the lowest total energy for a given system means that one has to apply more or less intelligent methods in the structure optimization and, furthermore, that there never will be any guarantee that the structure of the lowest total energy indeed is identified. Experience has shown that specially developed methods for unbiased structure optimization are the most efficient ones, whereas others like molecular-dynamics simulations and random searches often perform considerably worse. In this overview we have presented some few intelligent methods, including genetic algorithms, basin-hopping methods, and our own Aufbau/Abbau method. Alternatively, one may also consider only certain types of structures when aiming at studying special effects. The latter we illustrated through our study on AB semiconductor particles for which an interesting issue is the relative stability of cluster structures that can be derived from two different crystal structures (in this case, the zinc blende and wurtzite structures).

But even when applying such intelligent structure optimization methods, one often has to perform maybe even several 10^6 s of total-energy calculations for a given cluster. Therefore, it is often necessary to use approximate total-energy methods that then shall be chosen carefully according to the properties of the system of interest. In this presentation, we have discussed embedded-atom methods and Gupta potentials and similar approaches as useful methods for metallic systems with delocalized electrons, the DFTB method for systems with directional, covalent bonds, and the jellium model, also for metallic systems with delocalized electrons. Due to the approximations inherent in such approaches, it is always important to assess their accuracy through comparison with results of more accurate calculations or with other studies on similar systems.

Finally, we placed much emphasis in presenting various descriptors that we have developed in order to be able to extract general chemical or physical information from the, often enormous amount of, information from the calculations. This included descriptors for analyzing stability, shape, spatial distribution of atoms and orbitals, and structural similarity.

Acknowledgements The work that has been presented in this overview would not have been possible without the active participation of a number of persons, including (in alphabetic order) Abu Asaduzzaman, Markus Burkhard, Yi Dong, Valeri G. Grigoryan, Elisaveta Hristova (née Kasabova), Jan-Ole Joswig, Mohammad Molayem, Francisco Morales Morillon, Yong Pang, Sudip Roy, Pranab Sarkar, Sashi Satpathy, Gotthard Seifert, Violina Spurk (née Tevekeliyska), Michael Veith, and Ingolf Warnke. Moreover, we have benefited from financial support through the German Research Council (DFG).

Bibliography

- Alamanova, D., Grigoryan, V. G., & Springborg, M. (2006). Theoretical study of structure and energetics of gold clusters with the EAM method. *Zeitschrift für Physikalische Chemie*, *220*, 811–829.
- Alamanova, D., Grigoryan, V. G., & Springborg, M. (2008). Deposition of copper clusters on the Cu(111) surface. *Surface Science*, *602*, 1413–1422.
- Alamanova, D., Springborg, M., & Grigoryan, V. G. (2007). Theoretical study of the structure and energetics of silver clusters. *Journal of Physical Chemistry C*, *111*, 12577–12587.
- Asaduzzaman, A. Md., & Springborg, M. (2006). Structural and electronic properties of Si/Ge nanoparticles. *Physical Review B*, *74*, 165406.
- Baletto, F., & Ferrando, R. (2005). Structural properties of nanoclusters: Energetics, thermodynamic, and kinetic effects. *Reviews of Modern Physics*, *77*, 371–423.
- Blessing, N. (1998). Diplomarbeit, University of Konstanz, Konstanz.
- Brack, M. (1993). The physics of simple metal clusters: Self-consistent jellium model and semiclassical approaches. *Reviews of Modern Physics*, *65*, 677–732.
- Bulusu, S., Li, X., Wang, L.-S., & Zeng, X. C. (2006). Evidence of hollow golden cages. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 8326–8330.
- Cleri, F., & Rosato, V. (1993). Tight-binding potentials for transition metals and alloys. *Physical Review B*, *48*, 22–33.
- Dance, I. G. (1992). Geometric and electronic structures of $[Ti_8C_{12}]$: Analogies with C_{60} . *Journal of the Chemical Society, Chemical Communications*, *1992*, 1779–1780.
- Daw, M. S., & Baskes, M. I. (1983). Semiempirical, quantum mechanical calculation of hydrogen embrittlement in metals. *Physical Review Letters*, *50*, 1285–1288.
- Daw, M. S., & Baskes, M. I. (1984). Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals. *Physical Review B*, *29*, 6443–6453.
- Daw, M. S., Foiles, S. M., & Baskes, M. I. (1993). The embedded-atom method: A review of theory and applications. *Materials Science Reports*, *9*, 251–310.
- de Heer, W. A. (1993). The physics of simple metal clusters: Experimental aspects and simple models. *Reviews of Modern Physics*, *65*, 611–676.
- Dong, Y., Burkhardt, M., Veith, M., & Springborg, M. (2005). Electronic and structural properties of nanostructured HAIO. *Journal of Physical Chemistry B*, *109*, 22820–22829.
- Dong, Y., & Springborg, M. (2007). Unbiased determination of structural and electronic properties of gold clusters with up to 58 atoms. *Journal of Physical Chemistry C*, *111*, 12528–12535.
- Dong, Y., Springborg, M., & Pang, Y., Morales Morillon, F. (2013a). Analyzing the properties of clusters: Structural similarity and heat capacity. *Computational and Theoretical Chemistry*, *1021*, 16–25.
- Dong, Y., Springborg, M., & Warnke, I. (2013b). Structural and thermodynamic properties of Au_{2-58} clusters. *Progress in Theoretical Chemistry and Physics*, *27*, 181–193.
- Ferrando, R., Jellinek, J., & Johnston, R. L. (2008). Nanoalloys: From theory to applications of alloy clusters and nanoparticles. *Chemical Reviews*, *108*, 846–910.

- Finnis, M. W., & Sinclair, J. E. (1984). A simple empirical N -body potential for transition metals. *Philosophical Magazine A*, 50, 45–55.
- Foiles, S. M., Daw, M. S., & Baskes, M. I. (1986). Embedded-atom-method functions for the fcc metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys. *Physical Review B*, 33, 7983–7991.
- Frenzel, J., Joswig, J.-O., Sarkar, P., Seifert, G., & Springborg, M. (2005). The effects of organisation, embedding and surfactants on the properties of cadmium chalcogenide (CdS, CdSe and CdS/CdSe) semiconductor nanoparticles. *European Journal of Inorganic Chemistry*, 2005, 3585–3596.
- Goldberg, D. E. (1989). *Genetic algorithms in search, optimization and machine learning*. Reading: Addison-Wesley.
- Goswami, B., Pal, S., Sarkar, P., Seifert, G., & Springborg, M. (2006). Study of structural, electronic, and optical properties of Zn_mSe_n clusters. *Physical Review B*, 73, 205312.
- Grigoryan, V. G., & Springborg, M. (2001). A theoretical study of the structure of Ni clusters (Ni_N). *Physical Chemistry Chemical Physics*, 3, 5125–5129.
- Grigoryan, V. G., & Springborg, M. (2003). Structure and energetics of Ni clusters with up to 150 atoms. *Chemical Physics Letters*, 375, 219–226.
- Grigoryan, V. G., & Springborg, M. (2004). Structural and energetic properties of nickel clusters: $2 \leq N \leq 150$. *Physical Review B*, 70, 205415.
- Grigoryan, V. G., & Springborg, M. (2010). Vibrational and thermodynamic properties of metal clusters with up to 150 atoms calculated by the embedded-atom method. *Physical Review B*, 83, 155413.
- Grigoryan, V. G., Alamanova, D., & Springborg, M. (2006). Structure and energetics of Cu_N clusters with ($2 \leq N \leq 150$): An embedded-atom-method study. *Physical Review B*, 73, 115415.
- Guo, B. C., Kearns, K. P., & Castleman Jr., A. W. (1992). $Ti_8C_{12}^+$ -metallo-carbohedrenes: A new class of molecular clusters? *Science*, 255, 1411–1413.
- Gupta, R. P. (1981). Lattice relaxation at a metal surface. *Physical Review B*, 23, 6265–6270.
- Holland, J. H. (1975). *Adaption in natural algorithms and artificial systems*. Ann Arbor: University of Michigan Press.
- Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. *Physical Review*, 136, B864–B871.
- Hristova, E., Dong, Y., Grigoryan, V. G., & Springborg, M. (2008a) Structural and energetic properties of Ni-Cu bimetallic clusters. *Journal of Physical Chemistry A*, 112, 7905–7915.
- Hristova, E., Grigoryan, V. G., & Springborg, M. (2008b) Structure and energetics of equiatomic K-Cs and Rb-Cs binary clusters. *Journal of Chemical Physics*, 128, 244513.
- Hristova, E., Grigoryan, V. G., & Springborg, M. (2009a) Theoretical study on the structural properties of alkali-metal heteroclusters. *European Physical Journal D*, 52, 35–38.
- Hristova, E., Grigoryan, V. G., & Springborg, M. (2009b) Structure and stability of Ag clusters on Ag (111) and Ni (111) surfaces. *Surface Science*, 603, 3339–3345.
- Jellinek, J., & Krissinel, E. B. (1996). Ni_nAl_m alloy clusters: Analysis of structural forms and their energy ordering. *Chemical Physics Letters*, 258, 283–292.
- Joswig, J.-O., Springborg, M. (2003). A genetic-algorithms search for global minima of aluminum clusters using a Sutton-Chen potential. *Physical Review B*, 68, 085408.
- Joswig, J.-O., & Springborg, M. (2007). Size-dependent structural and electronic properties of Ti_n clusters ($n \leq 100$). *Journal of Physics: Condensed Matter*, 19, 106207.
- Joswig, J.-O., & Springborg, M. (2008). The influence of C_2 dimers on the stability of Ti_mC_n metcar clusters. *Journal of Chemical Physics*, 129, 134311.
- Joswig, J.-O., Springborg, M., Seifert, G. (2000). Structural and electronic properties of cadmium-sulfide clusters. *Journal of Physical Chemistry B*, 104, 2617–2622.
- Joswig, J.-O., Springborg, M., & Seifert, G. (2001). Structural and electronic properties of small titanium-carbon clusters (metcars). *Physical Chemistry Chemical Physics*, 3, 5130–5134.
- Joswig, J.-O., Roy, S., Sarkar, P., & Springborg, M. (2002). Stability and band gap of semiconductor clusters. *Chemical Physics Letters*, 365, 75–81.

- Joswig, J.-O., Seifert, G., Niehaus, T. A., & Springborg, M. (2003). Optical properties of cadmium sulfide clusters. *Journal of Physical Chemistry B*, *107*, 2897–2902.
- Kasabova, E., Alamanova, D., Springborg, M., Grigoryan, V. G. (2007). Deposition of Ni₁₃ and Cu₁₃ clusters on Ni(111) and Cu(111) surfaces. *European Physical Journal D*, *19*, 425–431.
- Knight, W. D., Clemenger, K., de Heer, W. A., Saunders, W. A., Chou, M. Y., & Cohen, M. L. (1984). Electronic shell structure and abundances of sodium clusters. *Physical Review Letters*, *52*, 2141–2143.
- Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Physical Review*, *140*, A1133–A1138.
- Li, Z., & Scheraga, H. A. (1987). Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, *84*, 6611–6615.
- Lifshitz, E., Dag, I., Litvin, I., Hodes, G., Gorer, S., Reisfeld, R., Zelner, M., & Minti, H. (1998). Optical properties of CdSe nanoparticle films prepared by chemical deposition and sol-gel methods. *Chemical Physics Letters*, *288*, 188–196.
- Lloyd, S. (2002). Computational capacity of the universe. *Physical Review Letters*, *88*, 237901.
- Lloyd, L. D., Johnston, R. L., Salhi, S., & Wilson, N. T. (2004). Theoretical investigation of isomer stability in platinum-palladium nanoalloy clusters. *Journal of Materials Chemistry*, *14*, 1691–1704.
- Molayem, M., Grigoryan, V. G., & Springborg, M. (2011). Theoretical determination of the most stable structures of Ni_mAg_n bimetallic nanoalloys. *Journal of Physical Chemistry C*, *115*, 7179–7192.
- Molayem, M., Grigoryan, V. G., & Springborg, M. (2011). Global minimum structures and magic clusters of Cu_mAg_n nanoalloys. *Journal of Physical Chemistry C*, *115*, 22148–22162.
- Parks, E. K., Nieman, G. C., Kerns, K. P., & Riley, S. J. (1997). Reactions of Ni₃₈ with N₂, H₂, and CO: Cluster structure and adsorbate binding sites. *Journal of Chemical Physics*, *107*, 1861–1871.
- Parks, E. K., Kerns, K. P., & Riley, S. J. (1998). The structure of Ni₃₉. *Journal of Chemical Physics*, *109*, 10207–10216.
- Parks, E. K., Kerns, K. P., & Riley, S. J. (2001). The structure of Ni₄₆, Ni₄₇, and Ni₄₈. *Journal of Chemical Physics*, *114*, 2228–2236.
- Pearson, R. G. (1997). *Chemical hardness*. Weinheim: Wiley-VCH.
- Pittaway, F., Paz-Borbón, L. O., Johnston, R. L., Arslan, H., Ferrando, R., Mottet, C., Barcaro, G., & Fortunelli, A. (2009). Theoretical studies of palladium-gold nanoclusters: PdAu clusters with up to 50 atoms. *Journal of Physical Chemistry C*, *113*, 9141–9152.
- Porezag, D., Frauenheim, Th., Köhler, Th., Seifert, G., & Kaschner, R. (1995). Construction of tight-binding-like potentials on the basis of density-functional theory. *Physical Review B*, *51*, 12947–12957.
- Rohmer, M. M., Bénard, M., & Poblet, J. M. (2000). Structure, reactivity, and growth pathways of metallocarbohedrenes M₈C₁₂ and transition metal/carbon clusters and nanocrystals: A challenge to computational chemistry. *Chemistry Reviews*, *100*, 495–541.
- Roy, S., & Springborg, M. (2003). A theoretical study of structural and electronic properties of naked stoichiometric and non-stoichiometric indium phosphide clusters. *Journal of Physical Chemistry B*, *107*, 2771–2779.
- Roy, S., & Springborg, M. (2005). Theoretical investigation of the influence of ligands on structural and electronic properties of indium phosphide clusters. *Journal of Physical Chemistry A*, *109*, 1324–1329.
- Sarkar P., & Springborg, M. (2003). Density-functional study of size-dependent properties of Cd_mSe_n clusters. *Physical Review B*, *68*, 235409.
- Sarkar, P., Springborg, M., & Seifert, G. (2005). A theoretical study of the structural and electronic properties of Cd/Se and CdS/CdSe core/shell nanoparticles. *Chemical Physics Letters*, *405*, 103–107.

- Seifert, G., Porezag, D., & Frauenheim, Th. (1996). Calculations of molecules, clusters, and solids with a simplified LCAO-DFT-LDA scheme. *International Journal of Quantum Chemistry*, *58*, 185–192.
- Seifert, G., & Schmidt, R. (1992). Molecular dynamics and trajectory calculations: The application of an LCAO-LDA scheme for simulations of cluster-cluster collisions. *New Journal of Chemistry*, *16*, 1145–1147.
- Springborg, M. (1999). Electronic properties, stability, and length scales of Cs_N clusters. *Journal of Physics: Condensed Matter*, *11*, 1–18.
- Springborg, M. (2000). *Methods of Electronic-Structure Calculations*. Chichester: Wiley.
- Springborg, M. (2006). Determination of structure in electronic structure calculations. In Hinchliffe, A. (Ed.), *Specialist periodical reports: Chemical modelling, applications and theory* (Vol. 4, pp. 249–323). Cambridge, UK: Royal Society of Chemistry.
- Springborg, M. (2009). Nanostructures. In Springborg, M. (Ed.), *Specialist periodical reports: Chemical modelling, applications and theory* (Vol. 6, pp. 510–574). Cambridge, UK: Royal Society of Chemistry.
- Springborg, M., Satpathy, S., Malinowski, N., Zimmermann, U., & Martin, T. P. (1996). Electronic shell structure and relative abundances in alkali-coated C_{60} . *Physical Review Letters*, *77*, 1127–1130.
- Tevekelyiska, V., Dong, Y., Springborg, M., & Grigoryan, V. G. (2010). Using theory in determining the properties of metal clusters: Sodium as a case study. In Chattaraj, P. K. (Ed.), *Aromaticity and metal clusters* (pp. 161–185). Boca Raton: Taylor & Francis.
- Tsai, C. J., & Jordan, K. D. (1993). Use of an eigenmode method to locate the stationary points on the potential energy surfaces of selected argon and water clusters. *Journal of Physical Chemistry*, *97*, 11227–11237.
- ur Rehman, H., Springborg, M., & Dong, Y. (2011). Structural and electronic properties of Si_n , Ge_n , and Si_nGe_n clusters. *Journal of Physical Chemistry A*, *115*, 2005–2015.
- van Heijnsbergen, D., von Helden, G., Duncan, M. A., van Roij, A. J. A., & Meijer, M. (1999). Vibrational spectroscopy of gas-phase metal-carbide clusters and nanocrystals. *Physical Review Letters*, *83*, 4983–4986.
- Wales, D.J. (2003). *Energy landscapes with applications to clusters, biomolecules and glasses*. Cambridge, UK: Cambridge University Press.
- Wales, D. J., & Doye, J. P. K. (1997). Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *Journal of Physical Chemistry A*, *101*, 5111–5116.
- Weller, H., & Eychmüller, A. (1996). Preparation and characterization of semiconductor nanoparticles. In Kamat, P. V., & Meisel, D. (Eds.), *Semiconductor nanoclusters* (pp. 5–22). Amsterdam: Elsevier.
- Wille, L. T., & Vennik, J. (1985). Computational complexity of the ground-state determination of atomic clusters. *Journal of Physics A*, *18*, L419–L422.
- Yeh, C.-Y., Lu, Z. W., Froyen, S., & Zunger, A. (1992). Predictions and systematization of the zinc-blende–wurtzite structural energies in binary octet compounds. *Physical Review B*, *45*, 12130–12133.
- Yeh, C.-Y., Lu, Z. W., Froyen, S., & Zunger, A. (1992). Zinc-blende–wurtzite polytypism in semiconductors. *Physical Review B*, *46*, 10086–10097.

The Response of Extended Systems to Electrostatic Fields

32

Michael Springborg, Bernard Kirtman, and Mohammad Molayem

Contents

Introduction	1416
The Gauge	1419
The Dipole Moment for the Finite System	1420
Quasi-One-Dimensional Systems	1421
Higher Dimensions	1424
The Dipole Moment for the Periodic Case	1427
The Single-Electron Case	1428
The Many-Electron Case	1430
Quasi-One-Dimensional Systems	1433
Higher Dimensions	1437
The Single-Particle Equations in the Periodic Case	1439
The Dipole Moment Operator	1439
The Secular Equation	1442
The Vector Potential Approach	1443
Solving the Secular Equation	1444
Applications	1447
Structure	1447
Converse Piezoelectricity	1450
The Vector Potential and Time-(In)dependent Fields	1454
DFT Catastrophe	1454

Physical and Theoretical Chemistry, University of Saarland, Saarbrücken, Germany

School of Materials Science and Engineering, Tianjin University, Tianjin, China
e-mail: m.springborg@mx.uni-saarland.de

B. Kirtman

Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA, USA
e-mail: kirtman@chem.ucsb.edu

M. Molayem

Physical and Theoretical Chemistry, University of Saarland, Saarbrücken, Germany
e-mail: m.molayem@mx.uni-saarland.de

Conclusions	1455
Bibliography	1456

Abstract

We present an overview of our understanding of how extended materials exposed to an electrostatic field can be treated theoretically. We concentrate on materials that are regular so that they can be approximated as being infinite and periodic. We show how expressions for the dipole moment per unit that were suggested previously within the so-called modern theory of polarization can be obtained. Subsequently, we present the single-particle equations that describe how the electrons respond to an external, electrostatic field and present a numerically efficient method of solution. We also discuss how structural responses in the presence of an electrostatic field can be calculated automatically. We demonstrate that bulk properties of macroscopic systems will depend upon the surfaces and demonstrate how these properties may be evaluated from calculations on the infinite and periodic system even though the latter does not contain surfaces. As one example, we show how the surfaces will contribute to the so-called converse piezoelectric effect. In addition, the polarizabilities induced by the structural response may be larger than the corresponding electronic polarizabilities. Finally, problems related to the description of the responses within density-functional theory are briefly discussed.

Introduction

The response to external electric and magnetic fields provides a fundamental tool for studying and altering the properties of chemical systems with numerous attendant applications. An indispensable ingredient for effectively utilizing this tool is the ability to relate the properties to electronic structure through chemical computations. Such computations depend, of course, on having an appropriate theoretical basis along with suitable algorithms for implementing the theory. In the case of atoms and molecules much has been achieved in this regard and many aspects are very well established. On the other hand, for infinite periodic systems (or systems that, in first approximation, may be regarded as such), developments are in a much more primitive state. In the case of uniform electric fields, one can relate this issue to the properties of the scalar electronic interaction potential, which is usually employed in atomic and molecular calculations. This potential is proportional to the quantum-mechanical operator \mathbf{r} (i.e., the electronic position) and is, therefore, nonperiodic and unbounded from below.

This situation is illustrated in Fig. 1 where we show the potential for an atom, a small molecule, and a long-chain compound both in the absence and in the presence of an external electrostatic field. For the atom, the electrostatic field leads to a situation where the lowest point on the potential curve is found infinitely far from the nucleus so that the bound states are changed into resonances. In the case of the small

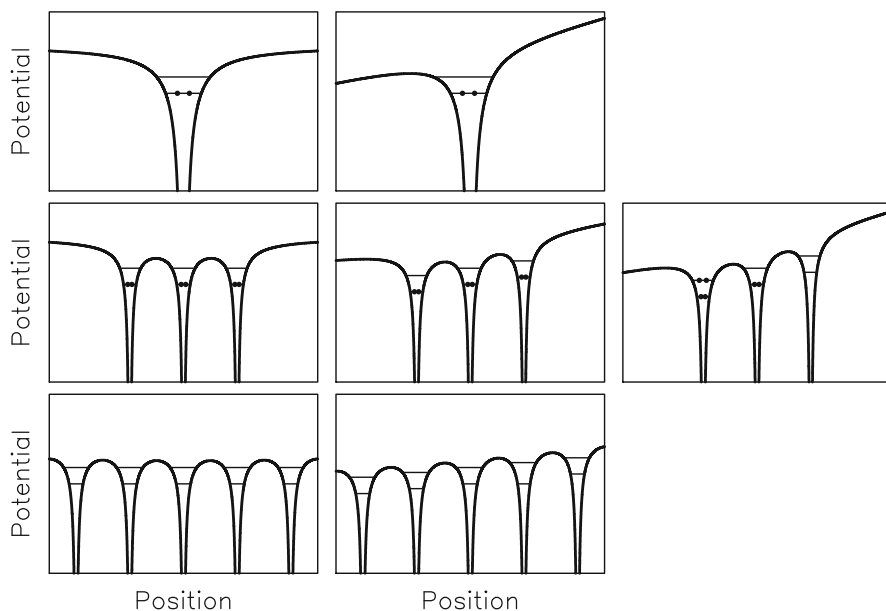


Fig. 1 Schematic representation of the potential for (*top row*) a single atom, (*middle row*) a three-atomic molecule, and (*bottom row*) a very large-chain compound. In each row, the leftmost panel shows the situation without an external electrostatic field, whereas a such is switched on in the other panels. The *short lines* symbolize electronic levels, and the *dark circles* their occupancies

molecule, a sufficiently strong field will cause the electrons to tunnel from one end to the other (Zener tunneling). Both these situations can be avoided in theoretical calculations by applying basis sets with functions that are localized to the system of interest and by applying sufficiently weak, external electrostatic fields. But neither of these methods seems able to allow for the treatment of the long-chain compound when it is exposed to an electrostatic field.

The systems that we shall study contain a large sequence of identical, regularly placed units for which deviations of this regularity only may occur near their boundaries. In that case, an alternative to the large, finite systems is to consider them as being infinite and periodic. Ultimately, any true system is finite, and, therefore, when treating it as being infinite and periodic, it is important to assure that the results are identical to those found for the large, finite system. In this presentation, we shall focus on the responses of such systems to electrostatic field and have, accordingly, a situation like that of the third example in Fig. 1 when treating them as being large and finite. As we shall see, for the infinite, periodic system the situation is much less obvious than what may intuitively have been imagined. Therefore, a close comparison between the results for large, finite systems and those for infinite, periodic ones becomes of ultimate importance.

A number of different methods have been suggested for dealing with the problems presented by the scalar interaction potential as applied to infinite periodic

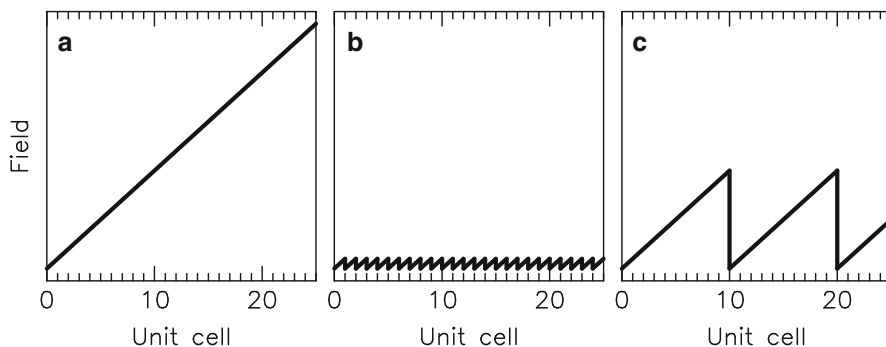


Fig. 2 Different proposals for how to approximate the potential from an electrostatic field. (a) shows the true potential, whereas in (b) and (c), approximations are shown that keep (b) the lattice or (c) BvK periodicity

systems. The most straightforward is to study finite clusters of increasing size and, then, extrapolate to the infinite size limit. At the first-principles level, this approach is prohibitively CPU demanding for 3D solids but can be manageable for quasilinear polymers (see, e.g., Champagne et al. (1997) for an early example of such studies). Often a key issue in such cases is how to extrapolate when the functional dependence of the property on chain length is unknown (Weniger and Kirtman 2003; Kudin et al. 2005; Berger et al. 2005).

A different approach, standard in solid-state physics, is to adopt Born-von Kármán (BvK) periodic boundary conditions in order to describe the infinite-system limit. Then, by construction, the system does not possess any surfaces, and any computed property is, by definition, “bulk.” At the same time, the electron position, \mathbf{r} , may be approximated as a sawtooth function with the periodicity of the BvK zone (Kunc and Resta 1983; Schmidt and Springborg 1999; Sebastiani and Parrinello 2003; Darrigan et al. 2003) (see Fig. 2). Despite some successes, we have shown (Springborg et al. 2004) (as discussed in further detail later on) that an ordinary sawtooth cannot fully account for the polarization. It is missing a contribution that may be expressed as being due to a flow of electronic charge throughout the system. This charge flow term is especially important when there is substantial electron delocalization. Recently, we have revisited the sawtooth potential and shown how it may be modified to yield polarization properties other than the dipole moment for the response to a finite field (Kirtman et al. 2009). In addition to these important exceptions, the convergence behavior of the modified potentials is poor, i.e., quite large systems need to be treated before the infinite-system limit can be identified.

Within the past 10–20 years, a “modern theory of polarization” (MTP) has become popular (King-Smith and Vanderbilt 1993; Resta 1994, 1998) in the physics community. The MTP is based on a discretized Berry-phase treatment and has been applied to calculate the static polarization (dipole moment) and linear polarizability of many 3D materials, most notably piezoelectrics and ferroelectrics. It does not include frequency-dependence and, until quite recently, was limited to infinitesimal

fields, i.e., fields small enough so that they can be treated by perturbation theory (de Gironcoli et al. 1989; Baroni et al. 2001; van Gisbergen et al. 1998; Nunes and Gonze 2001).

An alternative approach has been developed by Kirtman and coworkers (Kirtman et al. 2000; Bishop et al. 2001). Instead of dealing with the scalar potential or (explicitly) with the Berry phase, these authors have developed a time-dependent vector potential approach (VPA) based on a suggestion made originally by Genkin and Mednis (1968) (GM). Taken together with the prescription for the polarization operator given by GM, as well as Blount (1962), this approach has the virtue of immediately satisfying the periodicity requirement and of introducing frequency-dependence in a straightforward manner. Through a detailed comparison between the VPA and MTP for static fields, it has been shown that the two approaches are closely related (Springborg and Kirtman 2008).

In this contribution, we shall review our own work that is closely related to the MTP and VPA methodologies. We shall start with the large, finite system that will act as a reference for the approaches where the system is treated as being infinite and periodic. By requiring that a thermodynamic limit exists for the former, we obtain information on the responses of the system of interest to electrostatic fields. In particular, we shall see that surfaces and shapes are relevant. Nevertheless, we shall subsequently see that when treating the systems as being infinite and periodic, the effects of the surfaces are included.

The single-particle (Hartree-Fock or Kohn-Sham) equations for the infinite periodic system exposed to an electrostatic field have a nonstandard form. We shall therefore discuss how they can be solved. Finally, we give a few exemplary applications of our theory.

The Gauge

For an electron in an external electromagnetic field, the time-dependent (Hartree-Fock or Kohn-Sham) single-particle equation takes the form

$$\left\{ \frac{1}{2m} \left[-i\hbar\nabla + \frac{e}{c} \mathbf{A}_{\text{ext,E}}(\mathbf{r}, t) \right]^2 + V_{\text{ext,n}}(\mathbf{r}) + V_J(\mathbf{r}) + \hat{V}_{\text{xc}}(\mathbf{r}, t) + V_{\text{ext,E}}(\mathbf{r}, t) \right\} \psi_i(\mathbf{x}, t) = i\hbar \frac{\partial}{\partial t} \psi_i(\mathbf{x}, t). \quad (1)$$

Here, m is the mass of the electron and c is the speed of light. V_J is the coulombic potential due to electron-electron repulsion including self-interaction. Within density-functional theory, exchange and correlation effects are contained in \hat{V}_{xc} , which is a functional of the (time-dependent) electron density, whereas only exchange interactions are included in the Hartree-Fock approximation. In that case, \hat{V}_{xc} becomes an operator that depends on the occupied orbitals. $V_{\text{ext,n}}$ describes the electrostatic nuclear-electron attraction. The effects of the electric field are described

via the vector potential $\mathbf{A}_{\text{ext,E}}$ and the scalar potential $V_{\text{ext,E}}$. In the general case, an electric field, \mathbf{E} , and a magnetic field, \mathbf{B} , are given through

$$\begin{aligned}\mathbf{E}(\mathbf{r}, t) &= -\nabla V_{\text{ext,E}}(\mathbf{r}, t) - \frac{1}{c} \frac{\partial}{\partial t} \mathbf{A}_{\text{ext,E}}(\mathbf{r}, t) \\ \mathbf{B}(\mathbf{r}, t) &= \nabla \times \mathbf{A}_{\text{ext,E}}(\mathbf{r}, t).\end{aligned}\quad (2)$$

In the present discussion, we shall not be concerned with magnetic fields, so that $\mathbf{B} = \mathbf{0}$, which can be satisfied by taking

$$\mathbf{A}_{\text{ext,E}}(\mathbf{r}, t) = \nabla f_{\text{ext}}(\mathbf{r}, t), \quad (3)$$

where $f_{\text{ext}}(\mathbf{r}, t)$ is some function. Even then, however, the separation of \mathbf{E} into scalar and vector potential contributions is not unique: different choices define different gauges.

An example that we shall discuss below in some detail is that of a homogeneous, electrostatic field given by a constant \mathbf{E} everywhere in space. This can be described, for example, through a pure scalar potential,

$$\begin{aligned}V_{\text{ext,E}}(\mathbf{r}, t) &= -\mathbf{E} \cdot \mathbf{r} \\ \mathbf{A}_{\text{ext,E}}(\mathbf{r}, t) &= \mathbf{0},\end{aligned}\quad (4)$$

or, equivalently, through a pure vector potential,

$$\begin{aligned}V_{\text{ext,E}}(\mathbf{r}, t) &= 0 \\ \mathbf{A}_{\text{ext,E}}(\mathbf{r}, t) &= -c\mathbf{E} \cdot t.\end{aligned}\quad (5)$$

Linear combinations of the two cases are possible, too, but here we shall only be concerned with these two choices. Equation (1) depends on which of the two forms in Eqs. (4) and (5) is used, but, of course, physical observables must be independent of that choice. Below we shall exploit both possibilities.

The Dipole Moment for the Finite System

For any finite system exposed to a homogeneous, electrostatic field, the Hamilton operator changes according to

$$\hat{H} \rightarrow \hat{H} - \hat{\boldsymbol{\mu}} \cdot \mathbf{E}. \quad (6)$$

Here, \hat{H} is the Hamilton operator in the absence of the field, and we have used the scalar interaction potential with \mathbf{E} as the electrostatic field. Finally, $\hat{\boldsymbol{\mu}}$ is the dipole moment operator for the system of interest.

The dipole moment property (i.e., the expectation value of the operator) can be written as a sum of a contribution from the nuclei and one from the electrons,

$$\boldsymbol{\mu} = \sum_p Z_p \mathbf{R}_p - \int \rho(\mathbf{r}) \mathbf{r} d\mathbf{r}. \quad (7)$$

As everywhere else in this contribution going forward, we have used atomic units, in which case $|e| = 1$. Moreover, Z_p and \mathbf{R}_p are the charge and position of the p th nucleus, whereas $\rho(\mathbf{r})$ is the electron density.

As mentioned above, in the present study, we shall focus on systems that, besides being very large, are regular in the sense that they contain a very large number of identical units and only in the surface regions may one observe deviations from regularity. In that event, the thermodynamic limit has been reached. It is customary to approximate such systems as if they were infinite and periodic.

Quasi-One-Dimensional Systems

For the sake of simplicity, we shall start with a discussion of quasi-one-dimensional systems (chains) that are assumed to lie along the z axis. Then, the (intensive) property per unit corresponding to the extensive property $\xi(N)$, in which N is the number of units, may be determined as

$$\bar{\xi} = \lim_{N \rightarrow \infty} \frac{\xi(N)}{N} = \lim_{N \rightarrow \infty} \frac{1}{\Delta N} [\xi(N + \Delta N) - \xi(N)]. \quad (8)$$

Figure 3 shows a schematic representation of a long, regular chain. We see that the chain consists of a number of repeated units in the so-called central region marked by vertical blue lines). In this region, the presence of the boundaries/terminations is not felt. That is different for the terminal regions. The separation into a central region and terminations is not unique as can be seen by comparing the two topmost panels in Fig. 3. Nevertheless, the units of the central region are neutral regardless of the precise definition of those units. That this is so can be understood as follows. Imagine that these cells were not neutral but each carries some charge, q . When increasing the size of the system, the central units will, per construction, remain unchanged. Moreover, the charges in the terminal regions are required to stay unchanged when the system is sufficiently large. Then the total charge of the system would increase, which is in contrast to the assumption that we have reached the thermodynamic limit. Thus, in the thermodynamic limit $q = 0$.

If, in addition, the chain is neutral, then, the charges accumulated at the right and the left terminal regions must add to zero,

$$q_L = -q_R. \quad (9)$$

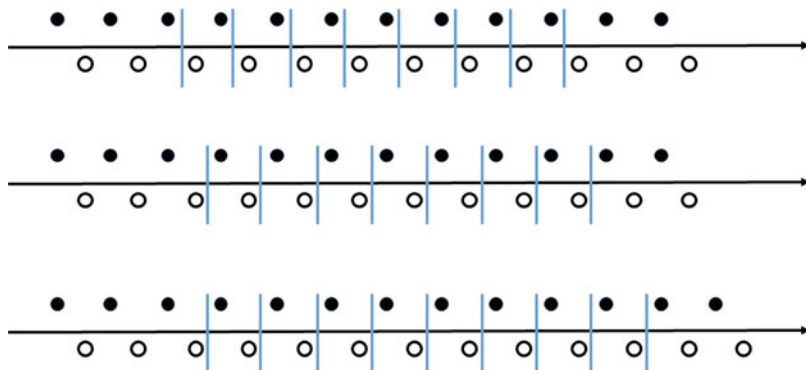


Fig. 3 Schematic representation of a long, regular chain. The two topmost panels differ in the way the repeated unit is defined (marked through the vertical lines), whereas the two bottommost panels differ in the length of the chain

We shall now discuss the dipole moment for the large chain with the ultimate goal of obtaining expressions for the dipole moment per unit using the last identity in Eq. (8). The total dipole moment of the system in the middle panel of Fig. 3 can be written as

$$\begin{aligned} \boldsymbol{\mu} &= \int \rho_{\text{tot}}(\mathbf{r})\mathbf{r}d\mathbf{r} = \int_C \rho_{\text{tot}}(\mathbf{r})\mathbf{r}d\mathbf{r} + \int_L \rho_{\text{tot}}(\mathbf{r})\mathbf{r}d\mathbf{r} + \int_R \rho_{\text{tot}}(\mathbf{r})\mathbf{r}d\mathbf{r} \\ &= K_C \boldsymbol{\mu}_C + (\mathbf{R}_R - \mathbf{R}_L) \int_R \rho_{\text{tot}}(\mathbf{r})d\mathbf{r} + \int_L \rho_{\text{tot}}(\mathbf{r})(\mathbf{r} - \mathbf{R}_L)d\mathbf{r} + \int_R \rho_{\text{tot}}(\mathbf{r})(\mathbf{r} - \mathbf{R}_R)d\mathbf{r}, \end{aligned} \quad (10)$$

where we have made explicit use of the spatial separation into right (R), left (L), and central (C) regions and used the neutrality of the central units as well as the overall neutrality. In Eq. (10), $\rho_{\text{tot}}(\mathbf{r})$ is the total charge density; \mathbf{R}_R and \mathbf{R}_L are typical positions within the right and left regions, respectively; and K_C is the number of units in C . Combining this relation with the second definition of Eq. (8) gives immediately (cf. the two lowest panels in Fig. 3)

$$\bar{\boldsymbol{\mu}} = \boldsymbol{\mu}_C + q_R \cdot \mathbf{a}, \quad (11)$$

in which q_R is the total charge of R , $\boldsymbol{\mu}_C$ is the dipole moment of one of the C units, and \mathbf{a} is the lattice constant of C . That is to say, upon increasing the size of the system, one contribution to the dipole moment per unit comes from the increased number of central units and another from the increased distance between the charges at the terminations.

It is clear that $\bar{\boldsymbol{\mu}}$ depends on the charge accumulated at the terminations. Thus, it might be thought that upon chemical substitution at the terminations, q_R can be varied essentially arbitrarily (within chemical limits), thereby making all values of $\bar{\boldsymbol{\mu}}$

possible. However, that is not the case as may be seen from the following argument (Vanderbilt and King-Smith 1993; Kudin et al. 2007).

The first step is to transform the occupied orbitals into a set of maximally localized orthogonal functions, $\{w_{np}(\mathbf{r})\}$, which in C will approach Wannier functions as the system becomes larger. Here, w_{np} is the p th function localized primarily to the n th cell (in the terminal regions, the division into cells is arbitrary) which, however, may have tails sticking out into other cells. We can, then, write $\underline{\mu}$ as

$$\begin{aligned}
 \mu &= \sum_n \left[-\sum_p \int |w_{np}(\mathbf{r})|^2 \mathbf{r} d\mathbf{r} + \sum_m \mathbf{R}_{nm} Z_m \right] \\
 &= \sum_{n \in C} \left[-\sum_p \int |w_{np}(\mathbf{r})|^2 \mathbf{r} d\mathbf{r} + \sum_m \mathbf{R}_{nm} Z_m \right] \\
 &\quad + \mathbf{R}_L \sum_{n \in L} \left[-\sum_p \int |w_{np}(\mathbf{r})|^2 d\mathbf{r} + \sum_m Z_m \right] \\
 &\quad + \sum_{n \in L} \left[-\sum_p \int |w_{np}(\mathbf{r})|^2 (\mathbf{r} - \mathbf{R}_L) d\mathbf{r} + \sum_m Z_m (\mathbf{R}_{nm} - \mathbf{R}_L) \right] \\
 &\quad + \mathbf{R}_R \sum_{n \in R} \left[-\sum_p \int |w_{np}(\mathbf{r})|^2 d\mathbf{r} + \sum_m Z_m \right] \\
 &\quad + \sum_{n \in R} \left[-\sum_p \int |w_{np}(\mathbf{r})|^2 (\mathbf{r} - \mathbf{R}_R) d\mathbf{r} + \sum_m Z_m (\mathbf{R}_{nm} - \mathbf{R}_R) \right] \quad (12)
 \end{aligned}$$

where Z_m is the charge of the m th nucleus (independent of n) and \mathbf{R}_{nm} is the position of the m th nucleus in the n th cell. In Eq. (12), all integrals are over the complete space, but the orbital densities are localized to different regions, whereas in Eq. (10), the integrals are over different regions, but the (electron) densities are completely delocalized. In the last identity in Eq. (12), only the first, second, and fourth terms depend (linearly) on the size of the system, and therefore only these terms will contribute to the long-chain limit of the dipole moment per unit.

Using the w functions, the density matrix for the occupied orbitals can be written in block form,

$$\underline{\underline{D}} = \begin{pmatrix} \underline{\underline{D}}_L & \underline{\underline{0}} & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{D}}_C & \underline{\underline{0}} \\ \underline{\underline{0}} & \underline{\underline{0}} & \underline{\underline{D}}_R \end{pmatrix} \quad (13)$$

where each block $\underline{\underline{D}}_X$ is the density matrix for the region X and the submatrices $\underline{\underline{0}}$ contain elements that are exponentially vanishing as a function of the distance from

the boundary between the regions. As the w functions become closer to Wannier functions in C , \underline{D}_C will approach a diagonal matrix. This density matrix for the occupied orbitals, \underline{D} , is idempotent, so that each block \underline{D}_X is idempotent as well. Therefore, in the formulation of Eq. (12), the w functions centered in the terminal regions are associated with an integral number of electrons.

We, finally, add that the separation of the dipole moment per unit into a bulk part and a contribution from the terminations, Eq. (11), is not unique. The two uppermost panels in Fig. 3 show the same system but with different definitions of the unit cells in the central region. Thereby, different values for μ_C result. However, also the precise definition of the terminations is affected by the definition of the bulk unit cells, and, accordingly, the charge we associate with those, q_R , changes, too. Ultimately, the total dipole moment per unit, $\bar{\mu}$, is unchanged but the individual components are not.

Higher Dimensions

For quasi-two-dimensional and three-dimensional systems, we can use an approach very similar to that presented above with, however, a couple of important modifications.

In Fig. 4 we show a schematic representation of a two-dimensional system with a rectangular shape. We have split the system into a central region, side regions, and corner regions. The unit cells of each of the side regions will all be equivalent but in most cases larger than those of the central region. This can happen because of surface reconstructions [which can make the surface units larger both in direction(s) parallel to the surface and in the direction perpendicular to the surface] as well as

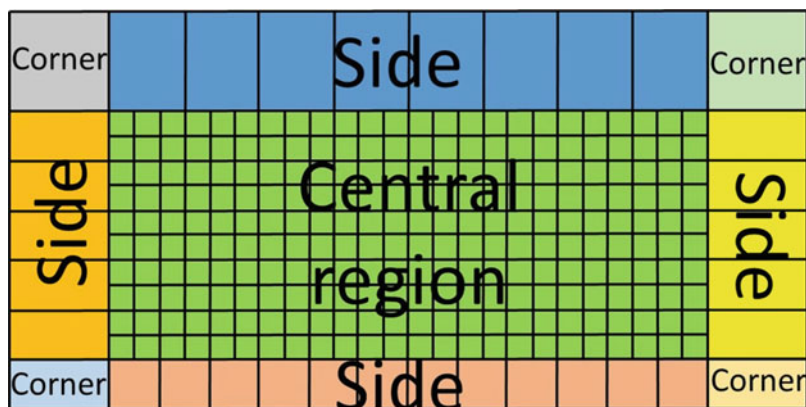


Fig. 4 Schematic representation of a two-dimensional system

the vacuum discontinuity (which will make the surface units larger/thicker in the direction perpendicular to the surface).

As above, the units of the central region are per construction neutral. This is not the case for either the side or the corner units, although each unit carries the same charge. With N_r being the number of surface units on the r th side and q_r being the charge of each of those units, charge neutrality gives that

$$\sum_r N_r q_r = 0, \tag{14}$$

where the r summation runs over all surfaces. This condition has the interesting consequence that two equivalent surfaces of two systems of the same material but differing in the shape may have different q_r values, i.e., the overall shape of the system dictates the possible surface charges on the individual sides.

In order to calculate the dipole moment per (bulk-size) unit, we proceed as in the previous subsection by studying the change in the dipole moment as the size of the system is increased. A general situation for a (quasi)-two-dimensional system is shown in Fig. 5. When the system is made larger, the r th surface is displaced the fraction f_r along the vector shown in the figure. This vector can be written as

$$\Delta \mathbf{R}_r = n_{ra} \cdot \mathbf{a} + n_{rb} \cdot \mathbf{b} \tag{15}$$

with \mathbf{a} and \mathbf{b} being the (bulk) lattice vectors and n_{ra} and n_{rb} being integers that describe the orientation of the r th side. The factors f_r are chosen so that the overall shape of the system remains unchanged. Through the displacement,

$$\Delta N_r = f_r N_r (n_{ra}^2 + n_{rb}^2) \tag{16}$$

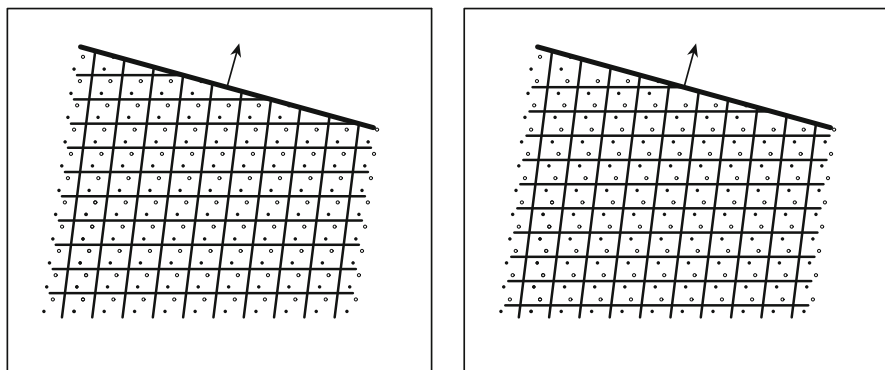


Fig. 5 Schematic representation of the surface of a two-dimensional system. The two different panels differ in the definition of the bulk unit cell

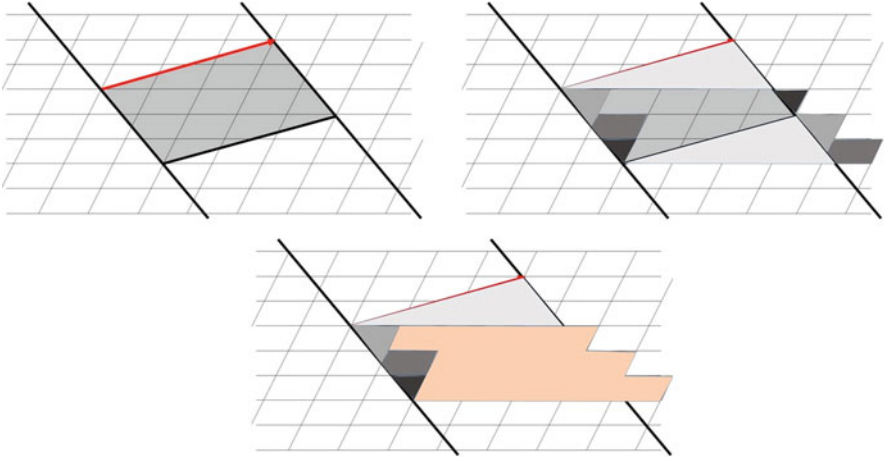


Fig. 6 A detailed illustration of how the number of unit cells in a two-dimensional system changes when the surface is displaced. The surface is characterized by $(n_{sa}, n_{sb}) = (3, 2)$. This vector is shown in red in the upper, left panel. In this panel, we also show (in grey) the ΔN_s unit cells that are added when the surface is displaced. In the upper, right panel, we have marked with different grey tones different parts that are replaced in order to facilitate the determination of ΔN_s as demonstrated in the lower panel

bulk units have been added to the system. In Fig. 6, we illustrate this through an example for which $f_r = 1$. Similarly, the total dipole moment has changed by

$$\Delta \boldsymbol{\mu}_r = f_r N_r q_r \quad \Delta \mathbf{R}_r = f_r N_r q_r (n_{ra} \mathbf{a} + n_{rb} \mathbf{b}). \tag{17}$$

The contribution to the dipole moment per unit from all surfaces is found by summing this expression over all surfaces, and dividing this by the total number of additional units, i.e., the sum of the terms in Eq. (16) and by also including the dipole moment for a single central unit, we have the total dipole moment per unit,

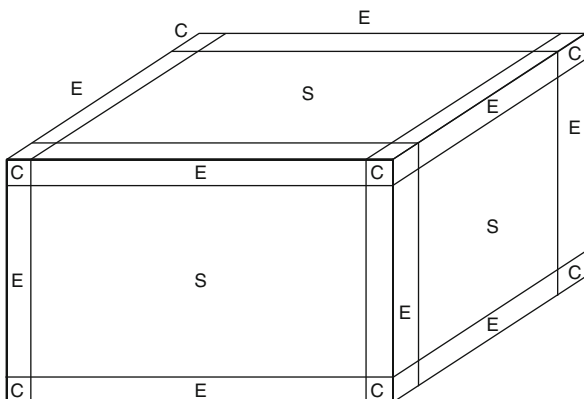
$$\bar{\boldsymbol{\mu}} = \boldsymbol{\mu}_C + \frac{\sum_r f_r N_r q_r (n_{ra} \mathbf{a} + n_{rb} \mathbf{b})}{\sum_r f_r N_r (n_{ra}^2 + n_{rb}^2)}. \tag{18}$$

For three-dimensional systems, the arguments above can be repeated with only minor modifications. Thus, the system will be split into a central region as well as surface, edge, and corner regions (see Fig. 7). Of those, only the central region (whose units are neutral) as well as the charges on the surfaces will contribute to the dipole moment per unit,

$$\bar{\boldsymbol{\mu}} = \boldsymbol{\mu}_C + \frac{\sum_r f_r N_r q_r (n_{ra} \mathbf{a} + n_{rb} \mathbf{b} + n_{rc} \mathbf{c})}{\sum_r f_r N_r (n_{ra}^2 + n_{rb}^2 + n_{rc}^2)}, \tag{19}$$

where now three integers are needed to describe the orientation of each surface. In this case as well, the charges q_r can take any value as long as Eq. (14) is satisfied.

Fig. 7 The separation of a 3D rectangular sample into surface (*S*), edge (*E*), and corner (*C*) regions. The central region is not shown



The Dipole Moment for the Periodic Case

Despite some obvious difficulties, it is of interest to consider whether the systems discussed above can be approximated as being infinite and periodic when subject to an external electrostatic field. Per construction, such systems do not possess any surfaces. However, as we have seen, the dipole moment contains a non-negligible contribution from the surfaces. Moreover, such fields destroy the translational invariance making it even more questionable whether the translational symmetry can be of any use.

We shall, however, see that it indeed is possible to reintroduce the translational invariance and that the effects of the surfaces are included in the calculations. The approach we shall use is not trivial and shall accordingly be discussed in some detail.

In the absence of any field, the translational symmetry can be used in classifying the single-particle orbitals according to their symmetry properties upon translation, i.e., to write them as Bloch functions,

$$\psi_j(\mathbf{r}) = \psi_j(\mathbf{k}, \mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_j(\mathbf{k}, \mathbf{r}), \tag{20}$$

where $u_j(\mathbf{k}, \mathbf{r})$ is a lattice-periodic function and where j is used in distinguishing different functions of the same \mathbf{k} . \mathbf{k} lies in the first Brillouin zone, but in a practical calculation one considers only a finite, discrete set of K \mathbf{k} values including $\mathbf{k} = \mathbf{0}$. K unit cells form the Born-von Kármán zone.

Often the orbitals of Eq. (20) are expanded in a set of atom-centered functions,

$$\begin{aligned} \psi_j(\mathbf{k}, \mathbf{r}) &= \sum_l c_{lj}(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r}) \\ \chi_l(\mathbf{k}, \mathbf{r}) &= \lim_{N \rightarrow \infty} \frac{1}{\sqrt{N}} \sum_{\mathbf{n}} e^{i\mathbf{k}\cdot\mathbf{R}_n} \chi_{l,\mathbf{n}}(\mathbf{r}). \end{aligned} \tag{21}$$

Here, $\chi_{l,\mathbf{n}}(\mathbf{r})$ is the l th atom-centered function of the unit cell at lattice point $\mathbf{R}_{\mathbf{n}}$, and N is the number of lattice points in the \mathbf{n} summation.

The Single-Electron Case

In order to obtain an expression for the position operator, we start with a single electron occupying an orbital of the form of Eq. (20). As observed by Blount (1962), the effect of the position operator acting on this orbital may be represented through the relation

$$\mathbf{r} [e^{i\mathbf{k}\cdot\mathbf{r}} u_j(\mathbf{k}, \mathbf{r})] = [i e^{i\mathbf{k}\cdot\mathbf{r}} \nabla_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}} - i \nabla_{\mathbf{k}}] [e^{i\mathbf{k}\cdot\mathbf{r}} u_j(\mathbf{k}, \mathbf{r})], \quad (22)$$

where $\nabla_{\mathbf{k}}$ is the gradient operator in \mathbf{k} space. However, in a practical calculation, we have only a discrete set of \mathbf{k} vectors at our disposal and will accordingly replace the \mathbf{k} derivatives with finite-difference approximations. Moreover, we shall consider a general function that is expanded in terms of the Bloch functions [ignoring an index j as in Eq. (21)],

$$\psi(\mathbf{r}) = \sum_l \sum_{\mathbf{k}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r}), \quad (23)$$

which corresponds to the Crystal Momentum Representation of Blount (1962). As Blount notes, the effect of using Eq. (23) for the position operator acting on a general ψ is given by

$$\begin{aligned} \mathbf{r}\psi(\mathbf{r}) &= \frac{1}{(2\pi)^3} \int e^{i\mathbf{k}\cdot\mathbf{r}} \mathbf{r} \sum_l c_l(\mathbf{k}) u_l(\mathbf{k}, \mathbf{r}) d\mathbf{k} \\ &= \frac{1}{(2\pi)^3} \int (-i) \sum_l \nabla_{\mathbf{k}} [c_l(\mathbf{k}) \psi_l(\mathbf{k}, \mathbf{r})] d\mathbf{k} \\ &\quad + \frac{1}{(2\pi)^3} \int i e^{i\mathbf{k}\cdot\mathbf{r}} \sum_l \nabla_{\mathbf{k}} [c_l(\mathbf{k}) u_l(\mathbf{k}, \mathbf{r})] d\mathbf{k}. \end{aligned} \quad (24)$$

It is convenient to take components along crystallographic axes. These may readily be transformed when necessary to obtain components along orthogonal directions. Then, if \hat{s}_1 refers to the finite-difference expression for the $i \nabla_{\mathbf{k}}$ term, we have

$$\begin{aligned} \hat{s}_1 \psi(\mathbf{r}) &= \hat{s}_1 \left[\sum_l \sum_{\mathbf{k}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r}) \right] = \sum_l \sum_{\mathbf{k}} \hat{s}_1 [c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r})] \\ &= i \sum_l \sum_{\mathbf{k}} \frac{c_l(\mathbf{k} + \Delta\mathbf{k}_s) \chi_l(\mathbf{k} + \Delta\mathbf{k}_s, \mathbf{r}) - c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r})}{\Delta k_s} \end{aligned} \quad (25)$$

where

$$\Delta k_s = \frac{2\pi}{a_s K_s} \quad (26)$$

is the spacing in \mathbf{k} space, the subscript s refers to the crystallographic direction, and a_s is the length of the lattice vector in direct space.

The finite-difference expression for the remaining gradient term may be written as

$$\begin{aligned} \hat{s}_2 \psi(\mathbf{r}) &= \hat{s}_2 \left[\sum_l \sum_{\mathbf{k}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r}) \right] = \sum_l \sum_{\mathbf{k}} \hat{s}_2 [c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r})] \\ &= i \sum_l \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \frac{e^{-i(\mathbf{k}+\Delta\mathbf{k}_s)\cdot\mathbf{r}} c_l(\mathbf{k}+\Delta\mathbf{k}_s) \chi_l(\mathbf{k}+\Delta\mathbf{k}_s, \mathbf{r}) - e^{-i\mathbf{k}\cdot\mathbf{r}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r})}{\Delta k_s}. \end{aligned} \quad (27)$$

It is now easy to show that

$$\begin{aligned} \hat{s}_1 \psi(\mathbf{r}) &= i \sum_l \sum_{\mathbf{k}} \frac{c_l(\mathbf{k} + \Delta\mathbf{k}_s) \chi_l(\mathbf{k} + \Delta\mathbf{k}_s, \mathbf{r}) - c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r})}{\Delta k_s} \\ &= i \frac{1}{\Delta k_s} \left[\sum_l \sum_{\mathbf{k}} c_l(\mathbf{k} + \Delta\mathbf{k}_s) \chi_l(\mathbf{k} + \Delta\mathbf{k}_s, \mathbf{r}) - \sum_l \sum_{\mathbf{k}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r}) \right] \\ &= i \frac{1}{\Delta k_s} \left[\sum_l \sum_{\mathbf{k}'} c_l(\mathbf{k}') \chi_l(\mathbf{k}', \mathbf{r}) - \sum_l \sum_{\mathbf{k}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r}) \right] \\ &= 0 \end{aligned} \quad (28)$$

since the \mathbf{k} and \mathbf{k}' summations go over all \mathbf{k} in the first Brillouin zone and we have cyclic boundary conditions in \mathbf{k} space. Similarly,

$$\begin{aligned} \hat{s}_2 \psi(\mathbf{r}) &= i \sum_l \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \frac{e^{-i(\mathbf{k}+\Delta\mathbf{k}_s)\cdot\mathbf{r}} c_l(\mathbf{k}+\Delta\mathbf{k}_s) \chi_l(\mathbf{k}+\Delta\mathbf{k}_s, \mathbf{r}) - e^{-i\mathbf{k}\cdot\mathbf{r}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r})}{\Delta k_s} \\ &= i \frac{1}{\Delta k_s} \left[\sum_l \sum_{\mathbf{k}} e^{-i\Delta\mathbf{k}_s\cdot\mathbf{r}} c_l(\mathbf{k}+\Delta\mathbf{k}_s) \chi_l(\mathbf{k}+\Delta\mathbf{k}_s, \mathbf{r}) - \sum_l \sum_{\mathbf{k}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r}) \right] \\ &= i \frac{1}{\Delta k_s} \left[e^{-i\Delta\mathbf{k}_s\cdot\mathbf{r}} \sum_l \sum_{\mathbf{k}} c_l(\mathbf{k}+\Delta\mathbf{k}_s) \chi_l(\mathbf{k}+\Delta\mathbf{k}_s, \mathbf{r}) - \sum_l \sum_{\mathbf{k}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r}) \right] \end{aligned}$$

$$\begin{aligned}
&= i \frac{1}{\Delta k_s} \left[e^{-i\Delta\mathbf{k}_s \cdot \mathbf{r}} \sum_l \sum_{\mathbf{k}'} c_l(\mathbf{k}') \chi_l(\mathbf{k}', \mathbf{r}) - \sum_l \sum_{\mathbf{k}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r}) \right] \\
&= i \frac{1}{\Delta k_s} [e^{-i\Delta\mathbf{k}_s \cdot \mathbf{r}} - 1] \sum_l \sum_{\mathbf{k}} c_l(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{r}) \\
&= \frac{i}{\Delta k_s} [e^{-i\Delta\mathbf{k}_s \cdot \mathbf{r}} - 1] \psi(\mathbf{r}). \tag{29}
\end{aligned}$$

The Many-Electron Case

We consider a system with

$$N = K \cdot N_e \tag{30}$$

electrons per Born-von Kármán (BvK) zone. Here, K is the number of unit cells per BvK zone, and N_e is the number of electrons per unit cell. The N -electron wave function is written as a single Slater determinant

$$\begin{aligned}
\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \dots & \psi_N(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \dots & \psi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \dots & \psi_N(\mathbf{x}_N) \end{vmatrix} \\
&= \frac{1}{\sqrt{N!}} \sum_P (-1)^P \psi_{l_1}(\mathbf{x}_1) \psi_{l_2}(\mathbf{x}_2) \dots \psi_{l_N}(\mathbf{x}_N) \\
&\equiv \hat{\mathcal{A}} [\psi_1(\mathbf{x}_1) \psi_2(\mathbf{x}_2) \dots \psi_N(\mathbf{x}_N)], \tag{31}
\end{aligned}$$

where \mathbf{x}_i is the usual combined position and spin coordinate and $\hat{\mathcal{A}}$ is the antisymmetrizing operator. Finally, the summation is over all permutations of the N indices of the single-particle orbitals, and P is the order of the permutation (either even or odd).

Each of the orbitals $\psi_j(\mathbf{x})$ is expanded in the Bloch functions,

$$\psi_j(\mathbf{x}) = \sum_l \sum_{\mathbf{k}} c_{lj}(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{x}). \tag{32}$$

Then, the terms in Eq. (31) can be written as

$$\begin{aligned}
\psi_{j_1}(\mathbf{x}_1) \psi_{j_2}(\mathbf{x}_2) \dots \psi_{j_N}(\mathbf{x}_N) &= \prod_n \psi_{j_n}(\mathbf{x}_n) \\
&= \prod_n \sum_{l_n} \sum_{\mathbf{k}_n} c_{l_n j_n}(\mathbf{k}_n) \chi_{l_n}(\mathbf{k}_n, \mathbf{x}_n). \tag{33}
\end{aligned}$$

Based on the same arguments as in the previous subsection, we define the N -electron operators \hat{S}_1 and \hat{S}_2 through

$$\begin{aligned} \hat{S}_1[\psi_{j_1}(\mathbf{x}_1)\psi_{j_2}(\mathbf{x}_2)\cdots\psi_{j_N}(\mathbf{x}_N)] &= \frac{i}{\Delta k_s} \left[\prod_n \sum_{l_n} \sum_{\mathbf{k}_n} c_{l_n j_n}(\mathbf{k}_n + \Delta \mathbf{k}_s) \chi_{l_n}(\mathbf{k}_n + \Delta \mathbf{k}_s, \mathbf{x}_n) \right. \\ &\quad \left. - i \prod_n \sum_{l_n} \sum_{\mathbf{k}_n} c_{l_n j_n}(\mathbf{k}_n) \chi_{l_n}(\mathbf{k}_n, \mathbf{x}_n) \right] = 0 \end{aligned} \quad (34)$$

and

$$\begin{aligned} &\hat{S}_2[\psi_{j_1}(\mathbf{x}_1)\psi_{j_2}(\mathbf{x}_2)\cdots\psi_{j_N}(\mathbf{x}_N)] \\ &= \frac{i}{\Delta k_s} \left[\exp\left(i \sum_n \mathbf{k}_n \cdot \mathbf{r}_n\right) \prod_n \sum_{l_n} \sum_{\mathbf{k}_n} \exp\left(-i \sum_n (\mathbf{k}_n + \Delta \mathbf{k}_s) \cdot \mathbf{r}_n\right) \right. \\ &\quad \left. \times c_{l_n j_n}(\mathbf{k}_n + \Delta \mathbf{k}_s) \chi_{l_n}(\mathbf{k}_n + \Delta \mathbf{k}_s, \mathbf{x}_n) \right. \\ &\quad \left. - i \exp\left(i \sum_n \mathbf{k}_n \cdot \mathbf{r}_n\right) \prod_n \sum_{l_n} \sum_{\mathbf{k}_n} \exp\left(-i \sum_n \mathbf{k}_n \cdot \mathbf{r}_n\right) c_{l_n j_n}(\mathbf{k}_n) \chi_{l_n}(\mathbf{k}_n, \mathbf{x}_n) \right] \\ &= \frac{i}{\Delta k_s} \left[\exp\left(-i \Delta \mathbf{k}_s \cdot \sum_n \mathbf{r}_n\right) - 1 \right] [\psi_{j_1}(\mathbf{x}_1)\psi_{j_2}(\mathbf{x}_2)\cdots\psi_{j_N}(\mathbf{x}_N)]. \end{aligned} \quad (35)$$

Thus,

$$\hat{S}_2\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{i}{\Delta k_s} \left[\exp\left(-i \Delta \mathbf{k}_s \cdot \sum_n \mathbf{r}_n\right) - 1 \right] \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N), \quad (36)$$

which means that

$$\begin{aligned} &\langle \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) | \hat{S}_2 | \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \rangle \\ &= \frac{1}{N!} \frac{i}{\Delta k_s} \sum_{P_1, P_2} (-1)^{P_1+P_2} \langle \psi_{m_1}(\mathbf{x}_1) | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}_1} | \psi_{j_1}(\mathbf{x}_1) \rangle \langle \psi_{m_2}(\mathbf{x}_2) | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}_2} | \psi_{j_2}(\mathbf{x}_2) \rangle \\ &\quad \cdots \langle \psi_{m_N}(\mathbf{x}_N) | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}_N} | \psi_{j_N}(\mathbf{x}_N) \rangle - \frac{i}{\Delta k_s} \\ &= \frac{i}{\Delta k_s} \sum_P (-1)^P \langle \psi_1(\mathbf{x}_1) | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}_1} | \psi_{j_1}(\mathbf{x}_1) \rangle \langle \psi_2(\mathbf{x}_2) | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}_2} | \psi_{j_2}(\mathbf{x}_2) \rangle \\ &\quad \cdots \langle \psi_N(\mathbf{x}_N) | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}_N} | \psi_{j_N}(\mathbf{x}_N) \rangle - \frac{i}{\Delta k_s} \end{aligned}$$

$$\begin{aligned}
 &= \frac{i}{\Delta k_s} \begin{vmatrix} \langle \psi_1 | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}} | \psi_1 \rangle & \langle \psi_1 | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}} | \psi_2 \rangle & \dots & \langle \psi_1 | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}} | \psi_N \rangle \\ \langle \psi_2 | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}} | \psi_1 \rangle & \langle \psi_2 | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}} | \psi_2 \rangle & \dots & \langle \psi_2 | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}} | \psi_N \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \psi_N | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}} | \psi_1 \rangle & \langle \psi_N | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}} | \psi_2 \rangle & \dots & \langle \psi_N | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}} | \psi_N \rangle \end{vmatrix} - \frac{i}{\Delta k_s} \\
 &\equiv \frac{i}{\Delta k_s} \left[\det \underline{S}_s^- - 1 \right]. \tag{37}
 \end{aligned}$$

The index s on \underline{S}_s^- specifies the direction in \mathbf{k} space so that \underline{S}_s^- can be interpreted as one of the components of a vector in three-dimensional space. Moreover, in the second identity, we have used the fact that one of the two summations P_1 and P_2 corresponds to interchanging the $\{\psi_{m_i}\}$ orbitals which is equivalent to interchanging the $\{x_i\}$ arguments of the orbitals. However, since these are “dummy” integration variables, the different permutations give the same contributions, and we can, accordingly, remove one of the summations when we simultaneously multiply with the number of such permutations, $N!$

We now make the approximation (Springborg et al. 2004) which is good in the limit $\Delta k_s \rightarrow 0$,

$$\begin{aligned}
 \frac{i}{\Delta k_s} \det \underline{S}_s^- &= \langle \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) | \frac{i}{\Delta k_s} \exp \left(-i \Delta \mathbf{k}_s \cdot \sum_n \mathbf{r}_n \right) | \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \rangle \\
 &\simeq \frac{i}{\Delta k_s} \exp \left(\langle \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) | -i \Delta \mathbf{k}_s \cdot \sum_n \mathbf{r}_n | \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \rangle \right), \tag{38}
 \end{aligned}$$

so that for the position operator per cell, we find

$$\begin{aligned}
 \frac{1}{K} \langle \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) | \frac{\Delta \mathbf{k}_s}{\Delta k_s} \cdot \sum_n \mathbf{r}_n | \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \rangle &\simeq \frac{i}{K \Delta k_s} \ln \det \underline{S}_s^- \\
 &= -\frac{1}{K \Delta k_s} \text{Im} \ln \det \underline{S}_s^-, \tag{39}
 \end{aligned}$$

since the expression on the left-hand side must be real.

A special case arises when the single-particle orbitals are Bloch functions with a unique \mathbf{k} , cf. Eq. (20), whereby

$$\begin{aligned}
 \langle \psi_{j_1}(\mathbf{k}_1, \mathbf{x}) | e^{-i \Delta \mathbf{k}_s \cdot \mathbf{r}} | \psi_{j_2}(\mathbf{k}_2, \mathbf{x}) \rangle &= \delta_{\mathbf{k}_1, \mathbf{k}_2 - \Delta \mathbf{k}_s} \langle u_{j_1}(\mathbf{k}_1, \mathbf{x}) | u_{j_2}(\mathbf{k}_2, \mathbf{x}) \rangle \\
 &\simeq \langle u_{j_1}(\mathbf{k}_1, \mathbf{x}) | u_{j_2}(\mathbf{k}_1, \mathbf{x}) \rangle + \Delta \mathbf{k}_s \cdot \nabla_k u_{j_2}(\mathbf{k}_1, \mathbf{x}) \\
 &= \delta_{j_1, j_2} + \langle u_{j_1}(\mathbf{k}_1, \mathbf{x}) | \Delta \mathbf{k}_s \cdot \nabla_k u_{j_2}(\mathbf{k}_1, \mathbf{x}) \rangle. \tag{40}
 \end{aligned}$$

Here, we have kept only the lowest-order terms in Δk_s and used the orthonormality of the Bloch waves.

We now insert the approximation of Eq. (40) into the expression for $\det \underline{S}_s^-$. Then, when evaluating $\det \underline{S}_s^-$, one obtains a sum of $N!$ terms, each involving the product of N factors of the type of Eq. (40). Upon keeping only terms that are at most linear in Δk_s , we see that only the product of diagonal terms survives. Then,

$$\det \underline{S}_s^- \simeq 1 + \sum_j \sum_{\mathbf{k}} \Delta \mathbf{k}_s \cdot \langle u_j(\mathbf{k}, \mathbf{x}) | \nabla_{\mathbf{k}} u_j(\mathbf{k}, \mathbf{x}) \rangle, \quad (41)$$

so that

$$\begin{aligned} & \frac{1}{K} \langle \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) | \frac{\Delta \mathbf{k}_s}{\Delta k_s} \cdot \sum_n \mathbf{r}_n | \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \rangle \\ & \simeq -\frac{1}{K \Delta k_s} \text{Im} \ln \det \underline{S}_s^- \\ & \simeq \frac{i}{K \Delta k_s} \ln \left[1 + \sum_j \sum_{\mathbf{k}} \Delta \mathbf{k}_s \cdot \langle u_j(\mathbf{k}, \mathbf{x}) | \nabla_{\mathbf{k}} u_j(\mathbf{k}, \mathbf{x}) \rangle \right] \\ & \simeq \frac{i}{K \Delta k_s} \sum_j \sum_{\mathbf{k}} \Delta \mathbf{k}_s \cdot \langle u_j(\mathbf{k}, \mathbf{x}) | \nabla_{\mathbf{k}} u_j(\mathbf{k}, \mathbf{x}) \rangle \\ & = \frac{i}{K} \sum_j \sum_{\mathbf{k}} \frac{\Delta \mathbf{k}_s}{\Delta k_s} \cdot \langle u_j(\mathbf{k}, \mathbf{x}) | \nabla_{\mathbf{k}} u_j(\mathbf{k}, \mathbf{x}) \rangle. \end{aligned} \quad (42)$$

This last expression can also be used in obtaining the energy due to an electrostatic field,

$$\begin{aligned} & \frac{1}{K} \langle \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) | \mathbf{E} \cdot \sum_n \mathbf{r}_n | \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \rangle \\ & = \frac{i}{K} \sum_j \sum_{\mathbf{k}} \mathbf{E} \cdot \langle u_j(\mathbf{k}, \mathbf{x}) | \nabla_{\mathbf{k}} u_j(\mathbf{k}, \mathbf{x}) \rangle. \end{aligned} \quad (43)$$

The expression of Eq. (39) was originally proposed by Resta (1994), whereas that of Eq. (42) was proposed by Vanderbilt and King-Smith (1993) (King-Smith and Vanderbilt 1993). These expressions, which are at the roots of the so-called modern theory of polarization, were originally given without derivation. More recently, we have presented the mathematical derivations outlined above (Springborg et al. 2004).

Quasi-One-Dimensional Systems

It is useful to illustrate the findings above on quasi-one-dimensional systems, i.e., chains. However, as we shall see, there is an important difference between quasi-one-dimensional and higher-dimensional systems. We assume that the chain is

placed along the z axis and consider accordingly the z component of the dipole moment per unit. This is separated into an electronic and a nuclear contribution,

$$\bar{\mu}_z = \bar{\mu}_{zn} + \bar{\mu}_{ze}. \quad (44)$$

As demonstrated above, within a single-particle (Hartree-Fock or Kohn-Sham) picture, the electronic contribution can be written either as (Resta 1994)

$$\bar{\mu}_{ze} = \bar{\mu}_R \equiv -\frac{a}{2\pi} \text{Im} \ln \det \underline{\underline{S}}^+ = \frac{a}{2\pi} \text{Im} \ln \det \underline{\underline{S}}^- \quad (45)$$

or as (King-Smith and Vanderbilt 1993; Vanderbilt and King-Smith 1993)

$$\begin{aligned} \bar{\mu}_{ze} = \bar{\mu}_{\text{KSV}} &\equiv -\frac{i}{K} \sum_{k=1}^K \sum_{j=1}^B \langle u_j(k, \mathbf{r}) | \frac{\partial}{\partial k} u_j(k, \mathbf{r}) \rangle \\ &= -\frac{i}{K} \sum_{k=1}^K \sum_{j=1}^B \langle e^{-ikz} \psi_j(k, \mathbf{r}) | \frac{\partial}{\partial k} e^{-ikz} \psi_j(k, \mathbf{r}) \rangle, \end{aligned} \quad (46)$$

where B is the number of occupied bands and spin degeneracy is not assumed [notice the difference in the sign compared to Eq. (42) which is due to the negative electronic charge]. These expressions are valid only for systems with an energy gap between occupied and unoccupied orbitals, i.e., semiconductors and insulators.

In Eq. (45), we have two different, but equivalent, expressions that can both be derived using the approach of the previous subsection. In this equation, the dimension of the matrices $\underline{\underline{S}}^\pm$ is equal to the number of electrons per BvK zone (i.e., KB), and the elements when expressed in terms of the Bloch functions of Eq. (20) become

$$\begin{aligned} [S_{j_1 j_2}(k_1 k_2)]^\pm &= \langle \psi_{j_1}(k_1, \mathbf{r}) | e^{\pm i \Delta k z} | \psi_{j_2}(k_2, \mathbf{r}) \rangle \\ &= \delta_{k_1, k_2 \pm \Delta k} \langle \psi_{j_1}(k_1, \mathbf{r}) | e^{\pm i \Delta k z} | \psi_{j_2}(k_2, \mathbf{r}) \rangle \\ &= \delta_{k_1, k_2 \pm \Delta k} \langle u_{j_1}(k_1, \mathbf{r}) | u_{j_2}(k_2, \mathbf{r}) \rangle, \end{aligned} \quad (47)$$

where $\Delta k = \frac{2\pi}{Ka}$ is the spacing between the k points. Upon organizing the occupied Bloch waves in order of increasing k , the $\underline{\underline{S}}^\pm$ matrices have the simple structure shown in Fig. 8 with square blocks above and below the $k_1 = k_2$ diagonal (except at the corners). Thus, when using expression (45), the dipole moment per unit couples neighboring k points. Accordingly, when including an external electrostatic field according to Eq. (6), lattice periodicity is destroyed and, instead, one has to deal with the complete BvK zone. This may be contrasted with the expression of Eq. (46) that does not destroy lattice periodicity.

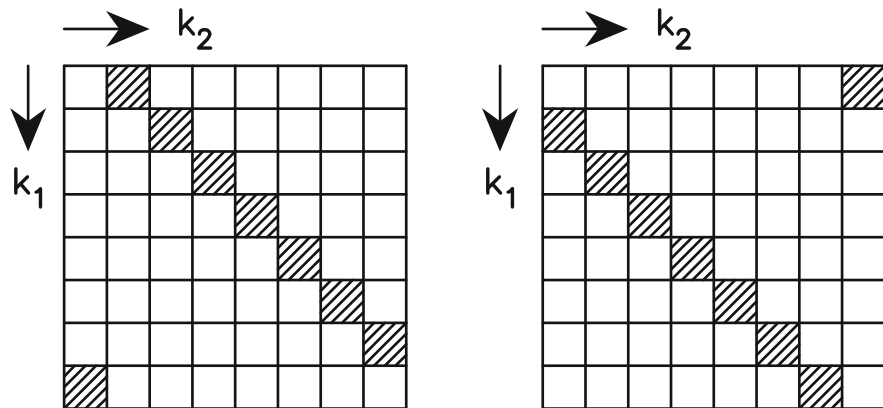


Fig. 8 Schematic presentation of the structure of the \underline{S}^\pm matrices when using Bloch functions as the basis. Only the elements inside the *shaded blocks* can be nonzero



Fig. 9 Schematic presentation of one set of equivalent Wannier functions for a quasi-one-dimensional system. The different functions are centered at different cells but otherwise equivalent. In the general case, there will be more sets of equivalent Wannier functions

In order to complete the picture, the nuclear contribution to the dipole moment per unit cell may be obtained simply as

$$\bar{\mu}_{zn} = \sum_m (Z_{mp} - pa) Q_m. \tag{48}$$

Here, Z_{mp} is the z component of the position of the m th nucleus in the p th unit cell, and Q_m is the nuclear charge. $\bar{\mu}_{zn}$ is independent of p .

It is useful to transform the set of K delocalized Bloch functions in Eq. (20) to a set of localized Wannier functions (see Fig. 9), i.e.,

$$w_{j,p}(\mathbf{r}) = \frac{1}{\sqrt{K}} \sum_{k=1}^K \psi_j(k, \mathbf{r}) e^{-ikap}$$

$$\psi_j(k, \mathbf{r}) = \frac{1}{\sqrt{K}} \sum_{p=1}^K w_{j,p}(\mathbf{r}) e^{ikap}. \tag{49}$$

Here, the Wannier function $w_{j,p}$ of the j th band is more or less localized to the p th unit inside the BvK zone. In terms of these functions, one can easily obtain

$$\begin{aligned}
 \bar{\mu}_{\text{KSV}} &= -\frac{i}{K} \sum_j \sum_k \langle u_j(k) | \frac{d}{dk} u_j(k) \rangle \\
 &= -\frac{i}{K^2} \sum_j \sum_{p_1, p_2} \sum_k e^{ika(p_2-p_1)} \langle w_{jp_1} | -i(z - ap_2) | w_{jp_2} \rangle \\
 &= -\frac{1}{K} \sum_j \sum_{p_1, p_2} \delta_{p_1, p_2} \langle w_{jp_1} | z - ap_2 | w_{jp_2} \rangle \\
 &= -\sum_{j=1}^B \langle w_{j,0}(\mathbf{r}) | z | w_{j,0}(\mathbf{r}) \rangle. \tag{50}
 \end{aligned}$$

(Notice the minus sign that originates from the negative charge of the electrons).

Alternatively, we may consider the LCAO approach of Eq. (21). Then

$$\begin{aligned}
 \bar{\mu}_{ze} &= \bar{\mu}_{\text{KSV}} \\
 &= -\frac{i}{K} \sum_j \sum_k \langle u_j(k, \mathbf{r}) | \frac{\partial}{\partial k} u_j(k, \mathbf{r}) \rangle \\
 &= -\frac{1}{K} \sum_{jm} \sum_k e^{ikma} \sum_{pq} c_{qj}^*(k) \left(\underbrace{\langle \chi_{q0} | z - ma | \chi_{pm} \rangle}_{\text{charge}} + i \underbrace{\langle \chi_{q0} | \chi_{pm} \rangle \frac{d}{dk}}_{\text{current}} \right) c_{pj}(k). \tag{51}
 \end{aligned}$$

Thereby we have split $\bar{\mu}_{ze}$ into so-called charge and current contributions. This separation can also be carried out including the nuclear part, which only affects the charge contribution, to give

$$\bar{\mu} = \bar{\mu}_{\text{charge}} + \bar{\mu}_{\text{current}}. \tag{52}$$

It can be shown analytically (Springborg and Kirtman 2011) that, for a large finite system, the charge contribution corresponds to that part of the dipole moment per unit that is associated with a central unit, i.e., μ_C of Eq. (11). Thus, the current contribution to the dipole moment per unit of the finite chain is due to the charge transfer between the termination regions, i.e., $q_R \cdot \mathbf{a}$ of Eq. (11), a proposition that has been verified numerically (Springborg and Kirtman 2008). This relation between the current term for the infinite, periodic system and the termination charges for the large, finite system leads to the interesting conclusion that covalently bonded systems with large band widths and delocalized electrons are those for which the largest surface charges can be found.

For our arguments, it is important to notice that each expression for the dipole moment per unit contains an unknown integer multiple of the lattice constant. In the expression of Eq. (45), $\text{Im} \ln \det \underline{S}^\pm$ is simply the phase of the complex number $\det \underline{S}^\pm$, which contains an unknown integer multiple of 2π . Thus, $\bar{\mu}_R$ is only

determined up to an integer times the lattice constant. Equivalently, in evaluating $\bar{\mu}_{\text{KSV}}$ of Eq. (46), we may modify each electronic orbital by a phase factor

$$\psi_j(k, \mathbf{r}) \rightarrow \psi_j(k, \mathbf{r})e^{i\phi_j(k)}. \quad (53)$$

The phases $\{\phi_j(k)\}$ are arbitrary except for satisfying

$$e^{i\phi_j(\pi/a)} = e^{i\phi_j(-\pi/a)} \quad (54)$$

or, equivalently,

$$\phi_j(\pi/a) = \phi_j(-\pi/a) + \tilde{n}_j \cdot 2\pi. \quad (55)$$

\tilde{n}_j is some integer. As a result, the dipole moment per unit is changed by an amount

$$-\tilde{n} \cdot a = -\sum_{j=1}^B \tilde{n}_j \cdot a. \quad (56)$$

This change arises entirely from the current term.

Finally, the expression for the dipole moment based on the Wannier functions in Eq. (49) also contains an equivalent additive integer. The reason is that the assignment of the localized functions to individual units is nonunique (see Fig. 9).

Higher Dimensions

In higher dimensions, we may use approaches similar to those above for each of the components of the dipole moment. Then, we will write

$$\mathbf{E} \cdot \bar{\boldsymbol{\mu}} = \sum_s E_s \bar{\mu}_s. \quad (57)$$

One possibility is to use expressions based on the $\underline{\underline{S}}_s^\pm$ matrices,

$$\bar{\mu}_s = \frac{\pm 1}{K \Delta k_s} \text{Im} \ln \det \underline{\underline{S}}_s^\pm. \quad (58)$$

In the d -dimensional case,

$$\Delta k_s \propto K^{-1/d}. \quad (59)$$

Therefore, the fact that $\text{Im} \ln \det \underline{\underline{S}}_s^\pm$ contains an undefined additive multiple of 2π means that $\bar{\mu}_s$ contains an undefined additive constant that only in the quasi-one-dimensional case ($d = 1$) becomes an additive multiple of the lattice constant.

Not surprising, this is also the case when using an expression like Eq. (46). We consider a substitution like that of Eq. (53) and replace the summation over all \mathbf{k} through an integration over the first Brillouin zone,

$$\begin{aligned}
 \mathbf{E} \cdot \bar{\boldsymbol{\mu}}_e &= -\frac{i}{K} \sum_{\mathbf{k}} \sum_{j=1}^B \langle e^{-i\mathbf{k}\cdot\mathbf{r}} \psi_j(\mathbf{k}, \mathbf{r}) | \mathbf{E} \cdot \nabla_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}} \psi_j(\mathbf{k}, \mathbf{r}) \rangle \\
 &\rightarrow -\frac{i}{K} \sum_{\mathbf{k}} \sum_{j=1}^B \left[\langle e^{-i\mathbf{k}\cdot\mathbf{r}} \psi_j(\mathbf{k}, \mathbf{r}) | \mathbf{E} \cdot \nabla_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}} \psi_j(\mathbf{k}, \mathbf{r}) \rangle \right. \\
 &\quad \left. + i \mathbf{E} \cdot \nabla_{\mathbf{k}} \phi_j(\mathbf{k}) \langle \psi_j(\mathbf{k}, \mathbf{r}) | \psi_j(\mathbf{k}, \mathbf{r}) \rangle \right] \\
 &= \mathbf{E} \cdot \bar{\boldsymbol{\mu}}_{e0} + \frac{1}{K} \sum_{\mathbf{k}} \sum_{j=1}^B \mathbf{E} \cdot \nabla_{\mathbf{k}} \phi_j(\mathbf{k}) \\
 &\rightarrow \mathbf{E} \cdot \bar{\boldsymbol{\mu}}_{e0} + \frac{1}{V_{\text{BZ}}} \sum_{j=1}^B \int_{1\text{BZ}} \mathbf{E} \cdot \nabla_{\mathbf{k}} \phi_j(\mathbf{k}) d\mathbf{k} \\
 &= \mathbf{E} \cdot \bar{\boldsymbol{\mu}}_{e0} + \frac{1}{V_{\text{BZ}}} \sum_{j=1}^B \int_{1\text{BZ}} \nabla_{\mathbf{k}} \cdot [\mathbf{E} \phi_j(\mathbf{k})] d\mathbf{k} \\
 &= \mathbf{E} \cdot \bar{\boldsymbol{\mu}}_{e0} + \frac{1}{V_{\text{BZ}}} \sum_{j=1}^B \int_{\delta 1\text{BZ}} \mathbf{1}_k \cdot [\mathbf{E} \phi_j(\mathbf{k})] d\mathbf{k} \\
 &= \mathbf{E} \cdot \bar{\boldsymbol{\mu}}_{e0} + \frac{1}{V_{\text{BZ}}} \sum_{j=1}^B \int_{\frac{1}{2}\delta 1\text{BZ}} \mathbf{1}_k \cdot \mathbf{E} [\phi_j(\mathbf{k}) - \phi_j(-\mathbf{k})] d\mathbf{k}. \quad (60)
 \end{aligned}$$

Here, $\bar{\boldsymbol{\mu}}_{e0}$ is the reference value before the substitution of Eq. (53), V_{BZ} is the volume of the first Brillouin zone, and we have used Gauss' theorem (the divergence theorem) to change the volume integral (labelled "1BZ") into an integral over the surface of the first Brillouin zone (labelled " $\delta 1\text{BZ}$ ") with $\mathbf{1}_k$ being an outwards pointing normal. We also utilize the fact that the integral over the complete surface of the first Brillouin zone can be written as a sum of two integrals, each over one half of the first Brillouin zone (labelled " $\frac{1}{2}\delta 1\text{BZ}$ "). Finally, based on the continuity in \mathbf{k} space,

$$\psi_j(\mathbf{k}, \mathbf{r}) = \psi_j(\mathbf{G} + \mathbf{k}, \mathbf{r}) \quad (61)$$

(with \mathbf{G} a lattice vector in reciprocal space), we define the integers

$$\phi_j(\mathbf{k}) - \phi_j(-\mathbf{k}) = 2\pi n_j(\mathbf{k}) \quad (62)$$

for \mathbf{k} on the boundary of the first Brillouin zone. Then, Eq. (60) becomes

$$\mathbf{E} \cdot \bar{\boldsymbol{\mu}}_e \rightarrow \mathbf{E} \cdot \bar{\boldsymbol{\mu}}_{e0} + \frac{2\pi}{V_{\text{BZ}}} \sum_{j=1}^B \int_{\frac{1}{2}\delta\text{1BZ}} \mathbf{1}_k \cdot \mathbf{E} n_j(\mathbf{k}) d\mathbf{k}. \quad (63)$$

As above, this expression also demonstrates that the dipole moment contains an additive constant that can take any value.

The Single-Particle Equations in the Periodic Case

For any large, finite, molecular system in the presence of a uniform electrostatic field, the electronic energy may be written, within the Born-Oppenheimer approximation, as $\langle \Psi | \hat{H}_e | \Psi \rangle - \langle \Psi | \mathbf{E} \cdot \hat{\boldsymbol{\mu}}_e | \Psi \rangle$. Here \hat{H}_e and $\hat{\boldsymbol{\mu}}_e$ are the usual field-free electronic Hamilton operator and the operator for the electronic part of the dipole moment, respectively. We use a single-particle (Hartree-Fock or Kohn-Sham) approach so that Ψ is the Slater determinant containing all the occupied single-particle orbitals. The latter are, then, expanded in a set of basis functions,

$$\psi_j(\mathbf{x}) = \sum_l c_{lj} \chi_l(\mathbf{x}). \quad (64)$$

As is standard, the expansion coefficients are determined by requiring that

$$\frac{\partial}{\partial c_{mp}^*} \left[\langle \Psi | \hat{H}_e | \Psi \rangle - \langle \Psi | \mathbf{E} \cdot \hat{\boldsymbol{\mu}}_e | \Psi \rangle - \sum_{mn} \lambda_{mn} (\langle \psi_m | \psi_n \rangle - \delta_{m,n}) \right] = 0. \quad (65)$$

For the infinite, periodic system, $\langle \Psi | \mathbf{E} \cdot \hat{\boldsymbol{\mu}}_e | \Psi \rangle$ becomes a nontrivial expression. We may choose between using expressions based either on the S_{\pm}^{\pm} matrices or on the ∇_k operators. In the former case, \mathbf{k} is no longer a good quantum number, and the secular equation becomes an eigenvalue equation with the dimension of the BvK zone (Springborg and Kirtman 2008). Moreover, the resulting single-particle equations become highly nonlinear in the orbital expansion coefficients. Therefore, we do not follow this approach but consider instead

$$\langle \Psi | \mathbf{E} \cdot \bar{\boldsymbol{\mu}}_e | \Psi \rangle = -\frac{i}{K} \sum_{\mathbf{k}} \sum_{j=1}^B \langle e^{-i\mathbf{k}\cdot\mathbf{r}} \psi_j(\mathbf{k}, \mathbf{r}) | \mathbf{E} \cdot \nabla_k | e^{-i\mathbf{k}\cdot\mathbf{r}} \psi_j(\mathbf{k}, \mathbf{r}) \rangle. \quad (66)$$

The Dipole Moment Operator

The expression of Eq. (66) corresponds to a single-particle, dipole moment operator for the electrons of the form

$$\hat{\boldsymbol{\mu}}_e = i e^{i\mathbf{k}\cdot\mathbf{r}} \nabla_k e^{-i\mathbf{k}\cdot\mathbf{r}}. \quad (67)$$

It is worthwhile to discuss the properties of this operator. In this connection, we consider two operators, namely,

$$\begin{aligned}\hat{S}_1 &= i \nabla_k \\ \hat{S}_2 &= i e^{i\mathbf{k}\cdot\mathbf{r}} \nabla_k e^{-i\mathbf{k}\cdot\mathbf{r}}\end{aligned}\quad (68)$$

and let them act on the functions

$$f_m(\mathbf{k}, \mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_m(\mathbf{k}, \mathbf{r}) \quad (69)$$

with $u_m(\mathbf{k}, \mathbf{r})$ being lattice periodic. We also define

$$\begin{aligned}S_{mn}(\mathbf{k}) &= \langle f_m(\mathbf{k}, \mathbf{r}) | f_n(\mathbf{k}, \mathbf{r}) \rangle \\ &= \langle e^{i\mathbf{k}\cdot\mathbf{r}} u_m(\mathbf{k}, \mathbf{r}) | e^{i\mathbf{k}\cdot\mathbf{r}} u_n(\mathbf{k}, \mathbf{r}) \rangle \\ &= \langle u_m(\mathbf{k}, \mathbf{r}) | u_n(\mathbf{k}, \mathbf{r}) \rangle.\end{aligned}\quad (70)$$

Then, we may obtain two alternative expressions for $i \nabla_k S_{mn}(\mathbf{k})$, i.e.,

$$\begin{aligned}i \nabla_k S_{mn}(\mathbf{k}) &= \langle -i \nabla_k f_m(\mathbf{k}, \mathbf{r}) | f_n(\mathbf{k}, \mathbf{r}) \rangle + \langle f_m(\mathbf{k}, \mathbf{r}) | i \nabla_k f_n(\mathbf{k}, \mathbf{r}) \rangle \\ &= -\langle f_n(\mathbf{k}, \mathbf{r}) | i \nabla_k f_m(\mathbf{k}, \mathbf{r}) \rangle^* + \langle f_m(\mathbf{k}, \mathbf{r}) | i \nabla_k f_n(\mathbf{k}, \mathbf{r}) \rangle \\ &= -\langle f_n(\mathbf{k}, \mathbf{r}) | i \nabla_k | f_m(\mathbf{k}, \mathbf{r}) \rangle^* + \langle f_m(\mathbf{k}, \mathbf{r}) | i \nabla_k | f_n(\mathbf{k}, \mathbf{r}) \rangle \\ &= -\langle f_n(\mathbf{k}, \mathbf{r}) | \hat{S}_1 | f_m(\mathbf{k}, \mathbf{r}) \rangle^* + \langle f_m(\mathbf{k}, \mathbf{r}) | \hat{S}_1 | f_n(\mathbf{k}, \mathbf{r}) \rangle\end{aligned}\quad (71)$$

as well as

$$\begin{aligned}i \nabla_k S_{mn}(\mathbf{k}) &= \langle -i \nabla_k u_m(\mathbf{k}, \mathbf{r}) | u_n(\mathbf{k}, \mathbf{r}) \rangle + \langle u_m(\mathbf{k}, \mathbf{r}) | i \nabla_k u_n(\mathbf{k}, \mathbf{r}) \rangle \\ &= -\langle u_n(\mathbf{k}, \mathbf{r}) | i \nabla_k u_m(\mathbf{k}, \mathbf{r}) \rangle^* + \langle u_m(\mathbf{k}, \mathbf{r}) | i \nabla_k u_n(\mathbf{k}, \mathbf{r}) \rangle \\ &= -\langle e^{-i\mathbf{k}\cdot\mathbf{r}} f_n(\mathbf{k}, \mathbf{r}) | i \nabla_k e^{-i\mathbf{k}\cdot\mathbf{r}} f_m(\mathbf{k}, \mathbf{r}) \rangle^* \\ &\quad + \langle e^{-i\mathbf{k}\cdot\mathbf{r}} f_m(\mathbf{k}, \mathbf{r}) | i \nabla_k e^{-i\mathbf{k}\cdot\mathbf{r}} f_n(\mathbf{k}, \mathbf{r}) \rangle \\ &= -\langle f_n(\mathbf{k}, \mathbf{r}) | i e^{i\mathbf{k}\cdot\mathbf{r}} \nabla_k e^{-i\mathbf{k}\cdot\mathbf{r}} | f_m(\mathbf{k}, \mathbf{r}) \rangle^* \\ &\quad + \langle f_m(\mathbf{k}, \mathbf{r}) | e^{i\mathbf{k}\cdot\mathbf{r}} i \nabla_k e^{-i\mathbf{k}\cdot\mathbf{r}} | f_n(\mathbf{k}, \mathbf{r}) \rangle \\ &= -\langle f_n(\mathbf{k}, \mathbf{r}) | \hat{S}_2 | f_m(\mathbf{k}, \mathbf{r}) \rangle^* + \langle f_m(\mathbf{k}, \mathbf{r}) | \hat{S}_2 | f_n(\mathbf{k}, \mathbf{r}) \rangle.\end{aligned}\quad (72)$$

Thus, only for functions satisfying

$$\nabla_k S_{mn}(\mathbf{k}) = \mathbf{0}, \quad (73)$$

can we say that the operators \hat{s}_1 and \hat{s}_2 are hermitian. That is true, for example, in the case of the orthonormal band orbitals. On the other hand,

$$\hat{s}_2 - \hat{s}_1 = \mathbf{r} \quad (74)$$

is always hermitian.

However, in a practical calculation, we ignore \hat{s}_1 of the reasons discussed in section “The Single-Electron Case” and let \hat{s}_2 act on Bloch functions formed from atom-centered orbitals, and as a result, Eq. (73) is not necessarily satisfied. Therefore, we make the replacement

$$\hat{s}_2 \rightarrow \frac{1}{2} (\hat{s}_2^\dagger + \hat{s}_2). \quad (75)$$

If the single-particle orbitals, $\{\psi_j\}$, are written as a linear combination of Bloch functions of atomic orbitals:

$$\psi_j(\mathbf{x}) = \psi_j(\mathbf{k}, \mathbf{x}) = \sum_l c_{lj}(\mathbf{k}) \chi_l(\mathbf{k}, \mathbf{x}) = \lim_{N \rightarrow \infty} \sum_l c_{lj}(\mathbf{k}) \frac{1}{\sqrt{N}} \sum_{\mathbf{n}} e^{i\mathbf{k} \cdot \mathbf{R}_n} \chi_{l,\mathbf{n}}(\mathbf{r}), \quad (76)$$

then

$$\begin{aligned} & \langle \Psi | \mathbf{E} \cdot \bar{\boldsymbol{\mu}}_e | \Psi \rangle \\ &= -\frac{i}{K} \sum_{\mathbf{k}} \sum_{j=1}^B \langle \psi_j(\mathbf{k}, \mathbf{r}) | \mathbf{E} \cdot \frac{1}{2} (\hat{s}_2^\dagger + \hat{s}_2) | \psi_j(\mathbf{k}, \mathbf{r}) \rangle \\ &= \lim_{N \rightarrow \infty} \frac{1}{2N} \sum_{\mathbf{k}} \sum_{j=1}^B \sum_{l_1, l_2} \sum_{\mathbf{n}_1, \mathbf{n}_2} e^{i\mathbf{k} \cdot (\mathbf{R}_{\mathbf{n}_2} - \mathbf{R}_{\mathbf{n}_1})} \\ & \quad \times \left\{ c_{l_1 j}^*(\mathbf{k}) c_{l_2 j}(\mathbf{k}) \left[\langle \chi_{l_1, \mathbf{n}_1} | \mathbf{E} \cdot (\mathbf{r} - \mathbf{R}_{\mathbf{n}_2}) | \chi_{l_2, \mathbf{n}_2} \rangle + \langle \chi_{l_1, \mathbf{n}_1} | \mathbf{E} \cdot (\mathbf{r} - \mathbf{R}_{\mathbf{n}_1}) | \chi_{l_2, \mathbf{n}_2} \rangle \right] \right. \\ & \quad \left. + i \left[-\mathbf{E} \cdot (\nabla_{\mathbf{k}} c_{l_1 j}^*(\mathbf{k})) c_{l_2 j}(\mathbf{k}) + c_{l_1 j}^*(\mathbf{k}) \mathbf{E} \cdot \nabla_{\mathbf{k}} c_{l_2 j}(\mathbf{k}) \right] \langle \chi_{l_1, \mathbf{n}_1} | \chi_{l_2, \mathbf{n}_2} \rangle \right\}. \quad (77) \end{aligned}$$

It is convenient to introduce the dipole matrix elements

$$\begin{aligned} \mathbf{M}_{l_1, l_2}(\mathbf{k}) &= \lim_{N \rightarrow \infty} \sum_{\mathbf{n}} e^{i\mathbf{k} \cdot \mathbf{R}_n} \langle \chi_{l_1, \mathbf{0}} | \mathbf{r} - \mathbf{R}_n | \chi_{l_2, \mathbf{n}} \rangle \\ &= \lim_{N \rightarrow \infty} \sum_{\mathbf{n}} e^{-i\mathbf{k} \cdot \mathbf{R}_n} \langle \chi_{l_1, \mathbf{n}} | \mathbf{r} | \chi_{l_2, \mathbf{0}} \rangle \quad (78) \end{aligned}$$

as well as the overlap matrix elements

$$\begin{aligned} S_{l_1, l_2}(\mathbf{k}) &= \lim_{N \rightarrow \infty} \sum_{\mathbf{n}} e^{i\mathbf{k} \cdot \mathbf{R}_n} \langle \chi_{l_1, \mathbf{0}} | \chi_{l_2, \mathbf{n}} \rangle \\ &= \lim_{N \rightarrow \infty} \sum_{\mathbf{n}} e^{-i\mathbf{k} \cdot \mathbf{R}_n} \langle \chi_{l_1, \mathbf{n}} | \chi_{l_2, \mathbf{0}} \rangle. \end{aligned} \quad (79)$$

In terms of these matrix elements, Eq. (77) becomes

$$\begin{aligned} &\langle \Psi | \mathbf{E} \cdot \bar{\boldsymbol{\mu}}_e | \Psi \rangle \\ &= \frac{1}{2} \sum_{\mathbf{k}} \sum_{j=1}^B \sum_{l_1, l_2} \left\{ c_{l_1 j}^*(\mathbf{k}) c_{l_2 j}(\mathbf{k}) \mathbf{E} \cdot [\mathbf{M}_{l_1, l_2}(\mathbf{k}) + \mathbf{M}_{l_2, l_1}^*(\mathbf{k})] \right. \\ &\quad \left. + i \left[-\mathbf{E} \cdot (\nabla_{\mathbf{k}} c_{l_1 j}^*(\mathbf{k})) c_{l_2 j}(\mathbf{k}) + c_{l_1 j}^*(\mathbf{k}) \mathbf{E} \cdot \nabla_{\mathbf{k}} c_{l_2 j}(\mathbf{k}) \right] S_{l_1, l_2}(\mathbf{k}) \right\}. \end{aligned} \quad (80)$$

The Secular Equation

The secular equation is obtained by taking the derivative of the total electronic energy with respect to an orbital expansion coefficient, $c_{l_p j_p}^*(\mathbf{k}_p)$, similar to what was done for the finite system in Eq. (65). Except for the term that depends explicitly on the electrostatic field, the “standard” approach may be followed. It is more difficult, however, to treat the electrostatic field term given in Eq. (80). For the sake of simplicity that is done below for the quasi-one-dimensional case:

$$\begin{aligned} &\langle \Psi | E \bar{\mu}_{ez} | \Psi \rangle \\ &= \frac{1}{2} \sum_k \sum_{j=1}^B \sum_{l_1, l_2} \left\{ c_{l_1 j}^*(k) c_{l_2 j}(k) E \cdot [M_{l_1, l_2}(k) + M_{l_2, l_1}^*(k)] \right. \\ &\quad \left. + i \left[-E \cdot \frac{d}{dk} c_{l_1 j}^*(k) c_{l_2 j}(k) + c_{l_1 j}^*(k) E \cdot \frac{d}{dk} c_{l_2 j}(k) \right] S_{l_1, l_2}(k) \right\}. \end{aligned} \quad (81)$$

If we use

$$\begin{aligned} &\frac{\partial}{\partial c_{l_p j_p}^*(k_p)} \sum_k \sum_{j=1}^B \sum_{l_1, l_2} i E \cdot \frac{d}{dk} c_{l_1 j}^*(k) c_{l_2 j}(k) S_{l_1, l_2}(k) \\ &\simeq \frac{\partial}{\partial c_{l_p j_p}^*(k_p)} \sum_k \sum_{j=1}^B \sum_{l_1, l_2} i E \cdot \frac{c_{l_1 j}^*(k + \Delta k) - c_{l_1 j}^*(k)}{\Delta k} c_{l_2 j}(k) S_{l_1, l_2}(k) \end{aligned}$$

$$\begin{aligned}
&= \sum_{l_2} iE \cdot \frac{c_{l_2 j_p}(k_p - \Delta k) S_{l_p, l_2}(k_p - \Delta k) - c_{l_2 j_p}(k_p) S_{l_p, l_2}(k_p)}{\Delta k} \\
&\simeq - \sum_{l_2} iE \cdot \frac{d}{dk} [c_{l_2 j_p}(k_p) S_{l_p, l_2}(k_p)]. \tag{82}
\end{aligned}$$

then

$$\begin{aligned}
&\frac{\partial}{\partial c_{l_p j_p}^*(k_p)} \langle \Psi | E \bar{\mu}_{ez} | \Psi \rangle \\
&= \frac{1}{2} \sum_{l_2} \left\{ c_{l_2 j_p}(k_p) E \cdot \left[M_{l_p, l_2}(k_p) + M_{l_2, l_p}^*(k_p) + i \frac{d}{dk} S_{l_p, l_2}(k_p) \right] \right. \\
&\quad \left. + iE [S_{l_p, l_2}(k_p) + S_{l_p, l_2}(k_p)] \frac{dc_{l_2 j_p}(k_p)}{dk} \right\}. \tag{83}
\end{aligned}$$

We have

$$M_{l_p, l_2}(k_p) + M_{l_2, l_p}^*(k_p) + i \frac{d}{dk} S_{l_p, l_2}(k_p) = 2M_{l_p, l_2}(k_p) \tag{84}$$

so that in the general one-, two-, or three-dimensional case, we have the following contribution to the secular equation from the electrostatic field

$$\sum_{l_2} \left[\mathbf{E} \cdot \mathbf{M}_{l_p, l_2}(\mathbf{k}_p) c_{l_2 j_p}(\mathbf{k}_p) + i S_{l_p, l_2}(\mathbf{k}_p) \mathbf{E} \cdot \nabla_{\mathbf{k}} c_{l_2 j_p}(\mathbf{k}_p) \right]. \tag{85}$$

This expression is actually the one we would obtain when using \hat{s}_2 directly without making the substitution of Eq. (75).

With this, the single-particle equations become

$$\left(\underline{\underline{F}}(\mathbf{k}) + \mathbf{E} \cdot \left[\underline{\underline{M}}(\mathbf{k}) + i \underline{\underline{S}}(\mathbf{k}) \nabla_{\mathbf{k}} \right] \right) \cdot \underline{\underline{c}}_j(\mathbf{k}) = \epsilon_j(\mathbf{k}) \cdot \underline{\underline{S}}(\mathbf{k}) \cdot \underline{\underline{c}}_j(\mathbf{k}). \tag{86}$$

Here, $\underline{\underline{F}}$ is the “standard” Fock (or Kohn-Sham) matrix although evaluated using the field-containing orbitals.

The Vector Potential Approach

Originally (Kirtman et al. 2000), Eq. (86) was derived using a vector potential approach in describing the effects of the external electric field (see section The Gauge). This approach allows for a simple treatment of dynamic, in addition to static, electric fields and shall therefore be briefly described here.

For the infinite, periodic system exposed to a time-dependent vector potential, the time-dependent single-particle equations are

$$\left(\hat{F} - i \frac{\partial}{\partial t}\right) \psi_{j_1}(\mathbf{k}, \mathbf{r}, t) = \sum_{j_2} \epsilon_{j_2 j_1}(k, t) \psi_{j_2}(\mathbf{k}, \mathbf{r}, t) \quad (87)$$

and, using the LCAO expansion of Eq. (64),

$$\begin{aligned} \underline{\underline{F}}(\mathbf{k}, t) \cdot \underline{\underline{c}}(\mathbf{k}, t) - \mathbf{E}(t) \cdot \left[\underline{\underline{M}}(\mathbf{k}) \cdot \underline{\underline{c}}(\mathbf{k}, t) + i \underline{\underline{S}}(\mathbf{k}) \nabla_k \underline{\underline{c}}(\mathbf{k}, t) \right] - i \underline{\underline{S}}(\mathbf{k}) \frac{\partial}{\partial t} \underline{\underline{c}}(\mathbf{k}, t) \\ = \underline{\underline{S}}(\mathbf{k}) \cdot \underline{\underline{c}}(\mathbf{k}, t) \cdot \underline{\underline{c}}(\mathbf{k}, t). \end{aligned} \quad (88)$$

Here, $\mathbf{E}(t)$ is the electric field obtained by taking the time derivative of the vector potential. As will be discussed in a later section, Eq. (88) offers an approach wherein the response to dynamic electric fields may be treated using perturbation theory. It is also valid in the limit of a static field, i.e., $\mathbf{E}(t)$ is time-independent. In that case, one arrives at Eq. (86). By solving for the linear and quadratic response to such fields, the linear and nonlinear static hyperpolarizabilities may be calculated. In addition, the same calculations determine infrared (Maschio et al. 2012, 2013) and Raman (Maschio et al. 2013) vibrational intensities. It should be noted that the latter are found by taking advantage of interchange-type relations that avoid having to solve equations for the response to the various vibrational displacements. This leads to a very efficient, fully analytical, procedure.

Solving the Secular Equation

We shall now return to the electrostatic field case. Because of the ∇_k term, Eq. (86) is not a standard matrix-eigenvalue problem. However, by using the normalization condition

$$\underline{\underline{1}} = \underline{\underline{c}}^\dagger(\mathbf{k}) \cdot \underline{\underline{S}}(\mathbf{k}) \cdot \underline{\underline{c}}(\mathbf{k}), \quad (89)$$

one can obtain

$$\underline{\underline{S}}(\mathbf{k}) \nabla_k \underline{\underline{c}}(\mathbf{k}) = \left\{ \underline{\underline{S}}(\mathbf{k}) \left[\nabla_k \underline{\underline{c}}(\mathbf{k}) \right] \underline{\underline{c}}^\dagger(\mathbf{k}) \underline{\underline{S}}(\mathbf{k}) \right\} \underline{\underline{c}}(\mathbf{k}), \quad (90)$$

whereby Eq. (86) can be cast into the standard form. In that case, the term in the curly brackets in Eq. (90) (as well as the Fock matrix) has to be treated self-consistently.

Even after the reformulation of Eq. (90), the \mathbf{k} derivatives of the orbital expansion coefficients are needed. As indicated, Eq. (86) conforms with the lattice periodicity so that \mathbf{k} is a good quantum number. Thus, in a practical calculation, one will solve

Eq. (86) [with the modification of Eq. (90)] for each \mathbf{k} point separately. Ultimately, this means that the orbital expansion coefficients may contain essentially arbitrary phase factors and, for \mathbf{k} where bands are crossing, more or less random linear combinations of the degenerate orbitals are determined. When calculating the \mathbf{k} derivatives of the orbital coefficients numerically (which is the way we shall do it), the phase factor and band-crossing issues can make the numerical derivatives highly inaccurate.

In order to circumvent the above problems, we have developed a smoothing technique (Springborg and Kirtman 2008) with which band-crossings and degeneracies are treated. Subsequently, the orbital expansion coefficients are modified according to

$$c_{lj}(\mathbf{k}) \rightarrow c_{lj}(\mathbf{k})e^{i\phi_j(\mathbf{k})} \quad (91)$$

so that the coefficients of neighboring \mathbf{k} points become maximally similar. This smoothing is a multistep procedure that will not be described in detail here.

In a practical calculation, we have an equidistant set of \mathbf{k} points in the three directions of \mathbf{k} space. Then, a derivative can be obtained as

$$\frac{\partial f(\mathbf{k})}{\partial k_s} \simeq \frac{1}{2\Delta k_s} \sum_{n=1}^{N_k} w_{n,N_k}^{(1)} [f(\mathbf{k} + n \cdot \Delta \mathbf{k}_s) - f(\mathbf{k} - n \cdot \Delta \mathbf{k}_s)]. \quad (92)$$

Here, s defines the direction in \mathbf{k} space and $\Delta \mathbf{k}_s$ is the spacing in that direction with the length Δk_s . A set of weight coefficients $\{w_{n,N_k}^{(1)}\}$ has been given by Dvornikov (2007) and by Khan and Ohba (2000),

$$w_{n,N_k}^{(1)} = \frac{1}{n\pi_n(N_k)} \quad (93)$$

with

$$\pi_n(N_k) = \prod_{k=1, k \neq n}^{N_k} \left(1 - \frac{n^2}{k^2}\right). \quad (94)$$

Alternative expressions (differing in the values of the weight factors $w_{n,N_k}^{(1)}$) were given later by Li (2005). Both formulas are based on expanding the function of interest $f(\mathbf{k})$ in some power or polynomial series about the point of interest at which the derivative(s) are calculated and, subsequently, differentiating the fitting function. However, as pointed out by Press et al. (1992), even if the function $f(\mathbf{k})$ is well behaved for real values of \mathbf{k} and we are only interested in such values, the existence of poles for complex \mathbf{k} can make such an approach unreliable. In fact, model calculations for a two-band system with an avoided band-crossing (in which all quantities are determined analytically) demonstrate that the treatments of Dvornikov, Khan and Ohba, and Li are inaccurate exactly at those \mathbf{k} that are

in the neighborhood of the avoided band-crossing where the derivatives $\frac{\partial f(\mathbf{k})}{\partial k_s}$ are numerically largest. So we use, instead, a rational approximation

$$f(\mathbf{k} + \Delta\mathbf{k}) \simeq \frac{a_0 + a_1|\Delta\mathbf{k}| + a_2|\Delta\mathbf{k}|^2 + \dots + a_N|\Delta\mathbf{k}|^N}{1 + b_1|\Delta\mathbf{k}| + b_2|\Delta\mathbf{k}|^2 + \dots + b_N|\Delta\mathbf{k}|^N}. \tag{95}$$

That the polynomials in the numerator and denominator are of the same order is not a necessity but we have made that choice. The $2N + 1$ unknown coefficients $a_0, a_1, a_2, \dots, a_N$ and b_1, b_2, \dots, b_N are determined by requiring that Eq. (95) is exactly satisfied for

$$\Delta\mathbf{k} = n\Delta\mathbf{k}_s, \quad n = -N, -N + 1, \dots, N - 1, N \tag{96}$$

for which we have the exact values of $f(\mathbf{k} + \Delta\mathbf{k})$. This condition can be written as a set of linear equations,

$$\sum_{j=0}^N a_j |n\Delta\mathbf{k}_s|^j - \sum_{j=1}^N b_j |n\Delta\mathbf{k}_s|^j f(\mathbf{k} + n\Delta\mathbf{k}_s) = f(\mathbf{k} + n\Delta\mathbf{k}_s), \tag{97}$$

that are solved using a singular value decomposition approach. These equations have to be set up and solved for each \mathbf{k} separately. From the approximate expression in Eq. (95), we can calculate the derivatives directly,

$$\frac{\partial f(\mathbf{k})}{\partial k_s} \simeq a_1 - b_1 a_0. \tag{98}$$

Finally, we mention an alternative suggestion (Kirtman et al. 2000), i.e., to write

$$\nabla_k \underline{\underline{c}}(\mathbf{k}) = \underline{\underline{c}}(\mathbf{k}) \cdot \underline{\underline{Q}}(\mathbf{k}). \tag{99}$$

Here, all matrix elements of $\underline{\underline{Q}}(\mathbf{k})$ can be determined analytically with a single exception: the imaginary part of the diagonal elements. This exception is due to the fact that the orbital expansion coefficients can be modified by arbitrary phase factors as shown, e.g., in Eq. (91).

For the one-dimensional case, $\frac{dc_{pj}(k)}{dk}$ in Eq. (51) is expressed in terms of the k dependent $\underline{\underline{Q}}$ matrix defined by

$$\frac{dc_{pj}(k)}{dk} = \sum_{j'} c_{pj'}(k) Q_{j'j}(k). \tag{100}$$

Then, using the orthonormality of the crystal orbitals

$$\sum_m e^{ikam} \sum_{pq} c_{qj'}^*(k) \langle \chi_{q0} | \chi_{pm} \rangle c_{pj}(k) = \sum_{pq} c_{qj'}^*(k) S_{qp}(k) c_{pj}(k) = \delta_{jj'}, \tag{101}$$

the current term in Eq. (51) becomes simply

$$\bar{\mu}_{\text{current}} = -\frac{i}{K} \sum_j Q_{jj}(k). \quad (102)$$

From the $\frac{d}{dk}$ derivative of the orthonormality condition in Eq. (101), it follows that

$$Q_{jj}(k) + Q_{jj}^*(k) = -\sum_{pq} c_{qj}^*(k) \frac{dS_{qp}(k)}{dk} c_{pj}(k). \quad (103)$$

This relation yields the real part of the diagonal elements of $Q_{jj}(k)$ but leaves the imaginary part unknown. In fact, the integers \tilde{n}_j of Eq. (55) are determined by the imaginary part of $Q_{jj}(k)$.

Applications

The approach we have sketched above has not yet been applied to a large number of systems or properties. However, some results have been obtained that will be discussed here briefly.

Structure

For the large finite system exposed to an electrostatic field, it is relatively easy to obtain analytical expressions for the forces acting on the nuclei, i.e.,

$$\mathbf{F}_p = -\nabla_p \left[\langle \Psi | \hat{H} - \hat{\mu}_e \cdot \mathbf{E} | \Psi \rangle + \frac{1}{2} \sum_{p_1 \neq p_2} \frac{Z_{p_1} Z_{p_2}}{|\mathbf{R}_{p_1} - \mathbf{R}_{p_2}|} + \sum_p Z_p \mathbf{R}_p \cdot \mathbf{E} \right] \quad (104)$$

with Z_p being the charge of the p th nucleus placed at \mathbf{R}_p and with ∇_p being the gradient operator with respect to the coordinates of the p th nucleus. The last two terms in the square brackets are the contribution of the nuclei to the total energy due to the nuclear-nuclear repulsion and to the electrostatic field, respectively.

For the infinite, periodic system, the electronic dipole moment operator becomes more complicated involving derivatives with respect to \mathbf{k} . Nevertheless, it has become possible to derive analytical expressions for the forces also in this case (Springborg and Kirtman 2008). Thereby, the changes in the structural parameters of a given system due to the electrostatic field can be calculated automatically.

A crucial test of the approach for the infinite, periodic system is to verify that electronic and structural responses of a very large system can be calculated using the infinite-system approach. To that end, we considered the quasi-one-dimensional system of Fig. 10 using a model Hamiltonian. The system contains repeated units of

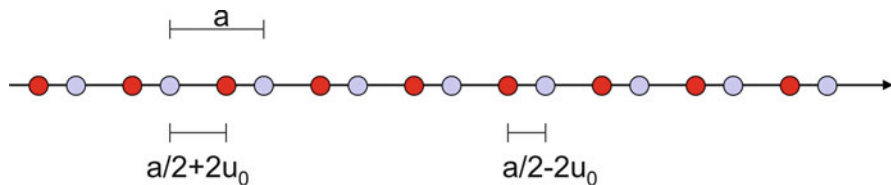


Fig. 10 Schematic presentation of a linear chain with alternating atoms and bond lengths

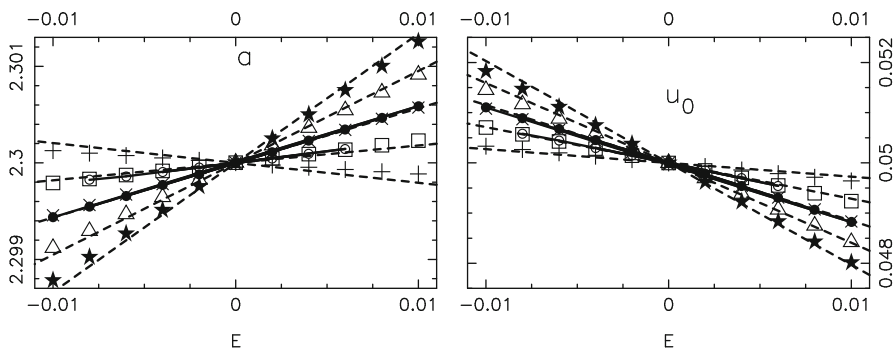


Fig. 11 Structural parameters obtained from a set of representative calculations for the model. The *open* and *filled circles* represent values for finite chains with 40/41 units. All other results are for periodic chains with $K = 80$. The finite chain calculations differ from one another by an end-to-end transfer of two electrons. For the periodic chains, \times , *triangles*, *squares*, *stars*, and $+$ mark results for the integer \tilde{n} being equal to 0, 2, -2 , 4, and -4 , respectively. The *straight*, *dashed lines* are linear approximations to the infinite periodic chain values

two different atoms with alternating nearest-neighbor bond lengths. We used a basis set of two functions per atom and parameterized the field-free Fock matrix. Through this approach, it became possible to treat fairly large, finite systems accurately as well as to remove problems due to summations in direct or reciprocal space.

We considered first, for fixed structure, the increment in the total energy of a large finite system upon adding one more diatomic unit:

$$\bar{\mathcal{E}}_{\text{tot}}(E) = \lim_{N \rightarrow \infty} [\mathcal{E}_{\text{tot}}(E, N + 1) - \mathcal{E}_{\text{tot}}(E, N)], \quad (105)$$

Here $\mathcal{E}_{\text{tot}}(E, N)$ is the total energy of the chain containing N units and exposed to an electrostatic field E . As required, the value of this increment was identical to the total energy per cell for the infinite, periodic system.

Of particular interest are the results of Fig. 11 where we show the optimized structures for the large, finite chains and for the infinite, periodic chain as a function of E . For the former, we use the structure of the unit cells in the middle of a chain that is made so large that a regular structure in the middle can be identified.

The figure demonstrates clearly that the two set of calculations lead to identical structural responses. Moreover, by modifying the charge at the terminations for the large finite system, different responses are obtained, a fact that we shall discuss

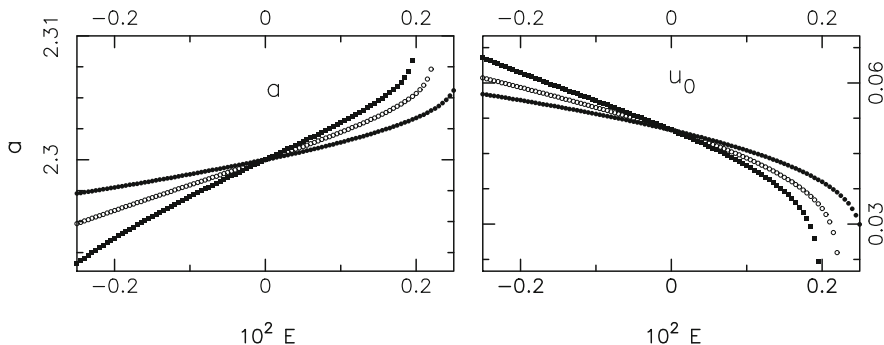


Fig. 12 Optimized values of the lattice constant, a , and bond-length alternation parameter, u_0 , for the infinite periodic chain as a function of the field, E . The different symbols mark results for different values of \tilde{n}

further in the next subsection. It is most noteworthy that the different periodic chain results can be modeled simply by varying the integer \tilde{n} of Eq. (56). This integer, which at first appears to be related solely to the mathematical formulation is, thereby, given a physical interpretation as well.

In Fig. 12, we show results from another set of model calculations whereby only the periodic system case is treated. In that figure, we can identify higher-order responses to the electrostatic field. Ultimately, the variation in the dipole moment or total energy per unit as a function of field strength can be fit by a power series in E . The coefficients yield the polarizabilities and hyperpolarizabilities, e.g., for the quasi-one-dimensional system, we may write

$$\bar{\mu} = \bar{\mu}^{(0)} + \bar{\alpha}E + \frac{1}{2}\bar{\beta}E^2 + \frac{1}{6}\bar{\gamma}E^3 + \dots, \quad (106)$$

where only the components parallel to the chain have been considered.

One may now distinguish between the electronic and the structural responses which, for instance, can be done theoretically by determining α , β , and γ for the field-free structure (thus giving an estimate for the electronic responses) and by allowing the structure to relax in the presence of the electrostatic field (thus giving the sum of electronic and structural responses).

From the model calculations, we found that the responses were largest for covalently bonded systems with large hopping integrals between neighboring atoms. These systems are also those for which the termination effects (quantified through \tilde{n}) are the largest. Moreover, for the higher-order responses (e.g., β and γ), the structural responses are dominant; they can easily account for 90% of the total.

Finally, we mention that a comparison of computation times for the finite system calculations with those for the periodic system (to obtain the results presented in Fig. 11) reveals that the latter are many (typically, 10) orders faster. From that point of view, then, the use of the infinite periodic system approach is highly recommended.

Converse Piezoelectricity

Piezoelectricity describes the ability of materials (most often, crystals) to produce an electric potential under the influence of an externally applied mechanical stress. For a material that is not short-circuited, the mechanical stress will lead to a separation of opposite electrical charges at opposite crystal surfaces which results in an electric voltage across the material. This effect is reversible in the sense that materials exhibiting the so-called direct piezoelectric effect, i.e., the production of electricity when a stress is applied, also exhibit the converse piezoelectric effect, i.e., the production of stress and/or strain when an electric field is applied. Thus, the results of Figs. 11 and 12 can be considered as manifestations of the converse piezoelectric effect for the systems of interest: upon applying an electrostatic field, the spatial dimension of the system changes. The fact that the figures show several different curves demonstrates that the converse piezoelectric effect depends on the terminations (or, for higher-dimensional systems, on the surfaces) of the system. We shall now discuss this in more detail.

Let us consider the model system of Fig. 10. The total energy per cell relative to that of the optimized structure in the absence of the electrostatic field (i.e., a_0 and u_0) may be approximated via a second-order Taylor expansion in $a - a_0$ and $u - u_0$ and a parameter that describes the external field, λ (λ will later be the field strength itself or the voltage)

$$\begin{aligned} \bar{\mathcal{E}}_{\text{tot}} \simeq & (a - a_0) \frac{\partial \bar{\mathcal{E}}_{\text{tot}}}{\partial a} + (u - u_0) \frac{\partial \bar{\mathcal{E}}_{\text{tot}}}{\partial u} + \lambda \frac{\partial \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda} + \frac{1}{2} (a - a_0)^2 \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2} \\ & + \frac{1}{2} (u - u_0)^2 \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} + (u - u_0)(a - a_0) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} + \frac{1}{2} \lambda^2 \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda^2} \\ & + \lambda (a - a_0) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda \partial a} + \lambda (u - u_0) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda \partial u}. \end{aligned} \quad (107)$$

We use the fact that

$$\frac{\partial \bar{\mathcal{E}}_{\text{tot}}}{\partial a} = \frac{\partial \bar{\mathcal{E}}_{\text{tot}}}{\partial u} = 0 \quad (108)$$

at the equilibrium structure in the field-free case and the same conditions for $\lambda \neq 0$ to find the optimized structure in the presence of a DC field. This yields

$$\begin{aligned} (a - a_0) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2} + \lambda \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda \partial a} + (u - u_0) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} &= 0 \\ \lambda \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda \partial u} + (a - a_0) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} + (u - u_0) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} &= 0 \end{aligned} \quad (109)$$

where all derivatives are taken at $u = u_0$, $a = a_0$, and $\lambda = 0$.

The solution of the linear simultaneous equations in Eq. (109) leads to expressions for a macroscopic, converse piezoelectric coefficient,

$$\eta = \left. \frac{da}{d\lambda} \right|_{\lambda=0} \quad (110)$$

as well as an “internal” converse piezoelectric coefficient for the internal degree of freedom,

$$\zeta = \left. \frac{du}{d\lambda} \right|_{\lambda=0}. \quad (111)$$

The result is

$$\begin{aligned} \eta &= \left[\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda \partial u} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} - \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda \partial a} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} \right] \cdot \left[\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} - \left(\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} \right)^2 \right]^{-1} \\ \zeta &= \left[\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda \partial a} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} - \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda \partial u} \right] \cdot \left[\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} - \left(\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} \right)^2 \right]^{-1}. \quad (112) \end{aligned}$$

We are interested in separating the contributions to η and ζ that arise from the charges at the terminations, cf. Eq. (11) from those that do not. To that end, the dipole moment per unit for the infinite, periodic system may be written as

$$\bar{\mu}_z = \bar{\mu}_{z0} + \tilde{n}a. \quad (113)$$

Thus, the first-order (linear in E) contribution to the total energy per unit for the infinite, periodic system exposed to an electrostatic field becomes

$$\bar{\mathcal{E}}_{\text{DC}} = E\bar{\mu}_{z0} + E\tilde{n}a. \quad (114)$$

At this point, we take

$$\lambda = E. \quad (115)$$

Then, in Eq. (112), all quantities except for

$$\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda \partial a} = \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial E \partial a} = \frac{\partial \bar{\mu}_{z0}}{\partial a} + \tilde{n} \quad (116)$$

are independent of the charges at the termination. In that case, the coefficients in Eq. (112) may be separated into terms that depend on \tilde{n} and terms that do not,

$$\begin{aligned} \eta &= \eta_0 + \tilde{n}\eta_1 \\ \zeta &= \zeta_0 + \tilde{n}\zeta_1 \end{aligned} \tag{117}$$

where

$$\begin{aligned} \eta_0 &= \left[-a \frac{\partial P_0}{\partial u} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} + \left(P_0 + a \frac{\partial P_0}{\partial a} \right) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} \right] \cdot \left[\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} - \left(\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} \right)^2 \right]^{-1} \\ \zeta_0 &= \left[- \left(P_0 + a \frac{\partial P_0}{\partial a} \right) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} + \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2} a \frac{\partial P_0}{\partial u} \right] \cdot \left[\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} - \left(\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} \right)^2 \right]^{-1} \\ \eta_1 &= \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} \cdot \left[\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} - \left(\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} \right)^2 \right]^{-1} \\ \zeta_1 &= - \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} \cdot \left[\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u^2} - \left(\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u \partial a} \right)^2 \right]^{-1} \end{aligned} \tag{118}$$

with

$$P_0 = \frac{\bar{\mu}}{a} \tag{119}$$

being the polarization, i.e., dipole moment per length (in three dimension, dipole moment per volume).

Alternatively, one may let λ be the potential drop over one unit cell,

$$\lambda = V = Ea, \tag{120}$$

whereby

$$\begin{aligned} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial \lambda \partial a} &= \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial (aE) \partial a} \\ &= \frac{\partial}{\partial a} \left[\frac{-\bar{\mu}_z}{a} + \frac{1}{E} \frac{\partial \bar{\mathcal{E}}_{\text{tot}}}{\partial a} \right] \\ &= \frac{\bar{\mu}_{z0}}{a^2} - \frac{1}{a} \frac{\partial \bar{\mu}_{z0}}{\partial a} + \frac{1}{E} \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a^2}, \end{aligned} \tag{121}$$

which is independent of \tilde{n} . Thus, in this case, the converse piezoelectric coefficient is independent of surface charges. Finally, we add that the straight, dashed lines in Fig. 11 were obtained using the expressions of Eqs. (110), (111), and (112).

The results of the model calculations suggest that the converse piezoelectric effect contains a non-negligible contribution from the surface charges. However, the model was not designed to represent any real system, and, accordingly, it is possible that this contribution is overestimated. It is, therefore, relevant to study a real material.

For that purpose, the so-called layered perovskites $R_2Ti_2O_7$ with R being Sm or Gd (Sayede et al. 2012) were considered. Although these systems possess many structural parameters, we reduced these to the two that are closely related to the spontaneous polarization (i.e., the permanent dipole moment per unit volume) that these materials possess. One of the two parameters is the lattice constant, a , in the direction of the spontaneous polarization, and the other, u , describes the distortion from the centrosymmetric structure at which the spontaneous polarization would vanish. All other structural parameters were relaxed to the optimized field-free value for a given a and u . For further details of the calculations, please consult Sayede et al. (2012).

Assuming that the second-order expansion on the right-hand side of Eq. (107) is sufficiently accurate in this case as well, one can calculate all required quantities approximately from the total energy and dipole moment per cell as functions of a and u using finite-difference approximations (see Sayede et al. 2012). The results for the (converse) piezoelectric parameters in Eq. (117) are $\eta_0 = 9.9$ (7.9) 10^{-10} (V/cm) $^{-1}$ and $\eta_1 = -7.5$ (-7.0) 10^{-11} [V/cm/(surface cell)] $^{-1}$ for R = Sm (R = Gd). Even if these numbers should be taken with some caution due to the finite-difference approximations involved in their determination, they do indicate a non-negligible effect of the surfaces. In practice, η_1 must be multiplied by the number of electrons per surface unit cell that are transferred from one surface to its opposite due, for example, to chemical modifications. For a transfer of ± 2 electrons/(surface cell), these results correspond to a change in the converse piezoelectric coefficient of about 20 % which is an effect that should be detectable experimentally.

Our treatment above can easily be generalized to the higher-dimensional case with several structural degrees of freedom. We shall assume that the number of atoms of the periodically repeated unit remains unchanged in the presence of the external field, although both the lattice vectors and the internal structural parameters may vary. The DC field vector is described by (E_1, E_2, E_3) . Let us assume that we have optimized the field-free structure. The lattice vectors for this structure can be described through six parameters $\{a_{i0}\}$, and the values of the internal structural degrees of freedom are $\{u_{n0}\}$. As before, we expand the total field-dependent energy, \mathcal{E} , around the field-free optimized structure (and zero field) as a function of geometry and field keeping terms only through second order,

$$\begin{aligned}
\bar{\mathcal{E}}_{\text{tot}} \simeq & \sum_i (a_i - a_{i0}) \frac{\partial \bar{\mathcal{E}}_{\text{tot}}}{\partial a_i} + \sum_n (u_n - u_{n0}) \frac{\partial \bar{\mathcal{E}}_{\text{tot}}}{\partial u_n} + \sum_k E_k \frac{\partial \bar{\mathcal{E}}_{\text{tot}}}{\partial E_k} \\
& + \frac{1}{2} \sum_{ij} (a_i - a_{i0})(a_j - a_{j0}) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial a_i \partial a_j} + \frac{1}{2} \sum_{nm} (u_n - u_{n0})(u_m - u_{m0}) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u_n \partial u_m} \\
& + \sum_i \sum_n (u_n - u_{n0})(a_i - a_{i0}) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial u_n \partial a_i} + \frac{1}{2} \sum_{kl} E_k E_l \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial E_k \partial E_l} \\
& + \sum_k \sum_i E_k (a_i - a_{i0}) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial E_k \partial a_i} + \sum_k \sum_n E_k (u_n - u_{n0}) \frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial E_k \partial u_n}. \quad (122)
\end{aligned}$$

In this case, the quantities $\frac{\partial^2 \bar{\mathcal{E}}_{\text{tot}}}{\partial E_k \partial a_i}$ will depend on the surface charges. However, for the higher-dimensional systems, this dependence is not through an additive lattice vector, but the additional contribution to the dipole moment per cell can take any value (cf. section ‘‘Higher Dimensions’’). Thereby, it becomes very difficult to compare calculated and measured values.

The Vector Potential and Time-(In)dependent Fields

As discussed in section ‘‘The Vector Potential Approach’’, it is also possible to use the vector potential approach, Eq. (86), for the effects of a homogeneous electrostatic field. With that approach, there is the added advantage that homogeneous, dynamic electric fields fall under the same rubric. Kirtman and coworkers have, in fact, presented a treatment for the linear and nonlinear frequency-dependent electronic responses of quasi-1D systems (Kirtman et al. 2000) based on the coupled perturbed Hartree-Fock method (CPHF). The extension to 2D and 3D systems, as well as to coupled perturbed Kohn-Sham (CPKS) time-dependent DFT, can readily be accomplished in the same way as it has already been done for the static responses (Ferrero et al. 2009; Orlando et al. 2010a, b). Indeed, an implementation in the CRYSTAL computer code (Dovesi et al. 2014) is currently well underway.

DFT Catastrophe

Finally, we shall discuss briefly an issue where the approach outlined here has given additional information. When carrying through a parameter-free electronic structure calculation, one may use either a wave function (e.g., Hartree-Fock, Møller-Plesset, CI, etc.) or density-based (Kohn-Sham) approach. In the latter case, one has to apply some approximate functional of the electron density to describe exchange and correlation effects.

It turns out that when calculating the linear and nonlinear responses to electrostatic fields [i.e., $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$, ... of Eq. (106)] for large, finite, oligomers, the values obtained are much larger when applying density-functional methods with present commonly used functionals than what is found when using Hartree-Fock and beyond wave function approaches (Champagne et al. 1998, 2000). Moreover, this overestimate increases dramatically with increasing chain length, occurs for higher-dimensional systems as well, and is larger for the higher-order responses. These problems of the density-functional calculations have been called the “DFT catastrophe.”

If one is interested in obtaining the thermodynamic limit, it is essential that sufficiently large systems are considered so that the systems contain a central region of identical, neutral unit cells. As we have found in our model calculations, this requirement often implies that finite chains of some 50 units are treated. For higher-dimensional systems, this number becomes correspondingly larger.

By having an approach for treating the infinite, periodic system exposed to an electrostatic field, it is possible to treat the thermodynamic limit directly. Moreover, the formulation of the responses to the field is different for the infinite, periodic system than for the large, finite system, although, of course, the same numerical values must be obtained in the two cases.

Using a simple model system, we were able to study both the large, finite system and the infinite, periodic system both with Hartree-Fock and with density-functional calculations (Vargas et al. 2014). The approach allowed us also to decompose the responses into cell contributions, whereby it could be seen that the higher-order responses were more delocalized. Moreover, the results suggested that so-called range-separated density-functional methods, despite a large sensitivity to the range separation parameter, could be those that will be able to remove (the largest parts of) the DFT catastrophe. However, further improvements are necessary considering some of the weaknesses that have already been identified (Bulik et al. 2013; Kirtman et al. 2008).

Conclusions

In this overview, we have described our current understanding of how to treat large, regular systems that are exposed to electrostatic fields. Such systems are best treated as being infinite and periodic, and we demonstrated that, even when the electrostatic fields are included, it is possible to derive a formulation where the translational symmetry is retained. However, the formulation is not trivial and involves operators that contain derivatives with respect to the k -vector. Consequently, special techniques had to be developed in order to calculate directly the responses – particularly, the dipole moment and related properties – to electrostatic fields.

An interesting feature is that the responses of the large, finite systems depend on the surfaces/terminations. Although these are not explicitly included in the infinite, periodic systems, their effects are nevertheless accessible in periodic system calculations. Here, however, an open issue remains: it is possible to understand that

different, periodic systems correspond to modeling samples of different shape, but it is not (yet) possible to describe precisely how a given shape can be described within the infinite, periodic system treatment.

Although our work has revealed a number of interesting, fundamental aspects, it is also highly relevant to apply our approach to real materials in order to determine their linear and nonlinear, electronic, and structural responses to electrostatic fields. Therefore, we are planning to implement our approach in established computer programs so that it can be broadly applied.

Acknowledgements The work has benefited from financial support through the German Research Council (DFG). Moreover, one of the authors (MS) is very grateful to the International Center for Materials Research, University of California, Santa Barbara, for generous hospitality. Finally, this work has benefited from collaborations with Emilie Bruyer, Yi Dong, Mauro Ferrero, Michel Rérat, Adlane Sayede, Violina Spurk (née Tevekeliyska), Jorge Vargas, Lorenzo Maschio, Roberto Orlando, and Roberto Dovesi.

Bibliography

- Baroni, S., de Gironcoli, S., Dal Corso, A., & Giannozzi, P. (2001). Phonons and related crystal properties from density-functional perturbation theory. *Reviews of Modern Physics*, *73*, 515–562.
- Berger, J. A., de Boeij, P. L., & van Leeuwen, R. (2005). A physical model for the longitudinal polarizabilities of polymer chains. *The Journal of Chemical Physics*, *123*, 174910.
- Bishop, D., Gu, F. L., & Kirtman, B. (2001). Coupled-perturbed Hartree-Fock theory for infinite periodic systems: Calculation of static electric properties of $(\text{LiH})_n$, $(\text{FH})_n$, $(\text{H}_2\text{O})_n$, $(-\text{CNH}-)_n$, and $(-\text{CH}=\text{CH}-)_n$. *The Journal of Chemical Physics*, *114*, 7633–7643.
- Blount, E. I. (1962). Formalisms of band theory. *Solid State Physics*, *13*, 305–373.
- Bulik, I. W., Zalesný, R., Bartkowiak, W., Luis, J. M., Kirtman, B., Scuseria, G. E., Avramopoulos, A., Reis, H., & Papadopoulos, M. G. (2013). Performance of density functional theory in computing nonresonant vibrational (hyper)polarizabilities. *Journal of Computational Chemistry*, *34*, 1775–1784.
- Champagne, B., Jacquemin, D., André, J.-M., Kirtman, B. (1997). Ab initio coupled Hartree-Fock investigation of the static first hyperpolarizability of model all-trans-polymethineimine oligomers of increasing size. *The Journal of Physical Chemistry A*, *101*, 3158–3165.
- Champagne, B., Perpète, E. A., van Gisbergen, S. J. A., Baerends, E.-J., Snijders, J. G., Soubra-Ghaoui, C., Robins, K. A., & Kirtman, B. (1998). Assessment of conventional density functional schemes for computing the polarizabilities and hyperpolarizabilities of conjugated oligomers: An ab initio investigation of polyacetylene chains. *The Journal of Chemical Physics*, *109*, 10489–10498.
- Champagne, B., Perpète, E. A., Jacquemin, D., van Gisbergen, S. J. A., Baerends, E.-J., Snijders, J. G., Soubra-Ghaoui, C., Robins, K. A., & Kirtman, B. (2000). Assessment of computational density functional schemes for computing the dipole moment and (hyper)polarizabilities of push-pull pi-conjugated systems. *The Journal of Physical Chemistry A*, *104*, 4755–4766.
- Darrigan, C., Rérat, M., Mallia, G., & Dovesi, R. (2003). Implementation of the finite field perturbation method in the CRYSTAL program for calculating the dielectric constant of periodic systems. *Journal of Computational Physics*, *24*, 1305–1312.
- de Gironcoli, S., Baroni, S., & Resta, R. (1989). Piezoelectric properties of III-V semiconductors from first-principles linear-response theory. *Physical Review Letters*, *62*, 2853–2856.
- Dovesi, R., Orlando, R., Erba, A., Zicovic-Wilson, C. M., Civalieri, B., Cassasa, S., Maschio, L., Ferrabone, M., De la Pierre, M., D'Arco, P., Noël, Y., Causà, M., Rérat, M., & Kirtman, B.

- (2014). CRYSTAL14: A program for the ab initio investigation of crystalline solids. *International Journal of Quantum Chemistry*, 114, 1287–1317.
- Dvornikov, M. (2007). Formulae of numerical differentiation. arXiv:math.NA/0306092 (unpublished). *Journal of Concrete and Applicable Mathematics*, 5, 77–88.
- Ferrero, M., Rérat, M., Kirtman, B., & Dovesi, R. (2009). Calculation of first and second static hyperpolarizabilities of one- to three-dimensional periodic compounds. Implementation in the CRYSTAL code. *The Journal of Chemical Physics*, 129, 214110.
- Genkin, V. N., & Mednis, P. M. (1968). Contribution to the theory of nonlinear effects in crystals with account taken of partially filled bands. *Soviet Physics JETP*, 27, 609–615.
- Khan, I. R., & Ohba, R. (2000). New finite difference formulas for numerical differentiation. *Journal of Computational and Applied Mathematics*, 126, 269–276.
- King-Smith, R.-D., & Vanderbilt, D. (1993). Theory of polarization of crystalline solids. *Physical Review B* 47, 1651–1654.
- Kirtman, B., Gu, F. L., & Bishop, D. (2000). Extension of the Genkin and Mednis treatment for dynamic polarizabilities and hyperpolarizabilities of infinite periodic systems. I. Coupled perturbed Hartree-Fock theory. *The Journal of Chemical Physics*, 113, 1294–1309.
- Kirtman, B., Bonness, S., Ramirez-Solis, A., Champagne, B., Matsumoto, H., & Sekino, H. (2008). Calculation of electric dipole (hyper)polarizabilities by long-range-correction scheme in density functional theory: A systematic assessment for polydiacetylene and polybutatriene oligomers. *The Journal of Chemical Physics*, 128, 114108.
- Kirtman, B., Ferrero, M., Rérat, M., & Springborg, M. (2009). Polarization of periodic systems in a static electric field. Sawtooth potential treatment revisited. *The Journal of Chemical Physics*, 131, 044109.
- Kudin, K. N., Car, R., & Resta, R. (2005). Longitudinal polarizability of long polymeric chains: Quasi-one-dimensional electrostatics as the origin of slow convergence. *The Journal of Chemical Physics*, 122, 134907.
- Kudin, K. N., Car, R., & Resta, R. (2007). Quantization of the dipole moment and of the end charges in push-pull polymers. *The Journal of Chemical Physics*, 127, 194902.
- Kunc, K., & Resta, R. (1983). External fields in the self-consistent theory of electronic states: A new method for direct evaluation of macroscopic and microscopic dielectric response. *Physical Review Letters*, 51, 686–689.
- Li, J. (2005). General explicit difference formulas for numerical differentiation. *Journal of Computational and Applied Mathematics*, 183, 29–52.
- Maschio, L., Kirtman, B., Orlando, R., & Rérat, M. (2012). Ab initio analytical infrared intensities for periodic systems through a coupled perturbed Hartree-Fock/Kohn-Sham method. *The Journal of Chemical Physics*, 137, 204113.
- Maschio, L., Kirtman, B., Rérat, M., Orlando, R., & Dovesi, R. (2013). Ab initio analytical Raman intensities for periodic systems through a coupled perturbed Hartree-Fock/Kohn-Sham method in an atomic orbital basis. I. Theory. *The Journal of Chemical Physics*, 139, 164101.
- Nunes, R. W., & Gonze, X. (2001). Berry-phase treatment of the homogeneous electric field perturbation in insulators. *Physical Review B*, 63, 155107.
- Orlando, R., Bast, R., Ruud, K., Ekström, U., Ferrabone, M., Kirtman, B., & Dovesi, R. (2010a). The first and second static electronic hyperpolarizabilities of zigzag boron nitride nanotubes. An ab initio approach through the coupled perturbed Kohn-Sham scheme. *The Journal of Physical Chemistry A*, 115, 12631–12637.
- Orlando, R., Lacivita, V., Bast, R., & Ruud, K. (2010b). Calculation of the first static hyperpolarizability tensor of three-dimensional periodic compounds with a local basis set: A comparison of LDA, PBE, PBE0, B3LYP, and HF results. *The Journal of Chemical Physics*, 132, 244106.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in Fortran 77: The art of scientific computing*. Cambridge/New York: Cambridge University Press.
- Resta, R. (1994). Macroscopic polarization in crystalline dielectrics: The geometric phase approach. *Reviews of Modern Physics*, 66, 899–915.
- Resta, R. (1998). Quantum-mechanical position operator in extended systems. *Physical Review Letters*, 80, 1800–1803.

- Sayed, A., Bruyer, E., & Springborg, M. (2012). Ab initio study of metastable layered perovskites $R_2Ti_2O_7$ ($R = Sm$ and Gd). *Physical Review B*, *86*, 125136.
- Schmidt, K., & Springborg, M. (1999). Static (hyper-)polarizabilities of infinite conjugated polymers in the framework of density-functional theory. *Physical Chemistry Chemical Physics*, *1*, 1743–1748.
- Sebastiani, D., & Parrinello, M. (2003). A new ab-initio approach for NMR chemical shifts in periodic systems. *The Journal of Physical Chemistry A*, *105*, 1951–1958.
- Springborg, M., & Kirtman, B. (2008). Analysis of vector potential for calculating linear and nonlinear responses of infinite periodic systems to a finite static external field. *Physical Review B*, *77*, 045102.
- Springborg, M., & Kirtman, B. (2011). On the dipolar electric field response of large systems. *Theoretical Chemistry Accounts*, *130*, 687–700.
- Springborg, M., Kirtman, M., & Dong, Y. (2004). Electronic polarization in quasilinear chains. *Chemical Physics Letters*, *396*, 404–409.
- van Gisbergen, S. J. A., Koostra, F., Schipper, P. R. T., Gritsenko, O. V., Snijders, J. G., & Baerends, E. J. (1998). Density-functional-theory response-property calculations with accurate exchange-correlation potentials. *Physical Review A*, *57*, 2556–2571.
- Vanderbilt, D., & King-Smith, R.-D. (1993). Electric polarization as a bulk quantity and its relation to surface charge. *Physical Review B*, *48*, 4442–4455.
- Vargas, J., Springborg, M., & Kirtman, B. (2014). Electronic responses of long chains to electrostatic fields: Hartree-Fock vs. density-functional theory: A model study. *The Journal of Chemical Physics*, *140*, 054117.
- Weniger, E. J., & Kirtman, B. (2003). Extrapolation methods for improving the convergence of oligomer calculations to the infinite chain limit of quasi-one-dimensional stereoregular polymers. *Computers & Mathematics with Applications*, *45*, 189–215.

Hande Toffoli, Sakir Erkoç, and Daniele Toffoli

Contents

Introduction	1460
Hierarchy of Methods	1460
0D Structures: Nanoparticles	1461
Global Optimization with Empirical Potentials	1464
Local Optimization with Higher Accuracy Methods	1466
1D Structures: Nanotubes, Nanowires, and Nanorods	1475
Elastic and Structural Properties	1476
Structural Properties	1479
Electronic, Magnetic, and Optical Properties	1481
2D Structures: Graphene and Derivatives	1485
Graphene, Nanosheets, Nanoribbons, Nanobelts, and Nanostrips	1486
Electronic and Mechanical Properties	1492
Magnetic and Optical Properties	1500
Adsorption Phenomena	1504
Bibliography	1506

Abstract

Materials properties show a dependence on the dimensionality of the systems studied. Due to the increased importance of surfaces and edges, lower-dimensional systems display behavior that may be widely different from their bulk counterparts. As a means to complement the newly developed experimental methods to study these reduced dimensional systems, a large fraction of

H. Toffoli (✉) • S. Erkoç

Department of Physics, Middle East Technical University, Ankara, Turkey

e-mail: hande@newton.physics.metu.edu.tr; ustunel@metu.edu.tr; erkoc@metu.edu.tr

D. Toffoli

Department of Chemistry, Middle East Technical University, Ankara, Turkey

e-mail: dtoffoli@units.it

the theoretical effort in the field continues to be channeled toward computer simulations. This chapter reviews briefly the computational methods used for the low-dimensional materials and presents how the materials properties change with dimensionality. Low-dimensional systems investigated are classified into a few broad classes: 0D nanoparticles, 1D nanotubes, nanowires, nanorods, and 2D graphene and derivatives. A comprehensive literature will guide the readers' interest in computational materials sciences.

Introduction

Hierarchy of Methods

Methods for materials modeling are based on the particular length and time scales used for the investigation of various materials properties that operate only over those scales. This has led to basically four independent methodological streams, which may be categorized as ab initio density functional theory, molecular dynamics, statistical methods based on Monte Carlo algorithms, and continuum mechanics (see Fig. 1 for the ranges of them). Each of these is computationally intensive within its own range. However, the expanding capabilities of computational methods due to the increasing power of computers and continuing development of efficient algorithms, together with advances in the synthesis, analysis, and visualization of

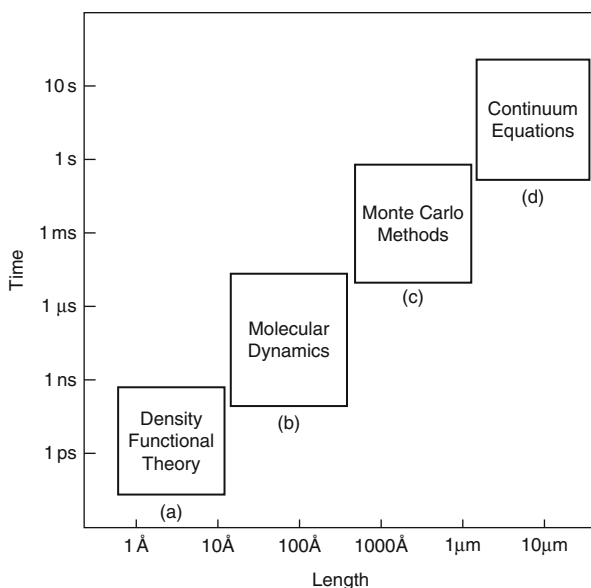


Fig. 1 Ranges of scales used for different methods in materials properties (Adapted from Vvedensky 2004)

materials at increasingly finer spatial and temporal resolutions, have spawned a huge effort in the modeling of materials phenomena.

Since the materials properties are directly related with the dimensionality of materials, particularly at the nanoscale, it would be better to classify the methods of modeling of nanostructures with respect to the dimensionality, such as zero-dimensional (0D), one-dimensional (1D), two-dimensional (2D), and three-dimensional (3D) materials. At the nanoscale the first three dimensionalities (0D, 1D, 2D) are commonly used, but the 3D materials are usually considered as macroscopic systems; therefore those materials are not the subject of this chapter.

0D Structures: Nanoparticles

The class of nanomaterials that may be termed *zero-dimensional* comprises systems that are confined within up to several hundreds of nanometers in all three dimensions. Although there exists no clear-cut size threshold at which a system switches from a zero-dimensional system to bulk, there is a rather well-defined class of systems that fit the above definition with unique and intriguing properties. The most commonly studied zero-dimensional systems are *quantum dots*, *nanoparticles* (or *clusters*), and *cage-like structures*. In this section, we shall begin with an overview of methods used to study such materials.

After the 1980s, the field of nanoparticles/Clusters has seen increasing activity. The results of these studies and their applications have been collected in several reviews (Balasubramanian 1990; Baletto and Ferrando 2005; Bonacic-Koutecky et al. 1991; Brack 1993; de Heer 1993; Halicioglu and Bauschlicher 1988; Morse 1986; Ustunel and Erkoc 2007; Weltner and Van Zee 1984; Oymak and Erkoc 2012). Studies on clusters have been concentrated on such diverse applications as catalysis, chemisorption and substrate adsorption, laser applications, magnetism, nucleation, photographic processes, and reactivity.

The term *cluster* may, in the broadest sense, be defined as an aggregate of atoms or molecules that form a size regime between molecules and bulk materials. It is far from a straightforward task to define the range of number of atoms which constitute a cluster. In contrast to their bulk counterparts, structural and electronic properties of clusters display a rather strong dependence on size and geometric configuration due to their large surface-to-volume ratio. Thus in order to understand the properties of clusters, a careful determination of the equilibrium structure is crucial. In general, the equilibrium geometry is determined by locating the global minimum of the *potential energy surface* (PES) of a cluster, which in principle can be mapped out by calculating the total energy as a function of all possible locations of the atoms. Unfortunately even the smallest clusters often have extremely complicated potential energy surfaces with practically countless local minima. This requires a thorough search over the entire PES for a realistic determination of the equilibrium structure (Farrow et al. 2014).

An added complication to the difficulty of determining the global minima is that clusters with different but similar sizes of the same material may present PESs that

are entirely different. It is likely then that the equilibrium shape of a cluster of a given material and size is significantly different from a different size cluster of the same material. Fortunately, some similarities may be found in the vast number of possible structures. Mass spectroscopy of clusters reveals that certain sizes of clusters are preferred over others, indicated by the high abundance detected (Martin 1996). The number of atoms or molecules in a preferred cluster is commonly referred to as a *magic number*. The origin of a magic number may be the completion of a *geometric* or an *electronic* shell of the cluster. The particular geometric or electronic shell being completed obviously depends upon the constituents of the clusters and the bonding between them. For instance, noble gases are found in polyhedral form, whereas metallic clusters prefer cubic or prolate structures (Johnston 2002). The search for magic numbers has been the subject of a very large number of works (Anagnostatos 1987; Balasubramanian 1990; Bonacic-Koutecky et al. 1991; Haberland 1994; Jarrold and Constant 1991; Moraga 1993; Yang et al. 2006).

In order to map out the energy landscape of nanoparticles, an adequate model for the interatomic interactions within the cluster needs to be employed. Empirical potentials, the tight-binding approximation, and density functional theory whose brief description has been given in previous sections are the most commonly used methods in literature. While empirical potentials (Erkoc 1997, 2001a) offer an efficient means of exploring the PES and determining the equilibrium structure of large clusters, calculations based on the tight-binding approximation and density functional theory are preferred for more accurate and detailed description of geometrical and electronic properties of small- to medium-sized clusters.

Once the method of choice is determined for interatomic interactions, the corresponding potential energy surface may be explored using *global search algorithms* to determine the lowest lying minima. A concise review of global optimization methods is given in the reviews (Baletto and Ferrando 2005; Dugan and Erkoc 2009; Wales and Doye 1997; Erkoc et al. 2003; Catlow et al. 2010). Here, we shall briefly describe a few of them.

The *basin-hopping algorithm* (Wales and Doye 1997) belongs to a class of optimization methods called *hypersurface deformation methods*. These methods aim to simplify the PES by applying a map that smooths it out. The global minimum of the simplified PES is then mapped back to that of the real PES assuming that it leads back to the global minimum of the real PES. In the basin-hopping algorithm, the original, complicated PES, $U(\vec{r})$, is transformed onto a steplike function $\tilde{U}(\vec{r})$ according to the rule (Baletto and Ferrando 2005; Wales and Doye 1997).

$$\tilde{U}(\vec{r}) = \min [U(\vec{r})], \quad (1)$$

where the min function implies that a local minimization is to be applied such as the conjugate gradients algorithm (Hestenes and Stiefel 1952).

In many of the cluster applications using basin hopping (Bromley and Flikkema 2005; Doye and Wales 1999; Hsu and Lai 2006; Wales and Doye 1997), this transformed PES is sampled with a Monte Carlo algorithm, sometimes with a fixed acceptance (Wales and Doye 1997).

The second method we shall mention is *the simulated annealing method* (Kirkpatrick et al. 1983), in which the system is evolved at high temperature and then gradually cooled down. Assuming that the system does not get trapped in a basin of attraction which is not the minimum, it reaches the state of lowest energy during the cooling process. Attempts to use simulated annealing to find the global energy minimum in larger systems are frustrated by high-energy barriers that trap the system in metastable configurations (Ma and Wang 2006).

Finally, we briefly describe a class of algorithms that has recently gained popularity, namely, *genetic algorithms*. Genetic algorithms (GAs) are based on ideas borrowed from natural evolution. As explained in Darby and coworkers (2002), they employ operators that are analogues of the evolutionary processes of genetic crossover, mutation, and natural selection. A GA can be applied to any problem where the variables to be optimized can be encoded to form a string, each string representing a trial solution to the system. In cluster terminology each *string* is a list of coordinates of a cluster. An initial number of clusters are chosen according to some generation scheme (usually random) and locally minimized.

After the creation of the initial generation of strings, each member of this generation is assigned a *fitness parameter* based on a fitness function. In Darby and coworkers (2002), for example, this function was chosen to be the negative exponent of the deviation of the total energy of each cluster from the cluster with minimum energy.

Once a fitness parameter is assigned to each member of the population, a new generation is created: Two *parents* are chosen from members with high fitness parameter for each *offspring* (member of the new generation) and then subjected to the *crossover* operation. A crossover operation for clusters consists of a particular rearrangement and combination of the atomic coordinates in the parent clusters. For an example, see Darby and coworkers (2002). After each new cluster is generated, a local minimization is performed. Mating continues in this way until the desired number of offsprings are generated. New fitness parameters are assigned to each offspring. In order to maintain diversity, new genetic material is introduced into the population by means of the *mutation* operator. After mutation in a cluster population, clusters are once again locally minimized.

Finally the next generation is completed by applying *natural selection*. The previous generation and the new generation are ranked in order of fitness and the best N individuals (clusters) are chosen, where N is the population of the original population. These N individuals constitute the new generation, and mating is performed again.

The process of mating, mutation, and selection is repeated for a predetermined number of generations or until a given criterion is reached. This criterion could be, for example, the convergence of the highest fitness to a plateau so that consecutive generations do not produce better results.

Other global search algorithms include parallel random tunneling, conformational space annealing, greedy search method, simulated annealing, quantum annealing, smoothing and hypersurface deformation techniques, lattice methods, growth sequence analysis, and replica exchange method (Oymak and Erkoc 2004).

Global Optimization with Empirical Potentials

In spite of the wealth of information they provide, global search algorithms ordinarily fail to identify all the minima of a given cluster. The best explored potential surface for clusters belongs to possibly the simplest empirical potential, namely, the Lennard-Jones potential. The Lennard-Jones potential is a simple model that captures the long- and short-range behavior of atoms and molecules. It was proposed in 1931 by Lennard-Jones (1931) and has been used in innumerable studies ever since. It has the following simple form

$$V(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (2)$$

where ϵ and σ are parameters that determine the depth and the width of the potential well, respectively, and r is the interatomic distance between atom pairs.

The four known structural forms for the LJ potential are (Hartke 2002) as seen in Fig. 2:

1. Icosahedral: formed by starting from the pentagonal bipyramid and adding a further layer of atoms

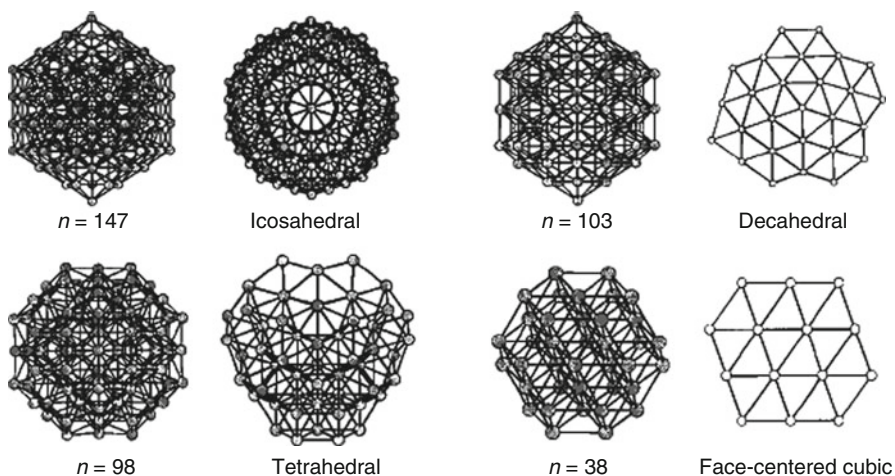


Fig. 2 The four basic structures of Lennard-Jones clusters (Reprinted with permission from Hartke (2002). © (2002) by Wiley-VCH Verlag GmbH & Co. KGaA)

2. Decahedral: icosahedral core surrounded by fcc stacks
3. Tetrahedral: tetrahedron of atoms in the innermost core
4. Face-centered cubic: sections of the fcc bulk structure

For $N < 1,000$, Lennard-Jones clusters follow an icosahedral pattern growth with magic numbers corresponding to *Mackay icosahedra* (Mackay 1962) for $N = 13, 55, 147, 309$, etc. In between these magic numbers, most of the structures are Mackay-like with incomplete outer layers. Exceptions occur when there are alternative structures with complete shells. These are mostly Marks decahedra (Doye 2003), but there are instances of an fcc truncated octahedron and a Leary tetrahedron (Noya and Doye 2006). The preference for icosahedral structures of Lennard-Jones clusters at small sizes is thought to be due to a trade-off between optimal bond distance and strain (Hartke 2002; Krainyukova 2006).

In spite of its simplicity, the Lennard-Jones potential may not be used for accurate description of cluster properties with the possible exception of noble gas clusters. Often empirical potentials that involve a larger number of parameters than the two-parameter Lennard-Jones potential are employed for structural determination (Farrow et al. 2014).

To understand their behavior at finite temperature, Wu and coworkers studied Zn clusters using the Gupta potential by means of molecular dynamics (Wu et al. 2005). Zn clusters are found to display peculiar structures that are very different from other sp-type metal clusters. The $N = 13$ cluster, for instance, has a disordered structure instead of the icosahedral structure found in most metals of the same size. The C_v curve also differs in that it has a double peak structure, whereas icosahedral structures display a single peak. The ΔE_2 study in Wu et al. (2005) reveals that magic structures are found at $N = 6, 13, 29, 33, 36, 42, 49$, and 53.

In a study by Rey and coworkers (1993), several transition metal clusters (Ni, Pd, Au, and Ag) were studied using different model potentials. Three different *embedded atom models* (EAM) and two different parameterizations of an N-body analytical potential were considered. For Ni clusters, although the binding energies provided by the various models showed differences, the peak stability was obtained for all the models at $N = 13$ (icosahedron) and $N = 19$ (double icosahedron). In this work (Rey et al. 1993), earlier experiments on large Ni clusters ($N = 49$ –2300) are also quoted where icosahedral growth was observed. For Ni the EAM that was fitted to bulk values (EAM1) gave poorer results than the one fitted to diatomic values. For Pd, Au, and Ag, only EAM1 and one of parameterizations of the analytical potential (TBM1) were compared. EAM1 failed to yield enhanced stability for $N = 13$ and $N = 19$, while TBM1 mostly gave icosahedral structures for both sizes. This shows the importance of parameterization for the accuracy of the model potentials.

In a recent study by Doye (2006), results for Pb clusters were calculated and compared for the Gupta potential and a glue potential in combination with a basin-hopping algorithm. Within the Gupta potential framework, Pb clusters, in contrast to Si, Ge, and Sn, do not show any tendency to form prolate or oblate structures. Their structures are mostly spherical with decahedra being the most favored structure

followed by close packed. This study also presents an interesting comparison between the Gupta potential and the glue potential, which reveals that there is virtually no agreement between the two potentials neither in the magic numbers nor in the resulting structures. This sharp discrepancy demonstrates that one should be very careful while comparing results from different methods. More disordered structures are favored for the glue potential than the Gupta potential because the embedding term dominates for the latter.

An interesting result from the work of Doye (2006) is that for the Gupta potential, the uncentered Mackay icosahedron with $N = 54$ is more stable than the complete $N = 55$ Mackay icosahedron. This is because of the strain applied by the external atoms on the center, making it unfavorable to be occupied.

As a final remark on empirical potentials, we note that the particular parameterization of the potential plays a crucial role in correctly identifying structural properties, and caution should therefore be used in interpreting results for empirical potentials. For instance, according to Michaelian et al. (1999), the global minima of Au clusters are difficult to localize because of the short range of the Au potential. As the range of a potential decreases, the number of global minima associated with the potential increases. As a result of this, different studies have identified different global minima for Au clusters. In addition different potential models for the metal yield different ordering of the lowest lying minima.

Local Optimization with Higher Accuracy Methods

Outside of the domain of simple empirical potentials, identifying the global minima of cluster PESs becomes a prohibitively demanding task. Therefore, most studies on the theoretical determination of cluster minima employ methods that either focus on local minima obtained through an adequate initial guess or a mixture of global search algorithms with simple potentials to reduce the number of minima followed by local minimization techniques that are more accurate.

The majority of elements, regardless of whether they prefer atomic shells or electronic shells, form clusters that are more or less spherical. The only elements that deviate from this behavior are the group IVA elements. This is especially remarkable in carbon, which goes from linear chains all the way to fullerenes. Clusters of Si and Ge first grow in one dimension and then abruptly change to form spherical clusters (Shvartsburg and Jarrold 1999). Some similarities exist between the growth patterns of Si and Ge cluster but major differences also occur especially concerning the size at which transition to bulk-like structures takes place (Shvartsburg and Jarrold 1999).

Because Si clusters undergo such a drastic change in cross section, an excellent experimental technique exists for studying the structure as a function of size. This technique (Jarrold and Constant 1991) is based upon the difference in the mobility of cluster ions that have different shapes. The ion mobility of a cluster depends on its rotationally averaged collision cross section, which depends on its size. Prolate and oblate geometries have a larger cross section and therefore smaller mobilities (Sieck

et al. 2003). Mobility studies indicate that Si_n clusters undergo a phase transition at about $n = 24\text{--}34$ from elongated to spherical, whereas for the Ge clusters, this transition size is between $n = 64$ and $n = 74$ (Jarrold and Constant 1991; Shvartsburg and Jarrold 1999).

The growth sequence of Si is realized by adding on small building blocks to smaller structures. In particular, Si_9 (Shvartsburg and Jarrold 1999) and Si_{10} (Baletto and Ferrando 2005) have enhanced stability and, therefore, act as the fundamental building blocks. On the other hand, medium-sized Si clusters are quasispherical but not crystalline. Crystallinity does not become apparent until N gets as large as a few hundred atoms (Baletto and Ferrando 2005).

In the simulated annealing and DFT-based tight-binding work by Sieck and coworkers (2003), a *prolateness parameter* is defined as follows

$$p = \sum_{j=1}^3 (I_j - \hat{I}) = -\frac{I_1}{2} + I_2 - \frac{I_3}{2}, \quad (3)$$

where $0 < I_1 \leq I_3$ and $\hat{I} = (I_1 + I_3)/2$. For an oblate structure, $p < 0$; for a prolate structure, $p > 0$; and for a spherical structure, $p = 0$. Sieck and coworkers (2003) studied various different isomers corresponding to three fixed sizes, namely, $N = 25$, 29, and 35. In Fig. 3, the cohesive energy versus the prolateness parameter for the low-lying isomers is displayed. An investigation into the most stable structures clearly indicates the tendency to go from elongated to spherical structures as the clusters grow in size.

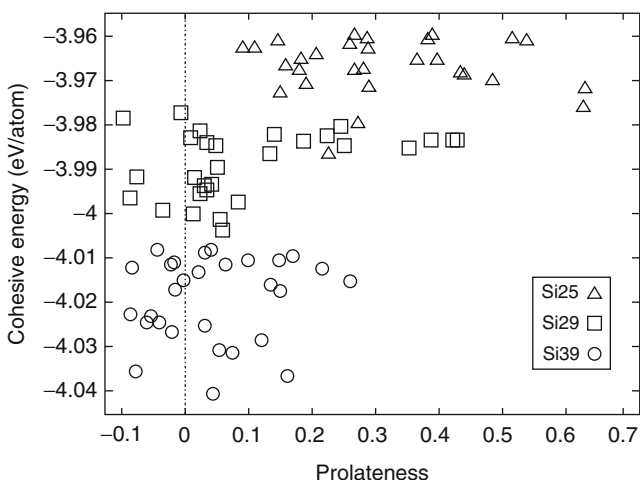


Fig. 3 Cohesive energy versus prolateness parameter defined in Eq. 3 for 30 lowest lying isomers of Si_{25} , Si_{29} , and Si_{35a} (Reprinted with permission from Sieck et al. (2003). © (2003) by Wiley-VCH)

An important point is noted by Sieck and coworkers (2003), that different methods, even different flavors of the same method (LDA/GGA in DFT), may disagree about the exact energetic ordering of clusters; however, usually the lowest energy structures are local minima in all of them.

As an example to the multiscale methods mentioned above, Bulusu and coworkers (2005) used a basin-hopping algorithm combined with DFT to explore low-lying minima for Ge_{12} – Ge_{20} . Global minima for the Ge clusters in this size range are compared to minima of Si clusters. Various checks were performed to make sure different starting points yield the same global minima, such as different seeding patterns. Unlike most studies in the literature, zero-point motion is taken into account while calculating the binding energies.

In the same work (Bulusu et al. 2005) of Si and Ge, clusters in the size range 12–20 were compared and clusters of both Si and Ge were shown to have prolate geometries. However, their growth patterns were found to diverge at $N = 13$. The global minima for Ge clusters of size 12–16 are obtained by adding atoms to the Ge_9 tetracapped trigonal prism (TTP) structure familiar from Si clusters (Bulusu et al. 2005). For low energy-Si clusters, TTP-to-sixfold puckered ring (six/six) transition occurs at $N = 16$ and clusters with higher number of atoms all contain the six/six pattern. In contrast, TTP-to-six/six pattern may occur at $N = 19$ for Ge clusters, and at $N = 20$, the magic number cluster Ge_{10} appears to be the preferred structural motif.

According to Shvartsburg and Jarrold (1999), clusters of Sn up to $n \approx 35$ follow the trend of germanium and prolate silicon clusters (Shvartsburg and Jarrold 1999). For $N > 20$, the mobility of Sn clusters shows larger fluctuations than either Si or Ge. This might indicate the presence of multiple isomers of Sn clusters in this size range (Shvartsburg and Jarrold 1999). Even though the $\alpha \Rightarrow \beta$ transition occurs between 286 and ≈ 310 K in the bulk, Sn clusters do not undergo such a transition even at higher temperatures. In fact, mobility measurements show that Sn clusters do not show significant changes in structure for a very broad temperature range. Transition from prolate to spherical growth in Sn clusters is not abrupt like the transition in Si and Ge clusters but occurs in steps. Clusters of $N \leq 35$ adopt a stacked prolate morphology much like Si and Ge clusters. This is unexpected because these highly noncompact structures are suitable for covalent materials, whereas bulk Sn under ambient conditions is a metal. In a sense, the covalent-to-metal transition that occurs between the fourth and fifth row of the periodic table for the carbon series in the bulk fails to occur in their clusters. Shvartsburg and Jarrold (2000) also find that Pb_n^- and Pb_n^+ clusters display different magic numbers.

Going down the group IV in the periodic table, a comprehensive set of data is given by Shvartsburg and Jarrold (2000). Their findings reveal that Pb clusters are structurally different from Si, Ge, and Sn cluster for $N < 25$. While the latter clusters have low mobilities in this size range indicating a prolate structure, Pb clusters exhibit much higher mobilities meaning they are quasispherical in shape. For $N > 25$, Si clusters also display mobilities in accord with a spherical shape. However, their mobilities are considerably smaller than the Pb clusters in the same

size range. This is attributed to the densely packed nature of the Pb clusters in contrast with the open, cage-like arrangement of the Si clusters.

Transition metals form the largest group of elements considered in this section. Even as bulk materials, they exhibit very interesting and diverse properties, in particular magnetism. Interesting questions therefore arise concerning whether such properties as magnetism are maintained or altered in clusters (Briere et al. 2002; Kabir et al. 2006). Appearance of magnetism in the clusters of $4d$ elements (such as Ru, Rh, and Pd) is a very interesting phenomenon as these elements are nonmagnetic in the bulk. Clusters of $3d$ elements (such as Fe, Co, and Ni), which are already magnetic in the bulk, exhibit enhanced magnetic moments in the cluster form due to narrower band widths and the increased localization of the electrons (Kumar and Kawazoe 2002; Pawluk et al. 2005). Indeed, a recent Stern–Gerlach experiment revealed that Mn clusters in the range $N = 11$ –99 display ferromagnetic ordering even though no such ordering is observed in the bulk phase (Knickelbein 2001). These results reveal PES minima at $N = 13$ and $N = 19$ and PES maxima at $N = 15$ and $N = 23$ –25.

We should mention here that most of the magnetism studies mentioned in this section take into account only the electronic spin contribution to the magnetic moment. This can be done only in the cases where the spin-orbit coupling can be neglected (Bansmanna et al. 2005).

In the work by Rodríguez-López and coworkers (2003), Co clusters in the range $N = 4$ –60 were studied by means of an evolutive algorithm based on the Gupta potential and tight binding. Experiments reveal that much like the Cr clusters (Payne et al. 2006), different isomers of Co clusters coexist with distinct magnetic moments. In this work two sets of isomers are identified for each size – the lowest and the second-lowest lying. For the lowest lying isomers, an icosahedral growth is observed with structures derived by adding atoms to the main icosahedral sizes at $N = 7, 13, 19, 23, 26, 34, 43,$ and 55 . For the second isomers, no particular growth pattern was identified. The stability of these sizes was also confirmed by the second energy difference ΔE_2 in addition to other intermediate icosahedral structures.

Relative thermodynamic populations of the lowest energy isomers were used to simulate possible experimental conditions. Isomers coexist particularly evenly between sizes of enhanced stability. This is due to the influence of the entropic contribution of the low-frequency normal modes of the isomers to the free energy. For both the global minimum and the second isomer, a nonmonotonic decrease is observed with increasing size. The greatest difference between the two sets of clusters arises in the range $N = 20$ –40, which corresponds to the range where the average interatomic distance and average coordination of the two sets show significant difference. In this size range, two effects seem to compete for determining the magnetization of the two sets of configurations. On the one hand, the average coordination is higher for the global minima, which should result in lower magnetic moments for the global minima. On the other hand, the average nearest-neighbor distance is higher for the global minima, which should yield higher magnetic moments. The results indicate that the average coordination number effect dominates.

For noble and transition metals, the interactions between atoms are not pairwise and simple empirical potentials are inappropriate (Barreateau et al. 2000). Therefore incorporating many body effects into the potentials is essential. Moreover, for magnetism studies, *ab initio* methods need to be employed, which render global optimization efforts extremely computation intensive. Therefore, most results we shall quote here will be based on restricted searches of the potential energy surface.

In the work by Barreateau and coworkers (2000), for instance, the relative stability of cuboctahedra and Mackay icosahedra is determined for Rh and Pd clusters for $N = 13, 55, 147, 309,$ and 561 using a tight-binding method. Since both structures have an identical sequence of magic numbers, it is interesting to determine the transition size. A continuous transition is possible from the Mackay icosahedron to the cuboctahedron, and for the Rh and Pd clusters, this pathway is explored. An analysis of the Mackay transition from the cuboctahedron to the icosahedron reveals that for $N = 13$, the cuboctahedron is unstable for both Rh and Pd, becoming metastable for larger sizes with an increasing activation barrier with size. The magnetic moment of rhodium was found to disappear for sizes more than 100 atoms, and palladium clusters were found to be hardly magnetic.

In Mn clusters, on the other hand, magnetism plays an important role in determining the ground state structures. As mentioned in the work by Briere and coworkers (2002), many spin isomers can lie close in energy. In this work, a few local geometric configurations of $N = 13, 15, 19,$ and 23 were studied using spin-polarized calculations. At all sizes except $N = 15$, the structure with the lowest energy was found to be icosahedral. For $N = 15$, a bcc configuration was found to be favorable. In terms of spin, all the structures were found to be ferrimagnetic with alternating domains of different spin configurations (see Fig. 4). Except for $N = 15$, the mean value of the integrated spin density was found to decrease with increasing size.

In the DFT work by Kabir et al. (2006), the ground state structure for Mn_{13} was found to be the icosahedron with the two pentagonal rings that are coupled antiferromagnetically. Therefore, the resulting magnetization is small, namely, $0.23\mu_B/\text{atom}$. This magnetization is considerably smaller than the neighboring sizes 12 and 14. The $N = 14$ structure differs from the $N = 13$ structure by a single capping atom. However, the presence of this atom changes the magnetization arrangement considerably. In this case the pentagonal rings are ferromagnetically coupled and the magnetization is $1.29\mu_B/\text{atom}$. The case of Mn_{15} is worth mentioning because of the discrepancy between two DFT studies by Briere and coworkers (2002) and Kabir et al. (2006). In the latter, the ground state structure was found to be icosahedral with a magnetic moment of $0.87\mu_B/\text{atom}$, whereas Briere et al. (2002) found a bcc structure with $0.20\mu_B/\text{atom}$. For $N = 19$, a double icosahedron was observed, which again has a smaller magnetic moment, $1.10\mu_B/\text{atom}$, than its neighboring clusters. The central pentagonal ring is AFM coupled to the neighboring pentagonal rings. This behavior is persistent in the $N = 20$ cluster, which has a magnetic moment of $1.50\mu_B/\text{atom}$. In the range $N = 11 - 20$, spin segregation is observed, where like spins tend to cluster. The binding energy is observed to increase monotonically with

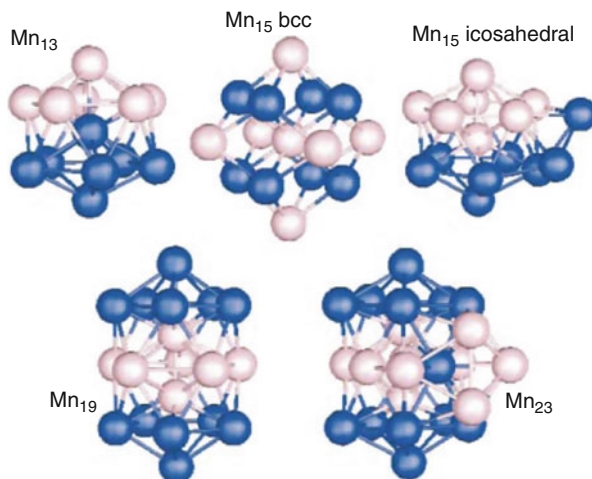


Fig. 4 Lowest energy structures of some Mn clusters. Relative spin alignments are marked with dark for spin up and light for spin down (Reprinted with permission from Briere et al. (2002). © (2002) by the American Physical Society)

increasing size. This is due to the increased sp. bonding. However, when compared with other transition metals, Mn clusters remain weakly bound.

An interesting property of Ta clusters was demonstrated in a recent study (Fa et al. 2006), where ferroelectricity and ferromagnetism were proven to coexist. Initial structures were obtained by simulated annealing using an empirical potential. These structures were later reoptimized with DFT calculations. The magic numbers for Ta clusters were found to be 4, 7, 10, 15, and 22. It was therefore deduced that Ta clusters do not prefer icosahedral growth. For $N = 13$, for instance, the lowest energy structure among those studied was found to be a distorted five-capped hexagonal bipyramid. For $N = 19$, the most stable structure is decahedral in contrast with the double icosahedron, which was found to be stable for many other clusters. No perfectly symmetric structures were found indicating that Jahn–Teller distortions play an important role in determining the ground state structures of Ta clusters. In the size range studied in this work (Fa et al. 2006), the atomic packing shows differences such that each size behaves like an individual system rather than steps of a continuous growth sequence. In addition, electronic dipole moment and magnetic moment were also calculated. The electronic dipole moment was found to have the same trend as the *inverse coordination number*, which is a parameter that reflects the asymmetry of the cluster. This agreement is attributed to the strong correlation between the structure and the electronic dipole moment of the clusters. Odd- N Ta clusters also display a magnetic moment of about $1\mu_B$, which suggests the possibility of the coexistence of ferroelectricity and ferromagnetism. The growth pattern of Ta was found to be very similar to that of Nb. However, when compared with vanadium clusters, this similarity is absent (Fa et al. 2006).

In a similar work by Fa and coworkers (2005), Nb clusters were also found to display ferroelectricity supported by a recent experimental study. This is an important discovery because ferroelectricity was never observed in single-element bulk materials. For $N \geq 38$, the electric dipole moment exhibits even-odd oscillations. This suggests that there is a strong correlation between the structure and the ferroelectricity.

In the work by Pawluk and coworkers (2005), the structure and stability of several Ir clusters were studied using DFT. Rather than using a global optimization algorithm, possible configurations both truncated from bulk and built independently were relaxed locally. The results indicate that Ir clusters mostly prefer cube-like structures up to $N = 13$, except for $N = 11$, which assumes an elongated structure. At $N = 13$, the lowest structure among the ones studied is the icosahedron. This is in contrast with Ru and Pt, which prefer simple cubic structures. When compared to clusters cut from the fcc bulk, simple cubic structures turn out to be more stable up to a size of $N = 48$. This transition occurs near $N = 40$ in Ru and $N = 13$ for Pt. Another interesting property studied in this work is the fluidity of clusters. The results indicate that while the Pt clusters exhibit a more fluidlike character and will thus easily coalesce with other clusters, Ir clusters are more rigid and have less tendency toward coalescence.

An interesting experimental result concerning Pt clusters was reported by Liu and coworkers (2006), who found that Pt_{13} clusters exhibit substantial magnetism (about $0.65\mu_B/\text{atom}$) even though bulk Pt is not magnetic.

Although for such clusters as Ni and Au, the icosahedral structure is calculated to be metastable with stability decreasing with size, experiment reveals that icosahedral structures are found in clusters containing several thousands of atoms. This means that kinetic effects are also very important in determining the structure of a cluster (Gafner et al. 2004). In order to simulate these kinetic effects, a 555 Ni cluster was studied using tight binding by Gafner and coworkers in (2004). The cluster was heated to 1,800 K ($T_{\text{melt}} = 1,145$ K) and subsequently cooled to 300 K. The melting and crystallization curves are determined from a sudden change in the potential energy as a function of temperature. They found that slow cooling results in an fcc structure, whereas fast cooling results in the formation of a metastable icosahedral structure.

In the study by Köhler and coworkers (2006), the potential surface of a few sizes ($N = 13, 53\text{--}57$) of Fe clusters was mapped out with respect to magnetization and volume change using a DFT-based tight-binding scheme. Icosahedra were found to be the most stable structures for the magic numbers $N = 13$ and $N = 55$. Two local minima were observed for the $N = 13$ icosahedron, one ferromagnetic and one antiferromagnetic. The PESs for clusters with $N = 53\text{--}57$ were mapped out using a genetic algorithm-based procedure. Derivatives of the $N = 55$ structure were considered for $N = 53, 54, 56$, and 57 . Generally icosahedra and icosahedron-derived structures have relatively small magnetic moments. The structures without apparent symmetry show higher magnetic moment than icosahedra. No ferromagnetic ordering was found for the Fe_{55} cluster.

Tiago and coworkers provide a very comprehensive explanation of the origin of magnetism in small Fe clusters in their article (Tiago et al. 2006). In the Fe atom, the magnetic moment is a result of exchange splitting. The $3d_{\uparrow}$ states are occupied by five electrons, while the $3d_{\downarrow}$ states are occupied by a single electron, which results in a rather high magnetic moment of $4\mu_B$. When the atoms come together to form a crystal, hybridization of the large $4s$ bands and the $3d$ bands reduces the magnetism down to $2.2\mu_B$. In clusters, hybridization is not so strong because of the reduced coordination numbers of the surface atoms. Because this hybridization depends on orientation, clusters with faceted surfaces are expected to have different magnetic properties than those with irregular faces. According to Tiago and coworkers (2006), this effect is the likely cause of the nonmonotonic suppression of magnetic moment as a function of size. Two classes of Fe clusters were considered in this work: faceted and nonfaceted. Nonfaceted structures are nearly spherical in shape and faceted structures are built using the conventional layer-by-layer growth model. The magnetic moment was calculated as the expectation value of the total angular momentum,

$$M = \frac{\mu_B}{\hbar} [g_s \langle S_z \rangle + \langle L_z \rangle], \quad (4)$$

where $g_s = 2$ is the gyromagnetic ratio of the electron. The magnetic moments of the clusters as a function of size are displayed in Fig. 5 for all classes of clusters considered. Suppression of the magnetic moment with increasing size is observed, in good agreement with experiment. Clusters with faceted surfaces indeed have a lower magnetic moment due to more efficient hybridization although the correlation of shape and magnetic moment is not always well defined. Icosahedral structures are predicted to have magnetic moments lower than bcc clusters.

Another good example of disagreement between different methods is given in Krüger and coworkers where Pd clusters having several high-symmetry structures including icosahedra, octahedra, and cuboctahedra were optimized using both the LDA and GGA in the DFT framework. GGA yields larger bond lengths in accord with the general expectation. But the difference is almost independent of size. Icosahedral structures tend to yield larger bond lengths than octahedral and cuboctahedral structures, also displaying a flatter variation with increasing coordination number. The accuracy of this *breathing mode* relaxation was confirmed by comparing to a full relaxation of the Pd₅₅ cluster.

The Pd₁₃ cluster was found in the work of Kumar and Kawazoe (2002) to have an icosahedral structure with a $0.62\mu_B/\text{atom}$ magnetic moment. This result agrees with some of the earlier studies and disagrees with some others. The central atom in the cluster is found to have a smaller magnetic moment than the surface atom only by a very small amount unlike the large difference in the Mn₁₃ study of Kabir and coworkers (2006). For $N = 14$ and 15, icosahedron-derived structures are favored over other high-symmetry structures. The magnetic moments for both sizes remain around $0.5\mu_B/\text{atom}$. Going from $N = 16$ to the double icosahedron at

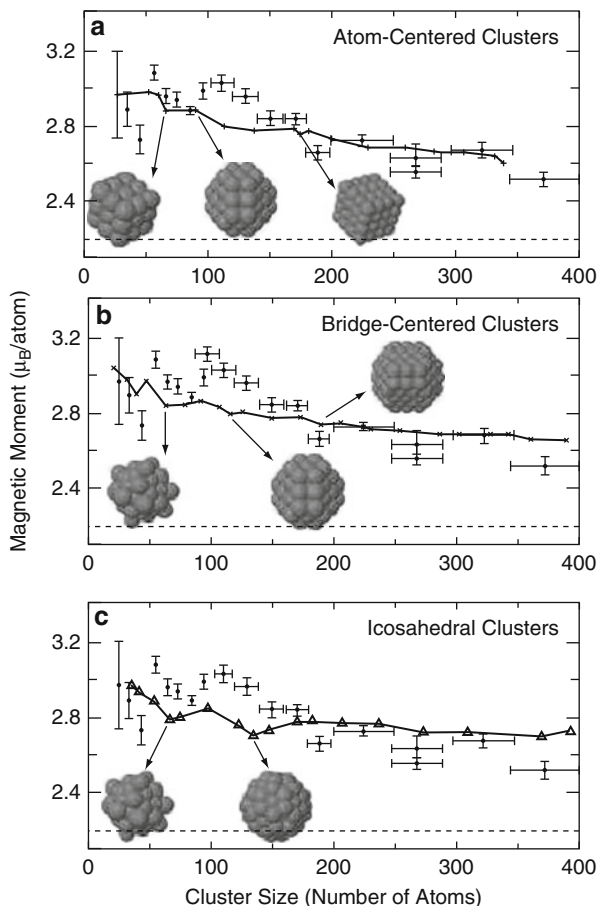


Fig. 5 Magnetic moments of faceted and nonfaceted Fe clusters in the icosahedral and bcc configurations: (a) atom-centered bcc, (b) bridge-centered bcc, and (c) icosahedral. Experimental data is displayed as *black diamonds* for comparison (Reprinted with permission from Tiago et al. (2006). © (2002) by the American Physical Society)

$N = 19$, the magnetic moment decreases from $0.38\mu_B/\text{atom}$ to $0.32\mu_B/\text{atom}$. As the cluster sizes increase, there is a decrease in local (atomic) magnetic moments. For $N = 55$ and $N = 147$, the lowest lying state was found to be the Mackay icosahedra. Cubic and decahedral structures were found to have the next highest energies, with $0.29\mu_B/\text{atom}$ and $0.41\mu_B/\text{atom}$, respectively. For the Pd_{55} cluster, structures with smaller magnetization lie very close to the ground state and would therefore be accessible at room temperature. So magnetic order can be easily lost under experimental conditions. In contrast to the results for Mn (Kabir et al. 2006), the icosahedral structures have a higher magnetic moment than other high-symmetry structures.

In work of Nava and coworkers (2003), Pd clusters were studied using the spin-polarized DFT method in the range $N = 2$ –309. The $N = 13$ cluster was found to have an icosahedral structure with a high spin state. It is seen to undergo a very slight Jahn–Teller distortion, which increases the cohesive energy only by 0.01 eV. The truncated decahedron and the cuboctahedron are found to be less stable.

Yao and coworkers (2007) utilized results from simulated annealing of Ni clusters using an empirical potential as starting configurations to further optimize them using a DFT code. This is another technique often used in cluster literature. In this work, Ni clusters with $N = 10$ –27 atoms were found to attain icosahedron-like structures with the $N = 13$ cluster being a perfect icosahedron and $N = 19$ the double icosahedron. Clusters in the range $N = 28$ –40 have very complicated structures because they are in the transition region between the $N = 13$ icosahedron and the $N = 55$ icosahedron. Around $N = 19$ and $N = 55$, clusters are mostly formed by adding or removing a few atoms from the corresponding perfect icosahedra. Dips are observed in the magnetic moment for $N = 13$ and 55 as expected from the compact structure of these clusters.

While all of the studies mentioned above deal with collinear spin, which singles out one direction along which the magnetic moment may be oriented, a very important class of clusters with high magnetic moments displays noncollinear spin. In these clusters, the magnetic moment is allowed point in an arbitrary direction and thus presents a new degree of freedom. A good example for the investigation of this effect is presented in the work by Du et al. (2010) where six-atom clusters of Co and Mn atoms with several different compositions but always in the octahedral geometry were investigated within the DFT theory including the noncollinear spin formalism. Low-lying isomers with up to more than 40° of average degree of noncollinearity were identified. For certain compositions and geometries, different tendencies for magnetic coupling (AFM vs. FM) are also found to cause a certain degree of spin frustration.

In the work by Zhang and coworkers (2004), Mo clusters were studied within the DFT framework. An interesting result was that the initial icosahedral structure for the $N = 13$ cluster was found to undergo a very large distortion. This distortion was explained by the tendency of Mo clusters to form Mo dimers. The strength of the bonds in Mo_{13} is covalent; therefore, at this size, the cluster shows nonmetallicity. At $N = 55$, the icosahedral structure is found to again undergo a large distortion. However in contrast to $N = 13$, the alternate O_h structure was not found to be much more stable than the distorted icosahedron. This was explained by the decrease of nonmetallicity with size.

1D Structures: Nanotubes, Nanowires, and Nanorods

In nanoscience literature, the name *one-dimensional* was coined to describe systems where one of the dimensions is several orders of magnitude larger than the other two dimensions. Much like the zero-dimensional case, the border between one- and two-dimensional systems is ill defined and system dependent. For this reason,

two-dimensional systems exhibit the same richness in structural and electronic properties as in the zero-dimensional case unlike their bulk counterparts.

One-dimensional structures may, in the broadest classification, be divided into nanotubes, nanowires, and nanoribbons. Graphene nanoribbons (GNRs) (Law et al. 2004; Hyun et al. 2013). *Nanotubes* are simply described as two-dimensional materials (such as graphene, BN sheets, TiO₂ sheets, and many others) seamlessly rolled into a hollow cylinder although the actual fabrication usually follows a much more involved procedure. Nanotubes may be regarded as a unique subset of one-dimensional structures that do not have a surface and are thus devoid of surface effects. In this respect they are analogous to the fullerenes in the zero-dimensional case. *Nanowires* on the other hand are extremely thin wires that are grown or extracted along well-defined crystal directions and may have widely different surfaces. The surfaces, however, rarely remain in their bulk configuration and often reconstruct to reduce strain or saturate broken bonds. This procedure is highly size and material specific, and nanowires therefore display a large variety of strongly surface-dependent properties (Dasgupta et al. 2014; Kempa et al. 2013). Finally, *nanoribbons* are thin strips of two-dimensional materials such as graphene. The edges may reconstruct or be saturated with different species to modify their properties (Ozdamar and Erkoc 2012; Ince and Erkoc 2011; Kilic and Erkoc 2013b).

In addition to structural and electronic properties that are explored in zero-dimensional materials, one-dimensional materials also exhibit rather interesting elastic properties. We shall begin this section with a brief review of elastic considerations regarding one-dimensional nanomaterials and afterward move onto structural and electronic properties.

Elastic and Structural Properties

In nanowire applications such as AFM tips, NEMS, and MEMS, which make use of mechanical properties, it is crucial to have a good understanding of the evolution of elastic properties all the way down to the nanoscale. Elastic properties that are ordinarily under investigation include elastic moduli, plasticity, crack propagation, buckling, and breaking points.

There are numerous examples of experimental determination of elastic properties of nanowires in the literature. A recent study by Barth and coworkers (2009) determines the Young's modulus of SnO₂ nanowires by anchoring and bending them with the help of an AFM tip. By mapping the bending amplitude to the Young's modulus through classical elasticity formulae, the Young's modulus is estimated at around 110 GPa for the samples studied.

Contrary to the macroscopic scale, the elastic properties of nanoscale one-dimensional systems are often seen to depend on their physical dimensions. This intriguing fact brings forth the necessity of studying, among others, the elastic properties of such systems as a function of their size. In fact, the theoretical literature is very rich in examples of such studies. Empirical potentials are particularly suitable for studies of size dependence since systems with large numbers of atoms may be

handled with relative ease, allowing the determination of convergence of elastic properties of nanomaterials to those of their bulk counterparts. A related question to this fact is to what extent the laws of continuum elasticity can be applied to nanoscale systems, which has received much attention both theoretically and experimentally. A recent study by Rudd and Lee (2008) opens with this very same question, where the Young's modulus of Ta (001) and Si (001) nanowires has been determined using a Finnis and Sinclair (1984) and Stillinger and Weber (1985) potential, respectively, in a molecular dynamics simulation. Nanowire radii were explored up to approximately 10 nm and both materials were found to display a strong dependence on the wire width eventually converging to the bulk value. Interestingly, the convergence trends are opposite where the Young's modulus of the Si nanowires increases while that of Ta nanowires decreases with radius. In the same study, DFT calculations were also performed on Si nanowires, partially confirming the outcome of the Stillinger–Weber calculations. However, the match with continuum theory was found to be much better for the DFT results.

Size dependence is also demonstrated in a study by Hu and coworkers (2008) on ZnO nanowires and nanotubes where they use an exp-6-type empirical potential including Coulomb interaction for Young's modulus calculations. The Young's moduli of nanowires and nanotubes show a very strong dependence on the radius and wall thickness, respectively.

Due to the extremely high surface-to-volume ratio in nanowires, the particular surfaces that are exposed at the outer edges of the nanowires play a crucial effect in the determination of their elastic properties (Ozdamar and Erkoç 2013; Kilic and Erkoç 2013a; Yagli and Erkoç 2014). Recent evidence for this fact was demonstrated by Wang and Li (2008) in their DFT-parameterized model study of Ag, Au, and ZnO nanowires with different surface terminations. Their results show discernible albeit small differences in the size dependence of the Young's modulus for different surface terminations of the nanowires in question.

Nanotubes have also received a great deal of attention from researchers due to their extraordinary elastic properties. They have been shown to possess an unusually high axial stiffness in addition to very high reversibility under large distortions. In addition to the large body of literature on experiments probing the elastic properties of nanotubes, many theoretical studies have also been conducted. Due to their varying radii and chirality, nanotubes present endless possibilities for the investigation of their elastic properties. Liang and Upmanyu (2006), for instance, have studied the radius (or equivalently curvature) and chirality dependence of the torsion induced by applied axial strength of the nanotubes. Their studies, which utilize the widely used second-generation reactive empirical bond-order potential (Brenner 1989), revealed a torsional response of up to 0.75 nm, which varied remarkably for different radii and chiral angles. Conversely, as reported in several early studies on nanotubes employing DFT calculations, Young's modulus is known to be largely independent of the chirality.

Elastic properties of nanotubes of many materials other than C have also been theoretically explored. In an exhaustive work, Baumeier and coworkers (2007) used DFT calculations to survey such properties as strain energy and Young's

modulus of SiC, BN, and BeO nanotubes. Baumeier and coworkers observe that being composed of two atomic species instead of one as in carbon nanotubes, the nanotubes made out of these materials exhibit a different relaxation pattern for the relevant anions and cations, suggesting possible different behavior than carbon nanotubes. Nevertheless, the results show that at least for (n, n) and $(n, 0)$ nanotubes of similar radii, the behavior is similar. All three materials display a decreasing Young's modulus as a function of radius converging to the sheet value for large radii. Among the three materials, BN nanotubes display a significantly high Young's modulus, followed by SiC and finally by BeO nanotubes.

The interwall attraction in multi-walled nanotubes opens up another possible avenue for the study of elastic properties. Zhang and coworkers (2008b) have studied the Young's modulus, Poisson ratio, and the buckling point of multi-walled nanotubes using a combination of second-generation reactive bond-order potential to model intralayer bonding and a Lennard-Jones potential for the interlayer interaction. The multi-walled nanotubes studied were divided into two sets, the first formed by embedding increasingly smaller (n, n) tubes into a large $(20, 20)$ nanotube (up to four walls) and the second by placing a $(5, 5)$ tube into increasingly larger (n, n) tubes (again up to four walls). A molecular dynamics algorithm was used during their calculation and the dependence of the moduli in addition to the buckled morphologies was presented. Young's modulus and Poisson's ratio turn out to follow a different trend for the two sets considered. While the set that grows inward displays increasing (decreasing) Young's modulus (Poisson ratio) for increasing number of walls, the set that grows outward follows the opposite trend.

As mentioned in the introduction to this section, one way to saturate the surface bonds of nanowires is through passivation by different species. The elastic behavior of passivated versus unpassivated nanowires raises an interesting question. Lee and Rudd (2007) studied several elastic properties of H-passivated nanowires of a large range of radii varying between 0.61 and 3.92 nm (see Fig. 6) using DFT. As expected, the Young moduli of H-passivated Si nanowires mostly follow predictions from continuum models regardless of the varying proportions of $\langle 100 \rangle$ and $\langle 001 \rangle$ surfaces exposed. This is attributed to the fact that the bulk-like covalent bonding character at the surfaces is preserved when H is used to passivate the dangling bonds at the surface.

Although microscopic modeling is of utmost importance in the understanding of nanoscale materials, there are a number of experimental situations of interest that cannot be handled by these time-intensive methods due to their large size. For large enough systems (such as large portions of nanotubes suspended over a trench), continuum methods may be employed (Ustunel et al. 2005). However, complete coarse graining is also not always a viable choice since one then loses detailed information on locally nonhomogeneous regions of the system such as defects and local deformations. In such cases, *multiscale methods* which apply different methods at different scales of the system are the methods of choice. An illustrative example was studied by Maiti (2008) where a micromechanical sensor made out of a nanotube was deformed by a Li needle. The bent but undeformed portions of the nanotube were modeled by a coarse-grained *molecular mechanics* simulation, while

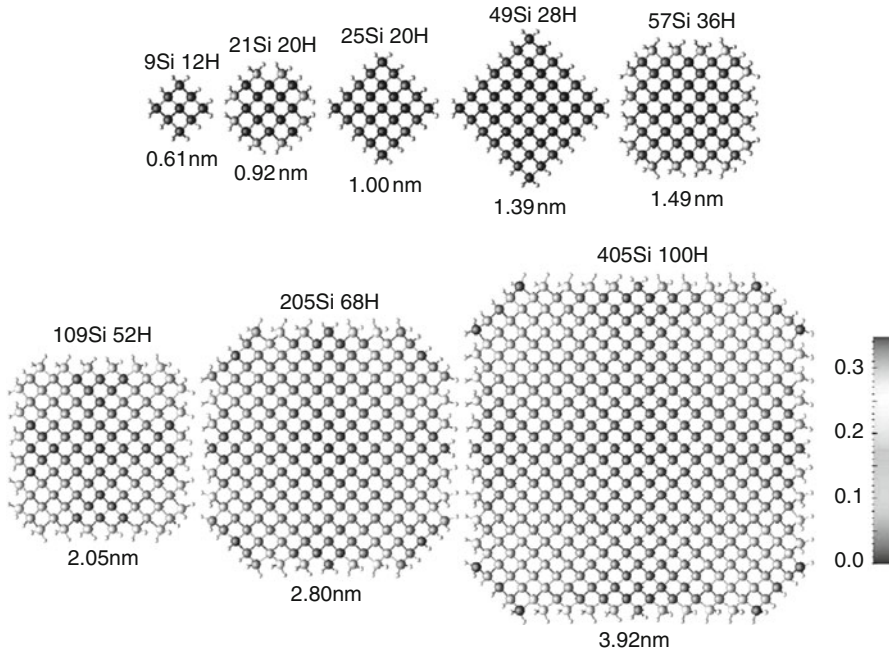


Fig. 6 H-passivated Si nanowires of different sizes studied by Lee et al. (Reprinted with permission from Lee and Rudd (2007). © (2007) by the American Physical Society)

the highly deformed midsection (enclosed in a box) was modeled using a quantum mechanical method. The two methods were then matched at the interface of the two regions.

Structural Properties

The novel electronic properties of nanotubes, nanowires, and nanobelts are inextricably linked to their structural properties. This is largely due to their high surface-to-volume ratio, where the bonding on the surface structure determines the electronic states which in turn determines such properties of the system as conductivity and magnetism.

As mentioned in the introductory section, nanowires nowadays may be manufactured from a great variety of materials. As fabricated, it is experimentally difficult, if not impossible, to intuitively infer their surface structure. Simulations on the other hand provide an inexpensive yet accurate way of studying this relation between structure and electronic properties.

The methods of preference in nanowire modeling are generally tight binding and to a larger extent DFT, since electronic properties depend sensitively on the structure requiring accurate calculations. Once a material (or combination of materials) is

chosen, the important parameters of nanowire modeling are the particular surfaces surrounding the nanowire and passivation, that is, the saturation of the bonds at the surfaces. Two of the most commonly studied surface terminations are surface reconstruction and H passivation.

Due to the axial periodicity of one-dimensional nanomaterials, plane wave-based methods which are traditionally used for crystals are very often employed. In the work of Arantes and Fazio (2007), where they study free and passivated Ge nanowires, the band gap of passivated and unpassivated nanowires was determined using a plane wave-based GGA-DFT method. The nanowires were grown in the $\langle 110 \rangle$ and $\langle 111 \rangle$ directions, and their band gaps were calculated as a function of nanowire diameter. In spite of the well-known underestimation of the band gap by LDA and GGA methods, a trend can be obtained rather reliably. The band gaps are seen to vary with respect to direction and size.

Though not often encountered, global search algorithms are also used for determining the structure of nanowires. Especially in the case of nanowires with smaller radii, the structure may be so different from bulk as to prevent any a priori prediction. Chan and coworkers (2006) conducted a genetic algorithm search based on the formation energy of H-terminated Si NWs where the formation energy is defined as

$$f = (E - \mu_H n_H) / n - \mu, \quad (5)$$

where E is the total energy of the NW in question, μ_H and μ are the chemical potentials of H and Si, respectively, and n is the number of Si atoms. The genetic algorithm was conducted in two stages. In the first stage, a long evolution through several generations was conducted using an empirical potential (of the Hansel-Vogel type). In the second stage, the outcome structures of the evolution for each size were relaxed using a DFT algorithm. The result, reminiscent of cluster structures, is that certain sizes of Si NWs are preferred over other sizes. These structures are once again termed *magic sizes*. In general, structures with even number of Si atoms are preferred over those with an odd number. The most stable structures are observed either in a *platelike* form where chains of Si hexagons join together to form flat structures, or a *hexagon-shaped* cross section.

One of the most important parameters in determining the geometric and therefore the electronic structure of nanowires is surface termination. The high-energy dangling bonds at the surface may either be saturated by a rearrangement and rebonding of the surface atoms or by attaching electropositive species (such as H). In general, while the reconstruction may significantly alter the structural and electronic properties of the material, passivation by other species leaves these properties relatively unaltered. In the recent work of Migas and Borisenko (2008), $\langle 001 \rangle$ -oriented Si nanowires were passivated by O, F, and H, and the structural and a large number of different sizes and geometries were studied using DFT. Under different combinations and coverages by these elements, the band gap of the Si nanowires considered (only nanowires with rhombic cross sections were considered) was

found to vary between 1.20 and 1.96 eV, which shows that termination can be used as a means to control the electronic properties.

Countless other studies in understanding the structure of experimentally relevant one-dimensional structures have been done. Bi nanowires (Qi et al. 2008), CdSe nanorods with hexagonal and triangular cross sections (Sadowski and Ramprasad 2007), and Te nanowires (Ghosh et al. 2007) are some studies that may be mentioned. In addition to materials that have already been manufactured, researchers have also been interested in the possible existence and properties of nanoscale systems that have not yet been experimentally realized. Rathi and Ray (2008), for instance, have examined the possibility of Si–Ge nanotubes, while Qi and coworkers (2008) investigated Bi nanotubes and hollow Bi rods. Recently it was demonstrated by model calculations that the stable carbon nanotube structure might be possible with non-graphene-like form (Erkoc 2004a). Furthermore, nanorod structures constructed from benzene rings only (called as *benzorods*) may also be possible. Their structural and electronic properties were investigated by performing model calculations (Erkoc 2003; Malcioglu and Erkoc 2004).

Electronic, Magnetic, and Optical Properties

As a result of the immense variation in structural properties of one-dimensional structures, one observes an equally diverse spectra of electronic properties. Perhaps the most intriguing property of one- and two-dimensional nanoscale systems is that after the required geometrical deformations in order to reach their equilibrium, one might observe a stark difference between the newly formed system and its bulk counterpart. A material which is an insulator in bulk may become a conductor when taken to the nanoscale. Similarly one- or two-dimensional nanostructures of a nonmagnetic material may have a nonzero magnetic moment. These differences usually stem from the new states introduced into the electronic structure of the material by terminating structures such as surfaces, steps, edges, and corners. The interest in these emergent properties is due to the possibilities of integrating these small-scale materials into technological applications and controlling their properties by controlling the structure.

Due to the ease of interfacing with current technology, *Si* nanowires have perhaps been the most intensely studied system. Rurali and Lorente (2005), for instance, explored a large range of surface reconstructions for $\langle 100 \rangle$ *Si* nanowires with a small radius of about 1.5 nm. They discovered that for certain reconstructions, the nanowires develop conducting states in their band gap, while for others the semiconducting behavior is retained. This is a prime example of effects of confinement on the electronic structure of a system.

Another parameter that has a strong effect on the electronic properties of nanowires is surface termination. As demonstrated by Rurali (2005) that both *Si* and C-terminated H-passivated SiC nanowires have a larger band gap than that of bulk SiC. As discussed by Rurali, this is purely a confinement effect since passivation prevents reconstruction and the related introduction of gap states. However, when

the nanowires are allowed to reconstruct, both species of SiC nanowires are seen to become conducting due to the surface states introduced into the band gap.

In addition to the conventional carbon nanotubes, several other materials were investigated as viable candidates for nanotube structures. BN (Chopra et al. 1995) and more recently (Sun et al. 2002) SiC nanotubes are two such materials which have been successfully synthesized in the nanotube form. Following their synthesis Gao and Kang (2008) conducted a DFT-PAW study of undoped and N-doped SiC nanotubes of varying sizes. For each size the nanotubes were doped initially with 2 and 4 N atoms and their structural and electronic properties were studied. For each doping level, several possibilities were investigated and the most stable structure was identified. As an extreme case, the case in which all the C atoms were replaced with N atoms, in other words SiN nanotubes, was considered. These nanotubes, instead of being circular, were found to have a staggered or starlike cross section. In all the cases considered, the nanotubes were found to be semiconducting with an indirect band gap. Recent model calculations on binary compounds BN (Erkoc 2001b), GaN (Erkoc et al. 2004), InP (Erkoc 2004b), and ZnO (Erkoc and Kokten 2005) nanotubes give reasonable results comparable with experimental findings.

Although by nature rather free of defects, the few existing defects in graphene and nanotubes change the electronic structure of their host material drastically (Pekoz and Erkoc 2008). In spite of the several experimental methods that have been developed to locate and study the properties of such defects, theoretical methods are an indispensable tool for creating isolated defects of the desired nature and studying their effects on electronic properties. A particularly interesting question is the stability of well-characterized graphene defects in nanotubes of varying radius and chirality. Since a graphene sheet can be viewed as a nanotube with an infinitely large radius, the formation energy of any nanotube defect should tend to the equivalent defect on a graphene sheet. Amorim et al. (2007) demonstrated an example of this behavior by studying the so-called 555777 defect, which is a combination of three pentagons and three heptagons. This defect is formed in two steps. In the first step, two divacancies coalesce to form a new defect (585) composed of two pentagons and an octagon. This step is followed by a further structural change, which yields the 555777 defect, see Fig. 7.

In graphene the 555777 defect is found to be more stable than the 585 defect by 0.8 eV. The same defects were then created in zigzag and armchair nanotubes of radius in the range of 7–15 Å. As in the graphene case, the 555777 defect is found to be more stable in all the nanotubes studied. The formation energy is, however, lower than in the graphene case. The expected tendency toward the corresponding graphene values is seen in both cases and convergence is estimated to occur around a radius of 40 Å.

The effect of defects on the electronic properties of their host substance is illustrated in this work by the conductance graph calculated using the Green's function density functional theory. The results indicate that the presence of both the 585 and 555777 reduces the conductance considerably while at the same time displaying different voltage dependence.

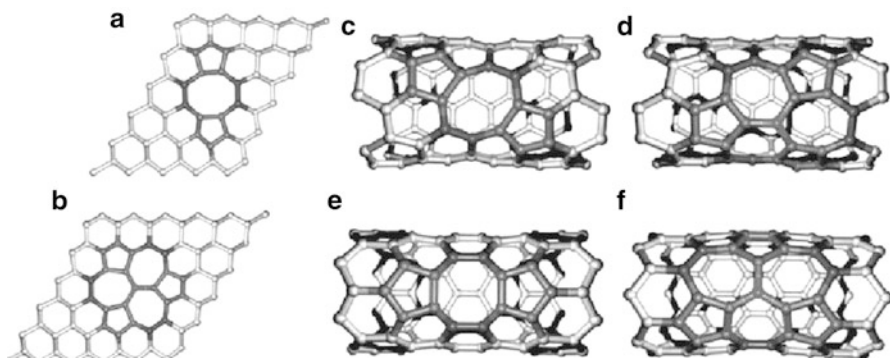


Fig. 7 585 (a) and 555777 (b) defects in graphene and 585 (c, e) and 555777 (d, f) defects in a nanotube in different orientations (Reprinted with permission from Amorim et al. (2007). © (2007) by the American Chemical Society)

Another investigation of defects in nanotubes for the purpose of application as a gas sensor was conducted by Andzelm and coworkers (2006). The particular question at hand is the binding of NH_3 molecule to nanotubes and whether or not binding is enhanced by defects. Three defects are considered: a Stone–Wales defect (a defect formed by rotating a given bond by 90° resulting in the formation of two pentagons and two heptagons), a monovacancy, and an interstitial C atom placed on top of a bridge. In addition, the case of an O_2 molecule dissociated on a SW defect was also considered to mimic the environmental effects. Two different orientations, straight and chiral, for the SW defect were considered. All-electron DFT calculations reveal that the monovacancy is the most stable. Calculation of reaction barriers also reveals that defects with preabsorbed O dissociate NH_3 into NH_2 and H.

The heterostructure problem, which has been widely studied in the bulk form, is becoming an increasingly popular topic also in the one-dimensional systems. The band alignment problem has been addressed recently in a DFT study conducted by Kagimura and coworkers (2007) where a Si–Ge interface was studied. The model for such a system is shown in Fig. 8. The band states contributed by the surface dangling bonds were investigated as possible candidates for the induction of a potential well.

The search for nanoscale materials that exhibit spontaneous magnetization has become an increasingly rich field in the past decade. Several materials such as doped nanotubes and nanowires, defective graphene, and nanoribbons of two-dimensional materials can be itemized as candidates considered in these studies. In order to identify the suitability of a material for spontaneous magnetization and electron density in the two spin channels, $n^\uparrow(\vec{r})$ and $n^\downarrow(\vec{r})$ are compared. If there is a significant difference in this distribution, the material is nominated for use in magnetic applications. If in addition the density of states in the two spin channels shows different characteristics at the Fermi level such that in one channel there is

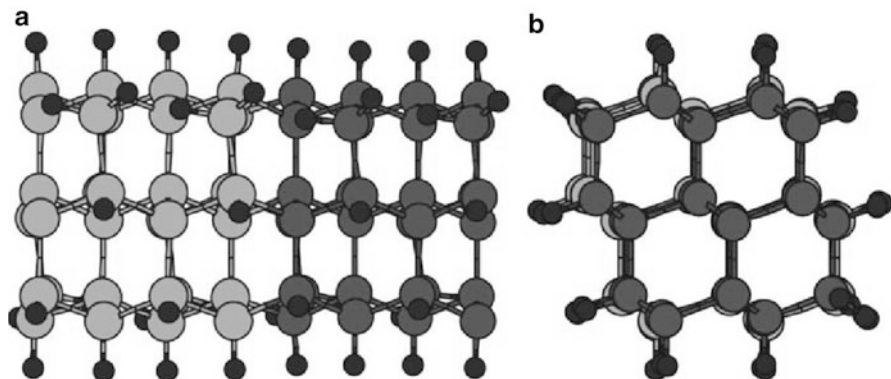


Fig. 8 A Si–Ge interface shaped into a nanotube. (a) Side view, (b) cross view (Adapted from Kagimura et al. 2007)

a significantly larger number of states than the other, then the material provides promise also in *spintronic* applications.

One such candidate material for magnetic applications is BN sheets and nanotubes doped with several different elements. In a recent example, Li and coworkers (2008a), in their DFT study on BN nanotubes with one, two, and three H atoms adsorbed at different locations, observed that some of the configurations considered may give rise to a magnetic moment of up to 2.0. The origin of this magnetic moment is evident from the band structures where the contribution of band gap states is due to only one of the spin channels. The states that are seen in the gap are also shown visually and their origin is unambiguously identified as due to the adsorbed hydrogen.

A rather novel and intriguing application of nanomaterials and its investigation using theoretical modeling is the subject of a recent work by de Oliveira and Miwa (2009). Their study discusses a phenomenon called *self-purification* of nanomaterials, namely, the expulsion of foreign species from the surface. Self-purification is attributed to the fact that nanoscale materials have a lower incorporation rate of impurities compared to their bulk counterpart. In this work SiC nanowires are studied as another potential example of self-purifying materials. B and N impurities were planted at different positions inside and on the surface of three SiC nanowire configurations $\langle 100 \rangle$ Si-coated, $\langle 100 \rangle$ C coated, and $\langle 111 \rangle$ Si and C coated at locations ranging from the center to the surface. The formation energies of the configurations thus formed were then calculated and compared. As a result, the self-purification process was found to be favorable for B-doped SiC nanowires irrespective of their orientation. The B atom was found to segregate favorably to the surface and thus expelled. The N impurities, however, were found to prefer the sites in the core of the nanowires, and therefore, N-doped SiC nanowires were found to fail as self-purifying materials.

One-dimensional systems are actively sought after also for their optical properties. Using the GW method, which is an accurate method for studying excited states, Bruno and coworkers (2007) demonstrated that optical properties of *Si* and *Ge* nanowires depend not only on the nanowire diameter but also on the orientation.

2D Structures: Graphene and Derivatives

Due to the remarkable ability of carbon to exist in different hybridization states, carbon-based materials display an unusually rich variety. Diamond, graphene, cage-like molecules and perhaps most notably carbon nanotubes are examples of the large selection of possibilities.

Graphene, which is a single layer of graphite, may be thought of as the building block of most of the allotropes listed above (Fig. 9). Nanotubes are geometrically rolled up versions of graphene, and fullerenes can be formed by introducing

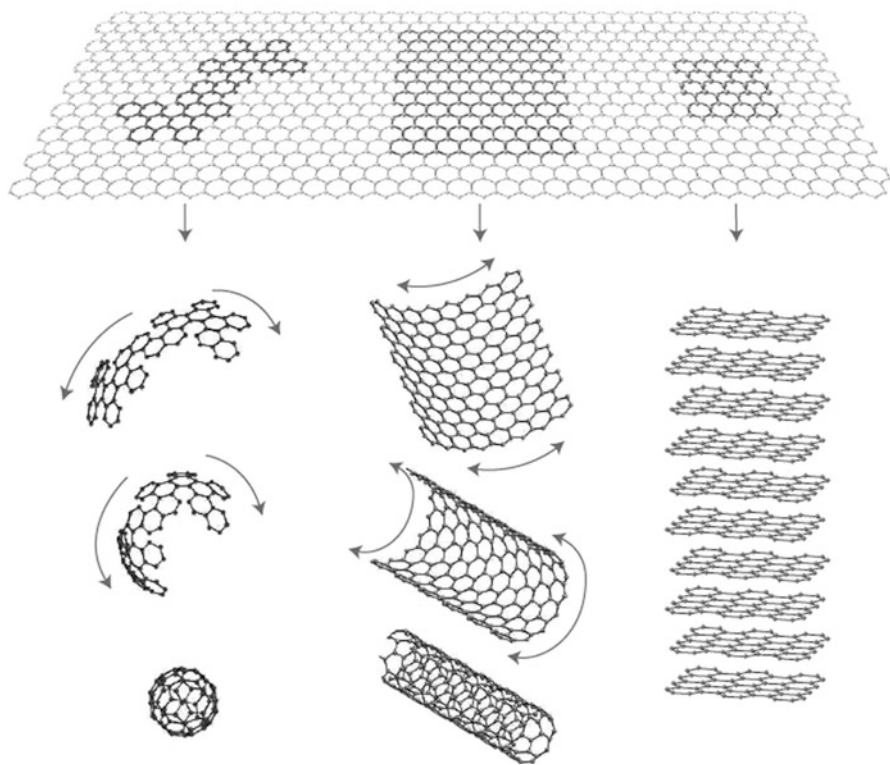


Fig. 9 Graphene is a 2D building material for carbon materials of all other dimensionalities. It can be wrapped up into 0D buckyballs, rolled into 1D nanotubes, or stacked into 3D graphite (Reprinted with permission from Geim and Novoselov (2007). © (2007) by Nature Publishing Group)

topological defects (heptagons and pentagons) into the perfect honeycomb lattice structure of graphene in order to introduce positive and negative curvature (Freitag 2008; Peres et al. 2006b; Kim and Rogers 2008; Tang and Zhou 2013).

Regardless of the immense attention that graphene received in the theoretical literature, it wasn't until 2004 (Novoselov et al. 2004) that a single layer of graphene was isolated experimentally. Since two-dimensional crystals were proven to be unstable theoretically (Landau and Lifshitz 1980; Mermin 1968), the discovery of free-standing graphene came as a surprise (Meyer et al. 2007). This apparent discrepancy, however, was lifted when upon closer inspection, the isolated graphene sheets were not perfectly flat but had corrugations reaching up to 1 nm in size (Rusanov 2014).

Since the discovery and isolation of graphene was achieved, many other materials were found to form two-dimensional structures. In this chapter, we give an overview of carbon-based two-dimensional materials including graphene, graphene nanoribbons, nanobelts, and strips in addition to two-dimensional structures of several other materials (Singh et al. 2011; Bonaccorso et al. 2015; Niu and Li 2015).

Graphene, Nanosheets, Nanoribbons, Nanobelts, and Nanostrips

With the advent of several sophisticated experimental techniques, the two-dimensional confinement of graphene layers was further extended to one dimensional in the form of *nanoribbons*. Nanoribbons are narrow strips of graphene that may exhibit quasi-metallic or semiconducting behavior depending on the geometry of their edges. Much like nanotubes, graphene nanoribbons (GNRs) are also termed *zigzag* or *armchair* based on the directionality of the bonds with respect to the long axis (see Fig. 10). Due to the dependence of their electronic properties on their geometry, it is important to control the morphology and crystallinity of these edges for practical purposes. It has been experimentally shown that (Jia et al. 2009) controlled formation of sharp zigzag and armchair edges in graphitic nanoribbons is possible by Joule heating. During Joule heating and electron beam irradiation, carbon atoms are vaporized, and subsequently sharp edges and step-edge arrays are stabilized, mostly with either zigzag- or armchair-edged configurations.

In addition to the edge geometry, the electronic properties, in particular the band gap, of nanoribbons also depend on their width (Han et al. 2007; Wu and Zeng 2008). The magnetic properties may also be severely altered upon reduction of size to a graphene fragment (Wang et al. 2008) which results in the emergence of giant spin moments.

Another path for controlling the electronic properties of nanoribbons is by an application of an external electric field or by chemical doping of the pristine samples. Half-metallicity, for instance, which has several applications in spintronics (Wu and Zeng 2008) may be introduced through functionalization with such species as H, COOH, OH, NO₂, NH₃, and CH₃ (Son et al. 2006).

Straight GNRs with zigzag, armchair, or mixed edges are proven to be semiconducting by the experiment. In addition, GNRs can be sculpted by attaching two

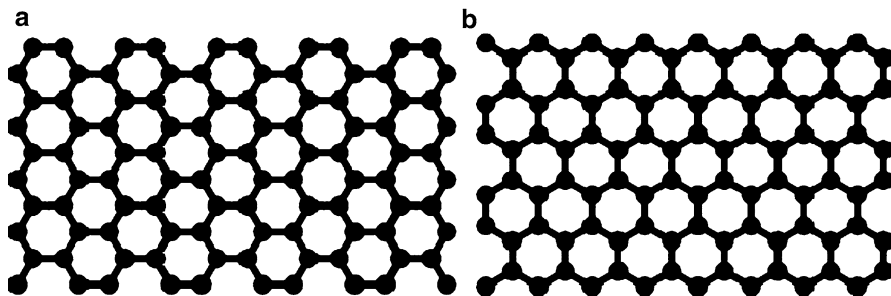


Fig. 10 (a) Armchair graphene ribbon model. (b) Zigzag graphene ribbon model

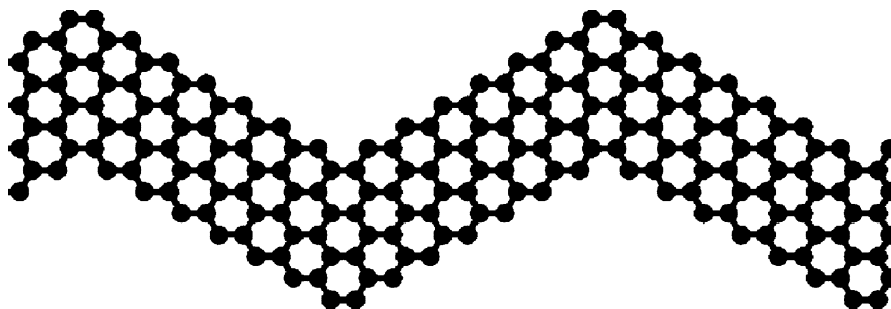


Fig. 11 The model structure of a sawtooth-like GNR (Adapted from Wu and Zeng 2008)

segments together that are manufactured to make a 120° angle with each other thereby forming a *sawtooth-like* nanoribbon (Fig. 11). The structure of a sawtooth-like GNR can be characterized by two integers (w , l). The first integer denotes the width of the nanoribbon, while the second integer describes its periodic length (Wu and Zeng 2008).

Stability is an important issue for a material which is intended to be used as a building block of device applications. Even though perfect two-dimensional crystals are proven to be unstable, graphene is found to be stabilized by corrugations in the third dimension. Understanding the effect of GNR width on the stability is therefore also a central issue for possible applications. Molecular dynamics computer simulations using empirical interatomic potentials predict that structural stability of graphene nanoribbons show dependence on size (width) and edge orientation (Dugan and Erkoç 2008; Ozdamar and Erkoç 2012). Figures 12 and 13 show, respectively, the relaxed structures of armchair- and zigzag-edged GNRs with various widths.

In the recent nanoscale literature, a wealth of materials other than C has been identified, both theoretically and experimentally, as viable candidates for future use as two-dimensional devices.

Boron nitride, for instance, having electronic properties that resemble carbon can exist in a hexagonal structure *h*-BN similar to the graphite-layered geometry.

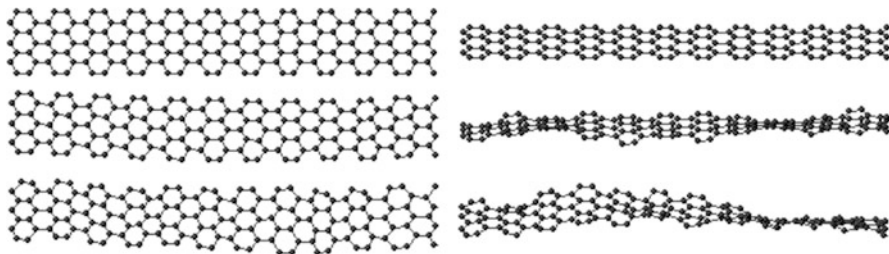


Fig. 12 Armchair-edged GNRs. *Left* column shows *top* view and *right* column shows tilted view. GNRs of width five hexagonal rings are shown. 1 K, 300 K, and final temperature images are given (Adapted from Dugan and Erkoc 2008)

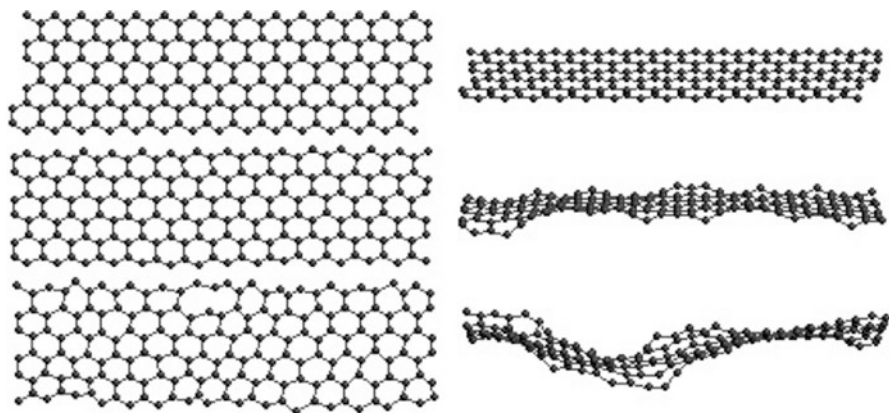
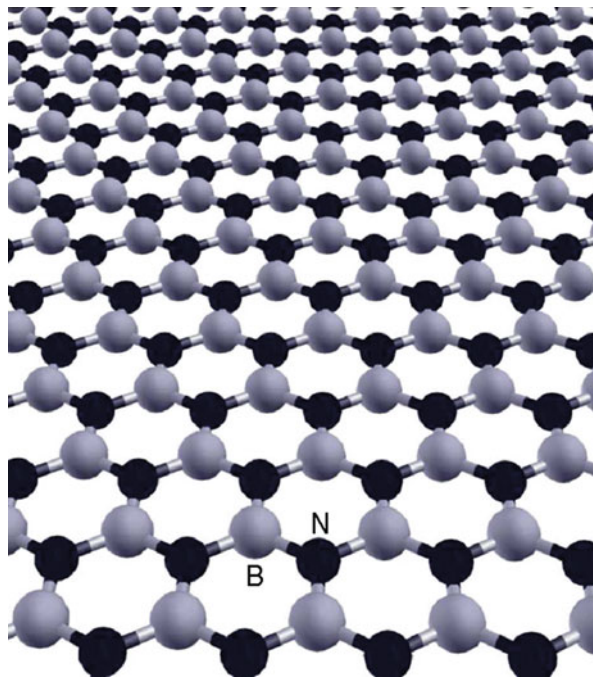


Fig. 13 Zigzag-edged GNRs. *Left* column shows *top* view and *right* column shows tilted view. Ribbons of width five hexagonal rings are shown. 1 K, 300 K, and final temperature images are given (Adapted from Dugan and Erkoc 2008)

Much like graphene sheets, BN sheets can be grown on more or less lattice-matched transition metal surfaces (Corso et al. 2004; Huda and Kleinman 2006). A model BN sheet is shown in Fig. 14.

BN, being a member of the III-V semiconductor family, indicates the possibility of nanoribbons or nanobelts made of other semiconductor families. Indeed, several oxide and other II-VI family nanobelts have been discovered and reported (Pan et al. 2001) (ZnO, SnO₂, CdO, Ga₂O₃, PbO₂, ZnS, CdSe, and ZnSe). Some of the oxides in this family such as ZnO and SnO₂, owing to their polarity and crystal structure, deform in novel morphologies such as rings, springs, and spirals in order to bring together the positive and negative charges, counteracting the charge imbalance (Yang and Wang 2006). These structures comprise a versatile set of nanomaterials that are promising candidates for various applications such as sensors, resonators, and transducers.

Fig. 14 The geometric structure of BN sheet (Adapted from Venkataramanan et al. 2009)



Nanobelts, a term coined by researchers working on these structures (Wang 2004), are described as nanowires with not only a well-defined growth direction but also well-defined top, bottom, and side edges and cross section. The nanobelt structures are usually obtained from functional oxides, which are semiconductor materials, such as ZnO, Ga_2O_3 , $t\text{-SnO}_2$, $o\text{-SnO}_2$, In_2O_3 , CdO, and PbO_2 . Pure metal nanobelt structures are also possible; Zn is one of the metals that form fine nanobelt structures (Wang 2004). ZnO, being one of the most versatile materials in nanoscale research alongside with carbon and BN, has once again been investigated at great depth in the context of nanobelts (Kulkarni et al. 2005; Wang 2004). Molecular dynamics computer simulations using empirical interatomic potentials (Kulkarni et al. 2005) reveal that ZnO nanobelts display properties that depend on their size and orientation. Depending on the growth direction, ZnO nanobelts may exhibit an interesting shell structure or a simple surface reconstruction (Wang 2009; Ozgur et al. 2005).

Nanostrips are similar to nanobelts; they are used with the same meaning with nanobelts. All these nanostructures (ribbons, belts, strips) are ideal materials for building nano-sized devices and sensors (Lin et al. 2009).

In addition to the materials mentioned above, recent theoretical studies have proposed a wealth of two-dimensional structures that are composed of less common materials. One such example is the recently proposed B_2C graphene (Wu et al. 2009). The optimized B_2C graphene structure is displayed in Fig. 15. Two neigh-

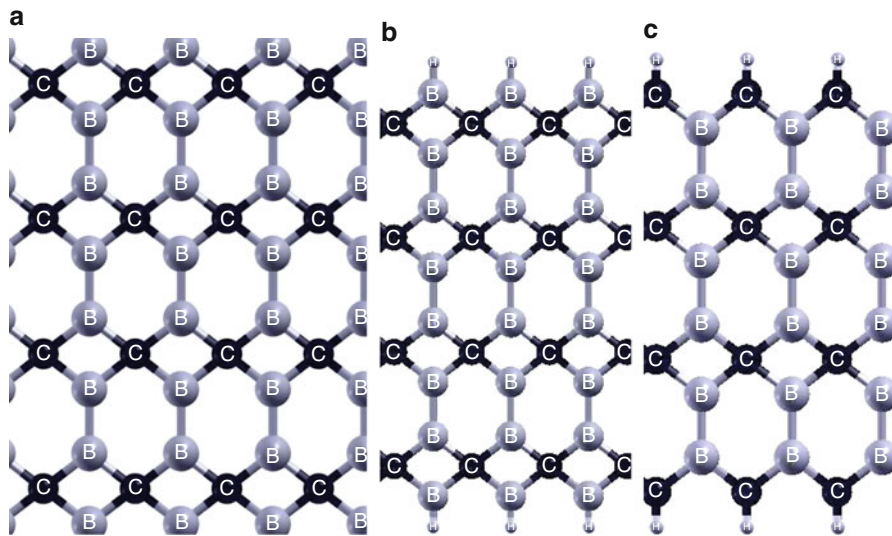


Fig. 15 (a) A $C_{2V} - CB_4$ motif structure. Each carbon atom is bonded with four boron atoms. Optimized structures of two types B_2C nanoribbons with two different edge configurations, referred as type I (b) and type II (c) (Adapted from Wu et al. 2009)

boring CB_4 motifs share two common boron atoms, giving rise to a hexagon and a rhombus. The mean B–C and B–B bond length is 1.557 and 1.685 Å, respectively. The sheet is slightly corrugated with a distance of only ~ 0.085 Å. The corrugation is a result of an excess of 2p electrons normal to the sheet relative to the gas-phase CB_4 molecule.

The quasi-one-dimensional B_2C nanoribbons are finite-size graphene with parallel edges. The width of the B_2C nanoribbon is defined by the number of C atoms normal to the long axis of the ribbon. As shown in Fig. 15, type I and II B_2C nanoribbons are displayed. The dangling bonds at the edges of B_2C nanoribbons can be passivated by either H atoms or CH groups.

TiO_2 with its uses in solar cell applications and surface catalysis has received much attention both theoretically and experimentally. Its nanostructures are of equal interest to technology due to their chemical inertness, endurance, strong oxidizing power, large surface area, high photocatalytic activity, non-toxicity, and low production cost. The titania nanostructures are constructed by “cutting” of TiO_2 monolayers into nanostrips and by rolling them into cylindrical nanotubes or nanorolls (Enyashin and Seifert 2005). There are two different topological nanostrip models constructed from titania. One model is obtained from (101) surface of titania, called as anatase layer, and the second model is obtained from (010) surface of titania, called as lepidocrocite layer. A view of the structures of (101) and (010) TiO_2 layers is shown in Fig. 16. By rolling of these strips, various nanostructures can

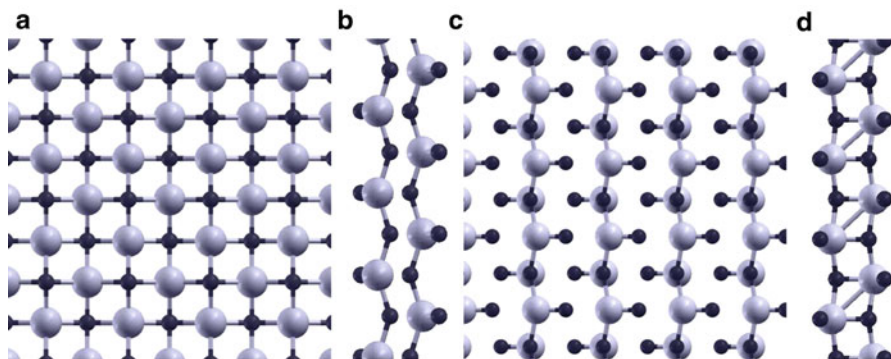


Fig. 16 Monolayers of TiO_2 in anatase (top (a) and side (b)) and lepidocrocite polymorphs (top (c) and side (d))

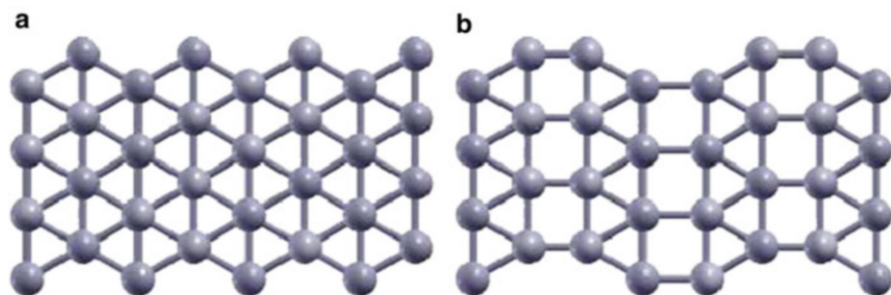


Fig. 17 Idealized 1212 (a) and reconstructed 1221. (b) Boron sheets

be generated, such as nanotubes and nanospirals (nanorolls) (Enyashin and Seifert 2005).

Non-carbon elemental sheet structures have also been investigated. Lau et al. (2006) proposed four possible configuration models for the boron sheet. According to this study, the flat form, denoted $\{1212\}$, seen in Fig. 17 is a triangular network, while the buckled $\{1212\}^b$ and pair-buckled $\{1212\}^{pb}$ configurations include chain-wise and pairwise out-of-plane displacements. Finally, a reconstructed $\{1212\}$ configuration is investigated with inversion symmetry in the unit cell. It can be considered as a triangular–square–triangular unit network.

The DFT calculations reveal (Lau et al. 2006) that the reconstructed $\{1221\}$ configuration is the most stable configuration by 0.23 eV/atom relative to the idealized $\{1212\}$ configuration. Both the $\{1212\}^b$ and $\{1212\}^{pb}$ configurations tend to converge to the idealized $\{1212\}$ configuration when relaxed during the geometry optimization.

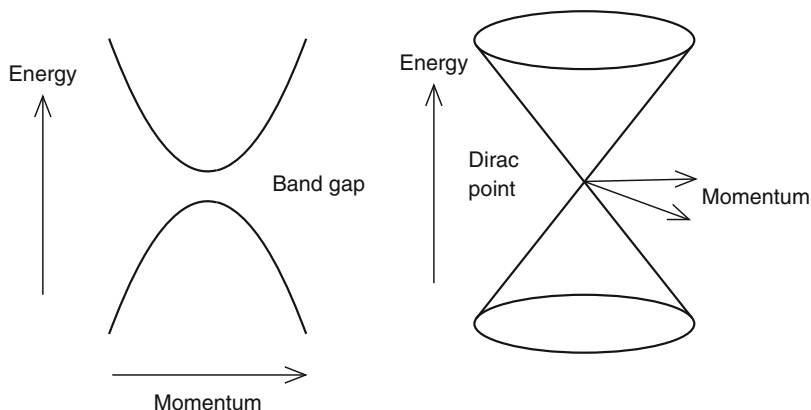


Fig. 18 The band structure of a representative three-dimensional solid (*left*) is parabolic, with a band gap between the lower-energy valence band and the higher-energy conduction band. The energy bands of 2D graphene (*right*) are smooth-sided cones, which meet at the Dirac point

Electronic and Mechanical Properties

Graphene has a unique and curious band structure which can be approximated by a double cone close to the six Fermi points at the corners of the Brillouin zone (see Fig. 18). Commonly referred to as Dirac electrons, the conduction electrons follow a linear energy-momentum dispersion and have a rather large velocity.

The conduction in graphene is enabled by the delocalized π -electrons above and below the plane. Due to their relative detachment from the tightly knit planar network, these electrons are free to move along the graphene sheet with rather high mobility. This is of course a rather desirable property for devices used in electronics (Li et al. 2008b).

A good yet simple method for understanding the band structure of graphene is the tight-binding formalism (Neto et al. 2009). Graphene is made out of carbon atoms arranged in hexagonal structure, as shown in Fig. 19. The structure can be seen as a triangular lattice with a basis of two atoms per unit cell. The lattice vectors can be written as

$$\mathbf{a}_1 = \frac{a}{2} (3, \sqrt{3}) \quad \mathbf{a}_2 = \frac{a}{2} (3, -\sqrt{3}),$$

where $a \sim 1.42 \text{ \AA}$ is the carbon-carbon distance. The reciprocal-lattice vectors are given by

$$\mathbf{b}_1 = \frac{2\pi}{3a} (1, \sqrt{3}) \quad \mathbf{b}_2 = \frac{2\pi}{3a} (1, -\sqrt{3}).$$

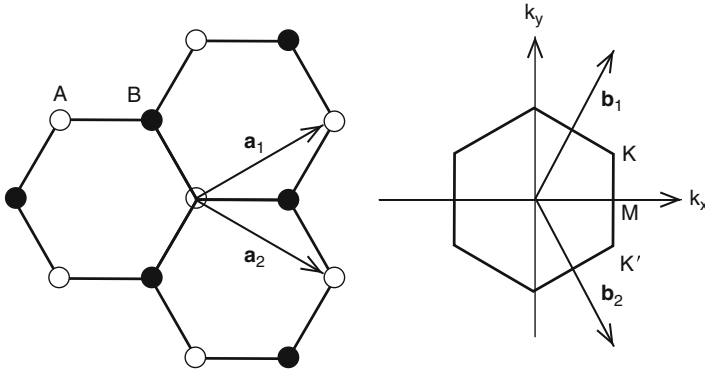


Fig. 19 Graphene lattice and its Brillouin zone. *Left*: lattice structure of graphene, made out of two interpenetrating triangular lattices (\mathbf{a}_1 and \mathbf{a}_2 are the lattice unit vectors). *Right*: corresponding Brillouin zone. The Dirac cones are located at the K and K' points

The two points K and K' at the corners of the graphene Brillouin zone (BZ) are named Dirac points. Their positions in momentum space are given by

$$\mathbf{K} = \left(\frac{2\pi}{3a}, \frac{2\pi}{3\sqrt{3}a} \right), \quad \mathbf{K}' = \left(\frac{2\pi}{3a}, -\frac{2\pi}{3\sqrt{3}a} \right).$$

The tight-binding Hamiltonian for electrons in graphene, considering that electrons can hop to both nearest- and next nearest-neighbor atoms, has the form (in units $\hbar = 1$)

$$H = -t \sum_{\langle i, j \rangle, \sigma} \left(a_{\sigma, i}^\dagger b_{\sigma, j} + H.c. \right) - t' \sum_{\langle\langle i, j \rangle\rangle, \sigma} \left(a_{\sigma, i}^\dagger a_{\sigma, j} + b_{\sigma, i}^\dagger b_{\sigma, j} + H.c. \right), \quad (6)$$

where $a_{i, \sigma}$ ($a_{i, \sigma}^\dagger$) annihilates (creates) an electron with spin σ ($\sigma = \uparrow, \downarrow$) on-site \mathbf{R}_i on sublattice A (an equivalent definition is used for sublattice B), t (~ 2.8 eV) is the nearest-neighbor hopping energy (hopping between different sublattices), and t' (~ 0.1 eV) is the next nearest-neighbor hopping energy (hopping in the same sublattice). The Hamiltonian in Eq. 7 is solved at various momenta and the energy bands are obtained as follows:

$$E_{\pm}(\mathbf{k}) = \pm t \sqrt{3 + f(\mathbf{k})} - t' f(\mathbf{k}) \quad (7)$$

with

$$f(\mathbf{k}) = 2 \cos(\sqrt{3}k_y a) + 4 \cos\left(\frac{\sqrt{3}}{2}k_y a\right) \cos\left(\frac{3}{2}k_x a\right), \quad (8)$$

where the plus sign applies to the upper (π^*) and the minus sign the lower (π) band.

As illustrated by this simple model, graphene is a semimetal or a zero-gap semiconductor. At low temperatures it does not possess superconducting properties; however, as demonstrated by Pathak et al. (2008) using variational Monte Carlo, there is a possibility that doped graphene may superconduct.

The graphene nanoribbons (GNRs) discussed in the previous sections may also be modeled rather easily using the tight-binding formalism taking as the basis the usual Schrödinger's equation (Ezawa 2006) or the massless particle Dirac equation (Sasaki et al. 2006). These models predict that armchair GNRs can be either metallic or semiconducting depending on their widths and that zigzag GNRs with zigzag-shaped edges are metallic regardless of the width.

The edges of GNRs are also suitable sites for chemical functionalization (Wang et al. 2007). Due to the existence of dangling bonds at the edges, the electronic properties of GNRs may be controlled by modifying these bonds by addition of other species.

The results of first-principles calculations using linear combination of atomic orbital density functional theory (DFT) method predict that the electronic band structures and band gaps of the GNRs show a dependence on the edge structure of the nanoribbons (Wu and Zeng 2008). Clearly, the electronic band structures are sensitive to the edge structure of the nanoribbons. All straight GNRs are semiconducting. Two distinct features can be seen concerning the band gap. The band gap of GNR with zigzag edges slightly decreases with increasing the width w , while that of GNR with armchair edges varies periodically as a function of w .

Like the straight GNRs, the calculated electronic band structures of the sawtooth-like GNRs (see Fig. 20) also show semiconducting characteristics with direct band gap. More interestingly, even though the sawtooth-like GNRs have zigzag edges, their band gaps show similar oscillatory behavior as those of straight GNRs with armchair edges, which depend on the width w . However, for most nanoribbons, their band gap reduces monotonically with increasing periodic length l . The band gaps of nanoribbons with $w = 1, 2$, or 4 reduce much rapidly, whereas those of nanoribbons with $w = 3$ or 5 reduce gradually.

For $(2, l)$ sawtooth-like nanoribbons, the band gap approaches zero rapidly as the l increases (Wu and Zeng 2008). These results show that the band gaps of the sawtooth-like GNRs are much more sensitive to their geometric structures, which is an opportunity for tuning the band gap.

Quantization of electric conductance under the action of an external magnetic field displays a rather interesting trend for GNRs. In the exhaustive tight-binding study conducted by Peres et al. (2006a), the number of plateaus in the quantized conductance was found to be even in armchair GNTs, while the same number is odd for the zigzag edge.

Rosales et al. (2008) investigated theoretically the effects of side-attached one-dimensional chains of hexagons pinned at the edges of the GNRs. These one-dimensional chains could be useful to simulate, qualitatively, the effects on the electronic transport of GNRs when benzene-based organic molecules are attached to the edges of the ribbons. They propose a simple scheme to reveal the main electronic properties and the changes in the conductance of such decorated planar structures.

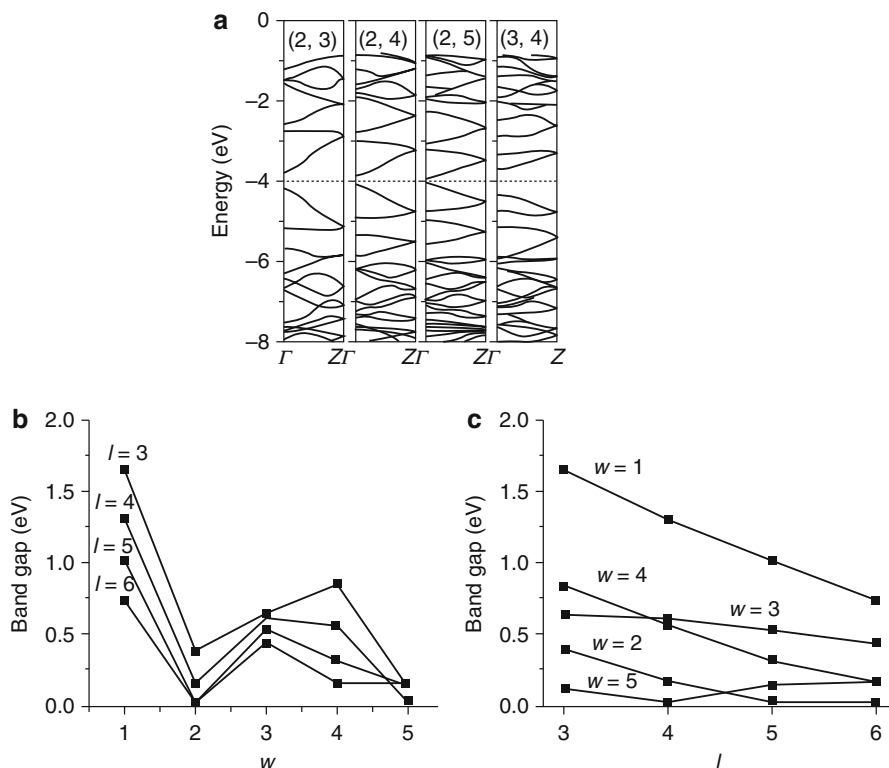


Fig. 20 (a) The electronic band structures of several sawtooth-like GNRs with various w and l . The band gaps versus (b) the width w and (c) the periodic length l (Reprinted with permission from Wu and Zeng (2008). © (2008) by Springer)

For simplicity, they consider armchair and zigzag nanoribbons and linear polyaromatic hydrocarbons (LPHC) and poly(paraphenylene), as the organic molecules. The attached molecules are simulated by simple one-dimensional carbon hexagonal structures connected to the GNRs.

These nanostructures are described using a single-band tight-binding Hamiltonian, and their electronic conductance and density of states are calculated within the Green's function formalism based on real-space renormalization techniques (Rosales et al. 2008).

As revealed by the theoretical analysis conducted by Nakabayashi et al., GNRs are a solution to the difficulty of producing an effective graphene switch for turning off the current. Zigzag nanotubes due to their peculiar band structures may be utilized for better current control rather than graphene. A single layer of graphene has no band gap, which makes it difficult to control current. However, Nakabayashi et al. (2009) found that when a sheet of graphene was in the form of nanoribbons just a few nanometers wide, its electronic structure changed so that currents could be

controlled in radically different ways in comparison to conventional semiconductor devices. Nakabayashi et al. (2009) showed that when a nanoribbon was cut so that its edges formed a zigzag structure with an even number of zigzag chains across its width, and an electrostatic barrier potential applied along part of its length, then it behaved as a so-called band-selective filter, preferentially scattering charges into either even- or odd-numbered bands of its electronic structure depending on the potential. And when two such filters with different potentials were connected in series, they showed it should be possible to completely shut off the flow of current through the nanoribbons.

Tight-binding studies have revealed countless interesting electronic properties regarding GNRs. GNRs, much like single-walled carbon nanotubes, can display metallic or semiconducting properties depending on their orientation and width. Similar to armchair nanotubes, zigzag ribbons are all metallic and may have magnetic properties. Bare and H-terminated ribbons, as studied by Barone et al., may show such effects as gap oscillations which make them a viable choice for possible band structure engineering applications. The results of this study, which is a careful investigation of GNRs of several orientations, widths, and terminations, are displayed in Fig. 21 along with the details of the models used and the band gap profiles of the corresponding models.

GNRs have localized edge states located near the Fermi level. By terminating the edges with different species, one can change the character of these states and consequently the properties of GNRs. In fact, a first-principles calculation within spin-unrestricted local-density functional formalism on zigzag-edged graphene nanostrips terminated with hydrogen and oxygen atoms as well as hydroxyl and imine groups shows that these different species have a significant impact on the electronic structure of these strips near the Fermi level (Gunlycke et al. 2007). Zigzag-edged nanostrips terminated with hydrogen atoms or hydroxyl groups exhibit spin polarization, while the nanostrips terminated with oxygen or imine groups are unpolarized. These differences of course result in very different conductance characteristics for these systems.

In further support of the aforementioned evidence, Zhang and Yang (2009) confirmed in their linear combination of atomic orbitals (LCAO) tight-binding study that H-terminated armchair GNTs exhibit size dependence in their conductance properties.

On the other hand, ZnO nanoribbons show different characteristics from that of carbon nanoribbons. ZnO nanoribbons grown along the [0001] direction can form two different planar monolayer, zigzag and armchair ribbons. First-principles calculations of zinc oxide nanoribbons show that the stability of armchair-edged structures is greater than the zigzag-edged configurations. Furthermore, single-layered armchair ribbons are semiconductors, whereas the zigzag counterparts are metallic (Botello-Mendez et al. 2007). An effect that is not present in purely carbon-based systems is the presence of two differently charged species offering different edge configurations. For the zigzag ZnO ribbons, for instance, the exposed atoms are oxygen atoms while the hidden are zinc. The opposite edge has the inverse

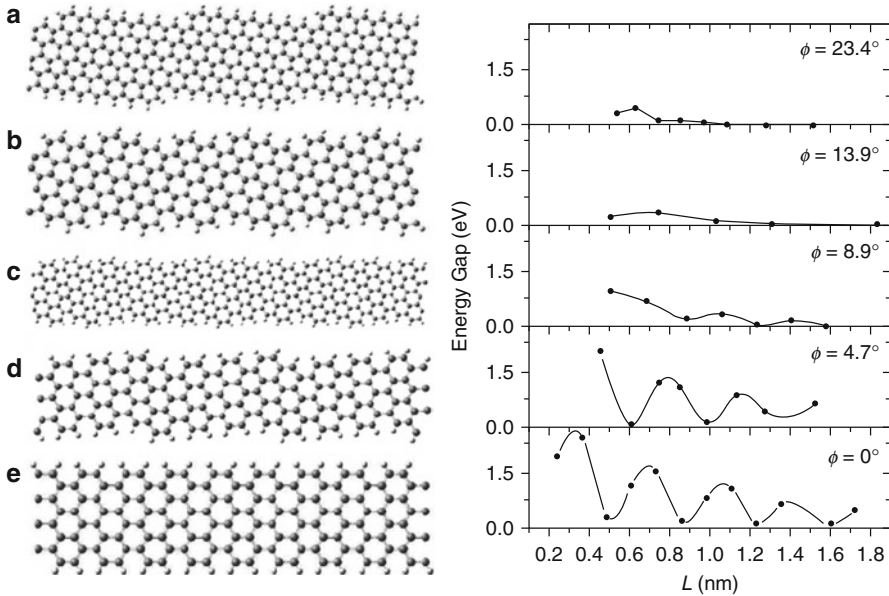


Fig. 21 *Left*: A representative set of semiconducting hydrogen-terminated GNRs, created by “unfolding” and “cutting” different types of CNTs. (a) A GNR with a chiral angle of $\phi = 23.4^\circ$ created by unfolding and cutting a C(6,4). (b) A GNR with a chiral angle of $\phi = 13.9^\circ$ created by unfolding and cutting a C(6,2). (c) A GNR with a chiral angle of $\phi = 8.9^\circ$ created by unfolding and cutting a C(20,4). (d) A GNR with a chiral angle of $\phi = 4.7^\circ$ created by unfolding and cutting a C(10,1). (e) An armchair GNR ($\phi = 0^\circ$) created by unfolding and cutting a zigzag CNT. *Right*: Dependence of the band gap on the width of hydrogen passivated chiral GNRs. The different panels correspond to the different CNRs presented in the *left* (Reprinted with permission from Barone et al. (2006). © (2006) by the American Chemical Society)

structure. The armchair ribbons are characterized by a ZnO pair at the outer edge and another pair at the inner edge. After relaxation, the far edge oxygen ions tend to shift outward on both sides of the ribbon. As in the case of graphene materials, understanding the effects of edge-dimensional variation on band gap energies will provide the new insights into the fundamental principle of architecture design of nanodevices fabricated with ZnO-nanostructured materials. A simple tight-binding model can be used to investigate the electronic properties of nonpolar ZnO nanobelts following the procedure described below.

Consider a simplified unit cell model (see Fig. 22) along the nanobelt growth z direction with Zn and O in the xy plane. This is the actual case for the ultrathin ZnO film oriented along the [001] direction.

In this case, the tight-binding Hamiltonian is given by (Yang and Wang 2006)

$$H = \sum_{r \in A} \varepsilon_r a_r^\dagger a_r + \sum_{r \in B} \varepsilon_r b_r^\dagger b_r + \sum_{r \in A, B} \sum_{i=1,2,\dots} (\tau_{r,i} a_r^+ b_{r+S_i} + \tau_{i,r}^* b_r^+ a_{r+S_i})$$

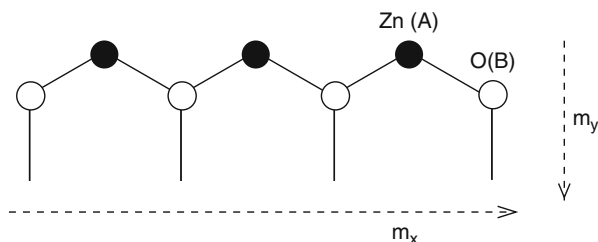


Fig. 22 Schematic illustration showing a tight-binding unit cell for a ZnO nanobelt (Adapted from Yang and Wang 2006)

where operators $a_r^\dagger(a_r)$ and $b_r^\dagger(b_r)$ create (annihilate) a state at sites A (for Zn) and B (for O) at their coordinates r , respectively. ε denotes the on-site self-interaction energy, whereas τ is the interatomic coupling term. Similarly, the wave functions for Zn and O at their respective lattice sites can be written as a linear combination of Bloch waves along the z direction and standing waves along x and y as the following (Yang and Wang 2006),

$$\left| \text{Zn}, k_x, k_y \right\rangle = \sum_{m_x=1}^{M_x} \sum_{m_y=1}^{M_y} \sum_{z_m} c_p e^{ik_z z_m} \sin(k_x m_x a) \sin(k_y m_y b) a_{m_x m_y}^+ \left| 0 \right\rangle$$

$$\left| \text{O}, k_x, k_y \right\rangle = \sum_{n_x=1}^{N_x} \sum_{n_y=1}^{N_y} \sum_{z_n} c_q e^{ik_z z_n} \sin(k_x n_x a) \sin(k_y n_y b) a_{n_x n_y}^+ \left| 0 \right\rangle$$

in which lattice vectors a and b are as defined in Fig. 19 and c_p and c_q are coefficients.

To simplify the model, one may consider only the coupling term between the nearest neighboring sites. In this case, on the same xy plane, the O-sites take the following coordinates around a Zn center: $n_x = m_x - 1$, m_x and $m_x + 1$; with $n_y = m_y - 1$, m_y and $m_y + 1$ in correspondence. Along the yz plane, there exist two O atoms that directly interact with the center Zn with $z_n = z_m \pm c$, respectively (Yang and Wang 2006).

In addition to the conventional strip-like geometry, other configurations based on GNRs have also been proposed as a part of a search for favorable transport properties. Various graphene nanojunctions based on the GNRs have been studied, such as L-shaped, Z-shaped, and T-shaped GRN junctions (Areshkin et al. 2007; Chen et al. 2008; Jayasekera and Mintmire 2007). Of particular interest is the hybrid junctions which are mixtures of armchair and zigzag GNRs. Some studies have shown that the transport properties of these hybrid GNR junctions are very sensitive to the details of the junction region (Xie et al. 2009). A model calculation based on tight binding and Green's function methods for L-shaped GNRs shows that the corner geometry of the L-shaped junction has great influence on the electron

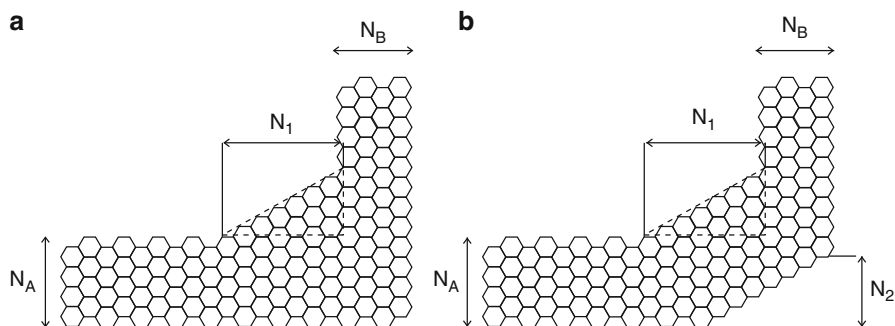


Fig. 23 (a) A deformed LGNR (called LGNR1) where a triangle graphene flake (in the *dashed-line triangle*) is connected to the inside corner of a standard right-angle LGNR. (b) A deformed LGNR (called LGNR2) in which a triangle graphene flake is cut from the outside corner of LGNR1 in (a). N_A and N_Z represent the widths of the semi-infinite AGNR and ZGNR, respectively; N_1 represents the side length of added triangle flake, while N_2 represents the side length of cut triangle flake (Adapted from Xie et al. 2009)

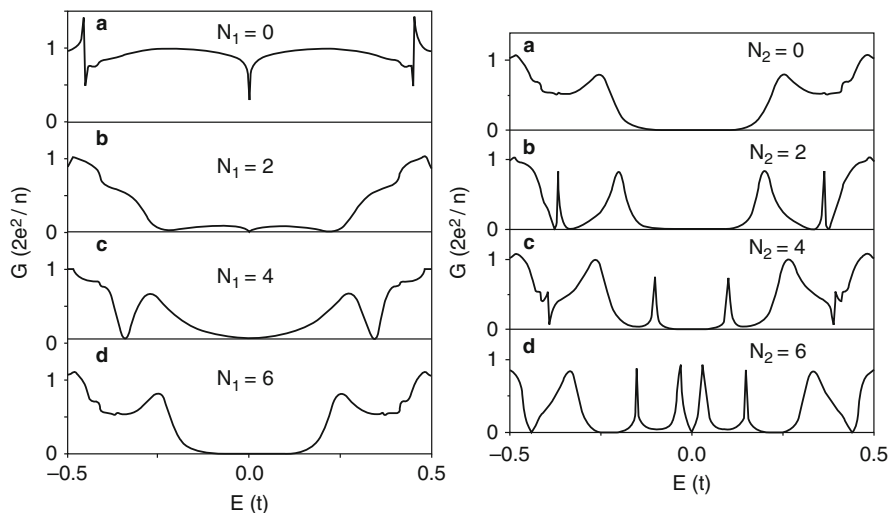


Fig. 24 *Left*: The conductance as a function of electron energy for the LGNR1 in Fig. 23a with different N_1 . Other parameters are $N_A = 11$ and $N_Z = 7$. *Right*: The conductance as a function of electron energy for the LGNR2 in Fig. 23b with $N_1 = 6$ and different N_2 . Other parameters are $N_A = 11$ and $N_Z = 7$ (Reprinted with permission from Xie et al. (2009). © (2009) by Elsevier)

transport around the Fermi energy (Xie et al. 2009). The L-shaped GNR models considered are shown in Fig. 23. The calculated conductance plots as a function of electron energy are shown in Fig. 24. Interestingly the results reveal that the corner is the decisive factor for the electronic properties of these junctions. As more carbon atoms are added to the inner side of the corner, the junction moves from being metal to a semiconductor, developing a band gap.

Carbon nanoribbons have also interesting mechanical properties. For instance, Raman spectra show a remarkable dependence on ribbon width (W) (Zhou and Dong 2008). For the GNRs whose widths are larger than 25 Å, the radial breathing-like mode (RBLM) frequencies follow the $1/W$ rule:

$$\omega = 3086.97 \times \frac{1}{W} + 1.08 \text{ in } 1/\text{cm}.$$

But for the narrow GNRs whose widths are less than 25 Å, their RBLM frequencies follow the $1/\sqrt{W}$ rule:

$$\omega = 1407.81 \times \frac{1}{\sqrt{W}} + 164.38 \text{ in } 1/\text{cm}.$$

A unified fitting function has been proposed (Zhou and Dong 2008), which can be suitable for all the GNRs:

$$\omega = 1584.24 \times \frac{1}{W} + 351.98 \times \frac{1}{\sqrt{W}} - 10.0 \text{ in } 1/\text{cm}.$$

Another interesting application of graphene is the antidot lattices. Antidot lattices are triangular arrangements of holes in an otherwise perfect graphene sheet. In each hexagonal unit cell of the lattice, a circular hole is introduced whose radius R can be adjusted along with the length of the hexagonal unit cell, L . Previously known applications of antidot lattices involve semiconductor lattices, while in a recent publication, the possibility of creating antidot structures on graphene using e-beam lithography has been reported (Pedersen et al. 2008b). The antidot lattice causes graphene to become a semiconductor, introducing a tuning parameter which is the dot size and shape. In addition spin qubit states can be formed in the antidot lattice by manipulating the antidot arrangement. Pedersen et al. (2008a) demonstrated that the hole shape plays a crucial role in determining the electronic structure. A triangular hole instead of a circular hole results in the appearance of a metallic state at the Fermi level. The different band structures are illustrated in Fig. 25.

Magnetic and Optical Properties

Lower-dimensional systems are expected to show electronic, magnetic, and optical properties that are not observed in their bulk counterparts. In two-dimensional structures, the inherent or emergent magnetic properties depend on many factors including shape, size, and the interaction between subregions (Jensen and Pastor 2003; Zhao et al. 2014).

GNRs are a typical example of emergent and controllable magnetic systems. Spin-polarized first-principles calculations by Gorjizadeh et al. (2008) have shown that doping GNRs with 3d transition metals result in the appearance of FM or AFM

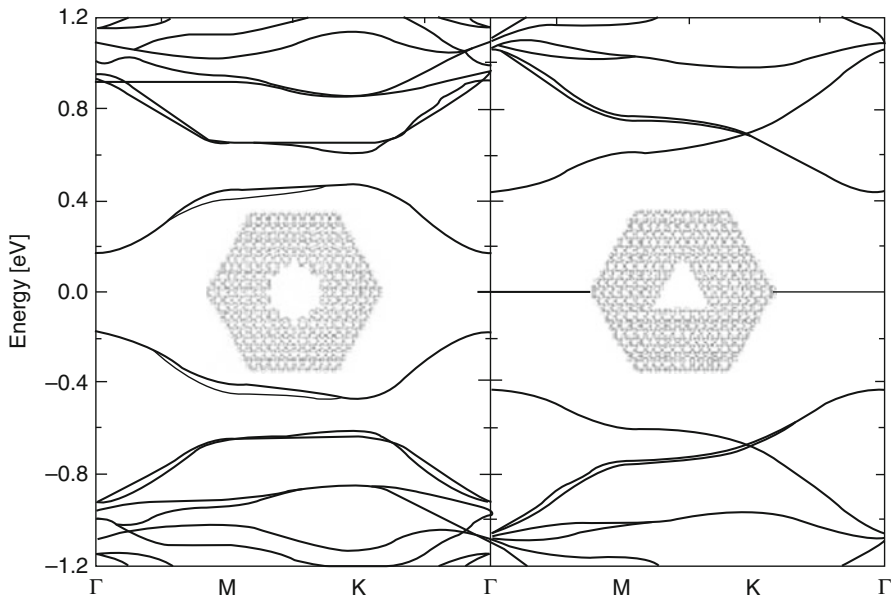


Fig. 25 Band structures of a $\{10, 3\}$ antidot lattice and similar structure having a triangular hole with zigzag edges. Note the dispersionless band at 0 eV in the triangular case (Reprinted with permission from Pedersen et al. (2008a). © (2008) by the American Physical Society)

states. In particular Cr and Co provide a large magnetic moment and doping with Fe or Mn at low densities yields half-metallic ribbons.

As discussed in the work by Enoki and Takai, as the dimension of the system decreases, the contribution coming from the edge states increases in proportion (Enoki and Takai 2009). The localized edge states in GNR, which are dispersionless states that appear at the Fermi level, give rise to interesting magnetic properties.

From the application's point of view, the magnetic properties of graphene nanoribbons are more interesting and studied both experimentally and theoretically. Particularly quasi-1D GNRs have magnetic properties depending on their size and symmetry. These are edge states of nanoribbons with opposite spin polarization and band gaps varying with the width of the ribbon (Topsakal et al. 2008).

Using first-principles plane wave calculations within the DFT method, Topsakal et al. (2008) predict that in addition to edge states, electronic and magnetic properties of graphene nanoribbons can also be affected by defect-induced states. In particular, when H-saturated holes are introduced into the GNT, the band structure is modified dramatically altering in return the electronic and magnetic properties. Similarly, vacancies and divacancies induce metallization and magnetization in nonmagnetic semiconducting nanoribbons due to the spin polarization of local defect states. Antiferromagnetic ground state of semiconducting zigzag ribbons can change to ferrimagnetic state upon creation of vacancy defects. In this study, the changes in electronic properties are studied as a function of the location and the

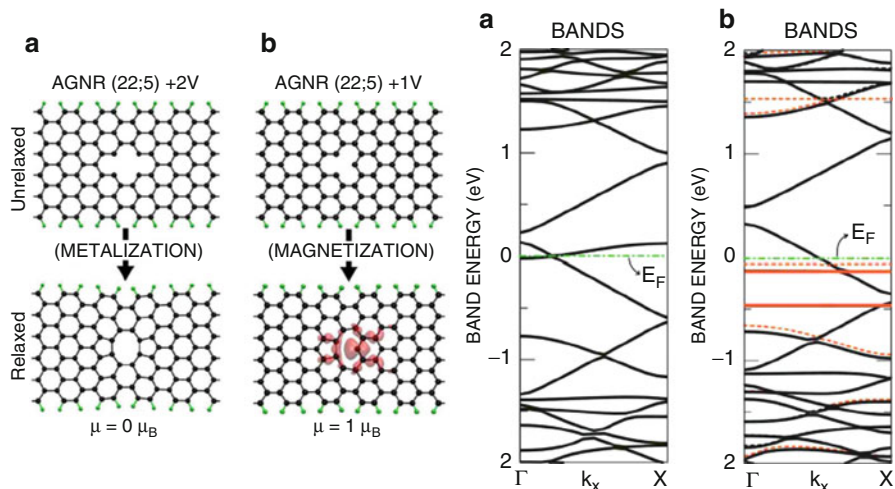


Fig. 26 (a) Metallization of the semiconducting AGNR(22) by the formation of divacancies with repeat period of $l = 5$. (b) Magnetization of the nonmagnetic AGNR(22) by a defect due to the single carbon atom vacancy with the same repeat periodicity. Isosurfaces around the vacancy correspond to the difference of the total charge density of different spin directions. Solid (blue) and dashed (red) lines are for spin-up and spin-down bands; solid (black) lines are nonmagnetic bands (Reprinted with permission from Topsakal et al. (2008). © (2008) by the American Physical Society)

geometry of the vacancies in different types of armchair GNRs. Due to the spin polarization of localized states and their interaction with edge states, magnetization may be introduced into the GNRs. Some of the representative results of this work are displayed in Fig. 26.

Not only armchair but also zigzag nanoribbons are strongly affected by defects due to single and multiple vacancies (Topsakal et al. 2008). When coupled with the magnetic edge states of the zigzag nanoribbons, the vacancy defect brings about additional changes. The magnetic state and energy band structure of these ribbons depend on the type and geometry of the defects. In a ZGNR(14;8), a single vacancy formation energy is lowered by 0.53 eV when the defect is situated at the edge rather than at the center of the ribbon. Furthermore, two defects associated with two separated vacancy and a defect associated with a relaxed divacancy exhibit similar behavior (Topsakal et al. 2008).

On the other hand, Pisani et al. (2007) investigated theoretically the electronic and magnetic properties of zigzag graphene nanoribbons by performing first-principles calculations within DFT formalism. They predict that the electronic structure of graphene ribbons with zigzag edges is unstable with respect to magnetic polarization of the edge states.

As can be seen from the model calculations of Topsakal et al. (2008, 2007), the energy band gaps and magnetic states of graphene nanoribbons can be modified

by defects due to single or multiple vacancies. Electronic and magnetic properties of finite length graphene nanoribbons also show similar behavior as infinite length nanoribbons. Finite length ribbons are usually referred to as *graphene quantum dots* (GQD). Tang et al. (2008) investigated theoretically the electronic and magnetic properties of a square graphene quantum dot. Electronic eigen-states of a GQD terminated by both zigzag and armchair edges are derived in the theoretical framework of the Dirac equation. They find that the Dirac equation can determine the eigen-energy spectrum of a GQD with good accuracy. By using the Hartree–Fock mean-field approach, they studied the size dependence of the magnetic properties of GQDs. They find that there exists a critical width between the two zigzag edges for the onset of the stable magnetic ordering. On the other hand, when such a width increases further, the magnetic ground state energy of a charge neutral GQD tends to a saturated value (Tang et al. 2008).

Magnetic properties of graphene show a sensitive dependence on single-atom defects; defect concentration and packing play an important role in magnetism. Singh and Kroll (2009) investigated the magnetism in graphene due to single-atom defects by using spin-polarized density functional theory calculations. Interestingly, they find that while the magnetic moment per defect due to substitutional atoms and vacancies depends on the defect density, it is independent of defect density for adatoms. The graphene sheet with B adatoms is found to be nonmagnetic, but with C and N adatoms, it is magnetic. The adatom defects cause a distortion of the graphene sheet in their vicinity. The distortion in graphene due to C and N adatoms is significant, while the distortion due to B adatoms is very small. The vacancy and substitutional atom (B, N) defects in graphene are planar in the sense that there is in-plane displacement of C atoms near the vacancy and substitutional defects. Upon relaxation the displacement of C atoms and the formation of pentagons near the vacancy site due to Jahn–Teller distortion depend upon the density and packing geometry of vacancies (Singh and Kroll 2009). The defect models considered by Singh and Kroll are shown in Fig. 27.

The optical absorption coefficient is one of the most important quantities in solids and is closely related to the electronic band structure. The Kubo formula-based calculations of Zhang et al. indicate that at high frequencies, the optical spectrum of graphene is highly anisotropic (Zhang et al. 2008a). It has also been shown that the

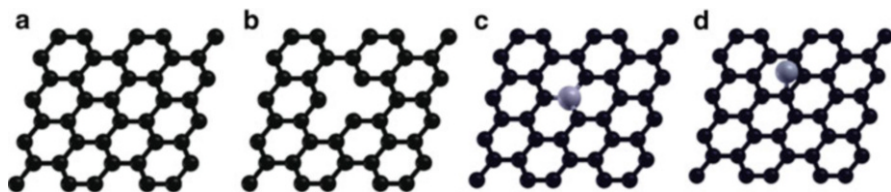


Fig. 27 The model samples of (a) ideal graphene, (b) graphene with one atom vacancy, (c) graphene with one substitutional atom, and (d) graphene with one adatom defect (Adapted from Singh and Kroll 2009)

weak optical response in graphene nanoribbons can be significantly enhanced in an applied magnetic field (Liu et al. 2008).

Another effect important for determining the optical properties of graphene-related materials is spin-orbit coupling. Although spin-orbit coupling is negligible at high frequencies, it is found to have a significant enhancing effect on the optical absorption at lower frequencies as shown in a direct solution of the spin-orbit Hamiltonian in the effective mass approximation by Wright et al. (2009).

Adsorption Phenomena

Adsorption of adatoms on graphene and/or graphene nanoribbons plays an important role in functionalizing graphene materials for various device applications, such as gas sensors and spin valves. Transition metal (TM) atom-decorated graphene shows different magnetic properties depending on the concentration and the species of TM atoms. There are several coverage models, such as one TM atom adsorbed on either (2×2) or (4×4) unit cells on only one side as well as on both sides, namely, above and below the graphene. Sevincli et al. (2008) investigated the possible adsorption sites of TM atoms on graphene and GNR. The geometrical configurations of the structures they considered include bridge (over a C–C bond), atop (on top of a C atom), and on center (over the center of a hexagon) adsorption sites for both perfect graphene and for AGNR of various width N_a .

In relation to another very important application, doping of transition elements was found to increase the hydrogen storage properties of materials (Shevlin and Guo 2006). Especially nickel and rhodium are widely used in the hydrogenation reaction and also in the synthesis of BN sheets. Considering the potential application of Ni and Rh nanoparticles in hydrogen storage and in catalysis, Venkataramanan et al. (2009) investigated the interaction of Ni and Rh atom on the BN sheets through first-principles calculations. They also analyzed the interaction between hydrogen molecules and the metal atoms adsorbed on the BN sheets, which might be useful to maximize the hydrogen storage capacity.

Hydrogen adsorption studies over the metal-doped BN sheets shows that both Ni and Rh atoms can hold three hydrogen molecules. In the case of Ni-doped BN sheet, all three hydrogen molecules are chemically bound and are intact. In the case of Rh-doped BN sheets, the first hydrogen molecule dissociates and the remaining hydrogen molecules are bound to the metal atom. The absorption energy for the first hydrogen was found to have a larger value for the Rh-doped BN sheet, whereas for the second and third hydrogen molecules, the absorption energies were higher for the Ni-doped BN sheets. Upon addition of a fourth hydrogen molecule to the Ni-doped BN sheets, the fourth hydrogen molecule moved to a distance of 3.855 Å. In the case of Rh-doped BN sheet, the Rh atom detached and acted in a similar way to a cluster. Thus Ni atoms are more stable on BN sheets and have higher absorption energy compared to the Rh-doped BN sheets (Venkataramanan et al. 2009).

Metallized graphene can be a potential high-capacity hydrogen storage medium. Graphene is metallized through charge donation by adsorbed Li atoms to its π^* bands. Each positively charged Li ion can bind up to four H_2 by polarizing these molecules. The storage capacity up to the gravimetric density of 12.8 wt% is possible with a favorable average H_2 binding energy of 0.21 eV (Ataca et al. 2008).

Water and gas molecules adsorbed on nanoscale graphene play the role of defects which facilitate the tunability of the band gap and allow one to control the magnetic ordering of localized states at the edges (Berashevich and Chakraborty 2009). The adsorbed molecules push the α -spin (up) and β -spin (down) states of graphene to the opposite (zigzag) edges such that the α - and β -spin states are localized at different zigzag edges. This breaks the symmetry that results in the opening of a large gap. The efficiency of the wavefunction displacement depends strongly on the type of molecules adsorbed on graphene (Berashevich and Chakraborty 2009). The influence of adsorption of water on the electronic and magnetic properties of graphene is based on calculation of the spin-polarized density functional theory, and the results of the calculations are depicted in Fig. 28.

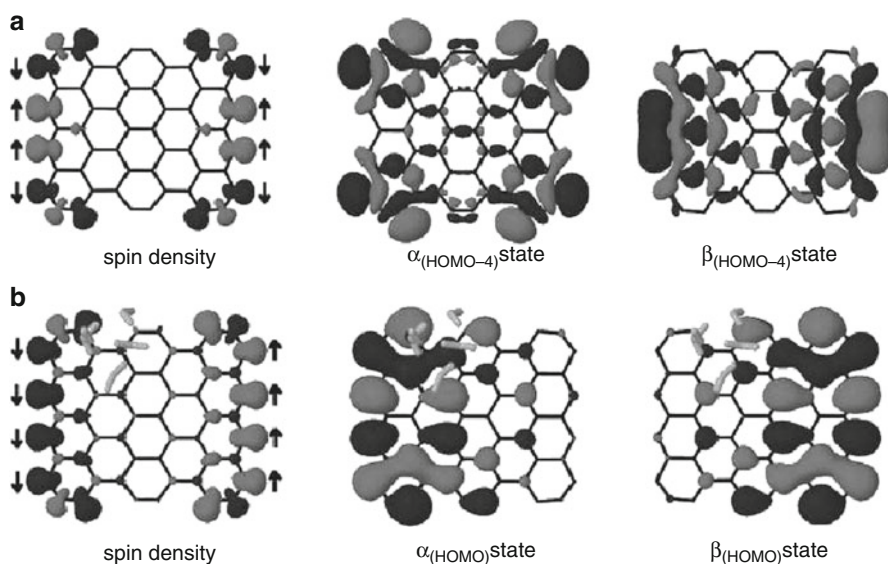


Fig. 28 Plots of the spin density and electron density for the α -spin (*up*) and β -spin (*down*) states in nanoscale graphene. For the spin density, different colors indicate the α - and β -states, while for the electron density distribution (α - and β -states), the different colors correspond to different signs of the molecular orbital lobes. (a) The electron density distribution in pure graphene. (b) The electron density distribution for graphene with adsorbed water (Reprinted with permission from Berashevich and Chakraborty (2009). © (2009) by the American Physical Society)

Bibliography

- Amorim, R. G., Fazzio, A., Antonelli, A., Novaes, F. D., & da Silva, A. J. R. (2007). Divacancies in graphene and carbon nanotubes. *Nano Letters*, 7, 2459.
- Anagnostatos, G. S. (1987). Magic numbers in small clusters of rare-gas and alkali atoms. *Physics Letters A*, 124, 85.
- Andzelm, J., Govind, N., & Maiti, A. (2006). Nanotube-based gas sensors – Role of structural defects. *Chemical Physics Letters*, 421, 58.
- Arantes, J. T., & Fazzio, A. (2007). Theoretical investigations of Ge nanowires grown along the [110] and [111] directions. *Nanotechnology*, 18, 295706.
- Areshkin, D. A., Gunlycke, D., & White, C. T. (2007). Ballistic transport in graphene nanostrips in the presence of disorder: Importance of edge effects. *Nano Letters*, 7, 204.
- Ataca, C., Akturk, E., Ciraci, S., & Ustunel, H. (2008). High-capacity hydrogen storage by metallized graphene. *Applied Physics Letters*, 93, 043123.
- Balasubramanian, K. (1990). Spectroscopic constants and potential-energy curves of heavy p-block dimers and trimers. *Chemical Reviews*, 90, 93.
- Baletto, F., & Ferrando, R. (2005). Structural properties of nanoclusters: Energetic, thermodynamic, and kinetic effects. *Reviews of Modern Physics*, 77, 371.
- Bansmanna, J., Bakerb, S. H., Binnsb, C., Blackmanc, J. A., Bucherd, J. P., Dorantes-Da`vilae, J., Dupuisf, V., Favref, L., Kechrakosg, D., Kleibert, A., Meiwes-Broera, K. H., Pastore, G. M., Perezf, A., Toulemonded, O., Trohidoug, K. N., Tuailonf, J., & Xiee, Y. (2005). Magnetic and structural properties of isolated and assembled clusters. *Surface Science Reports*, 56, 189.
- Barone, V., Hod, O., & Scuseria, G. E. (2006). Electronic structure and stability of semiconducting graphene nanoribbons. *Nano Letters*, 6, 2748.
- Barreteau, C., Desjonqueres, M. C., & Spanjaard, D. (2000). Theoretical study of the icosahedral to cuboctahedral structural transition in Rh and Pd clusters. *The European Physical Journal D*, 11, 395.
- Barth, S., Harnagea, C., Mathur, S., & Rosei, F. (2009). The elastic moduli of oriented tin oxide nanowires. *Nanotechnology*, 20, 115705.
- Baumeier, B., Kruger, P., & Pollmann, J. (2007). Structural, elastic, and electronic properties of SiC, BN, and BeO nanotubes. *Physical Review B*, 76, 085407.
- Berashevich, J., & Chakraborty, T. (2009). Tunable bandgap and magnetic ordering by adsorption of molecules on graphene. *Physical Review B*, 80, 033404.
- Bonaccorso, F., Colombo, L., Yu, G., Stoller, M., Tozzini, V., Ferrari, A. C., Ruoff, R. S., & Pellegrini, V. (2015). Graphene, related two-dimensional crystals, and hybrid systems for energy conversion and storage. *Science*, 347, 41.
- Bonacic-Koutecky, V., Fantucci, P., & Koutecky, J. (1991). Quantum chemistry of small clusters of elements of groups Ia, Ib, and IIa: Fundamental concepts, predictions, and interpretation of experiments. *Chemical Reviews*, 91, 1035.
- Botello-Mendez, A. R., Martinez-Martinez, M. T., Lopez-Urias, F., Terrones, M., & Terrones, H. (2007). Metallic edges in zinc oxide nanoribbons. *Chemical Physics Letters*, 448, 258.
- Brack, M. (1993). The physics of simple metal clusters: Self-consistent jellium model and semiclassical approaches. *Reviews of Modern Physics*, 65, 677.
- Brenner, D. (1989). Relationship between the embedded-atom method and Tersoff potentials. *Physical Review Letters*, 63, 1022.
- Briere, T. M., Sluiter, M. H., Kumar, V., & Kawazoe, Y. (2002). Atomic structures and magnetic behavior of Mn clusters. *Physical Review B*, 66, 064412.
- Bromley, S. T., & Flikkema, E. (2005). Columnar-to-disk structural transition in nanoscale (SiO₂)N Clusters. *Physical Review Letters*, 95, 185505.
- Bruno, M., Palummo, M., Ossicini, S., & del Sole, E. (2007). First-principles optical properties of silicon and germanium nanowires. *Surface Science*, 601, 2707.
- Bulusu, S., Yoo, S., & Zeng, X. C. (2005). Search for global minimum geometries for medium sized germanium clusters: Ge₁₂-Ge₂₀. *Journal of Chemical Physics*, 122, 164305.

- Catlow, C. R. A., Bromley, S. T., Hamad, S., Mora-Fonz, M., Sokola, A. A., & Woodley, S. M. (2010). Modelling nano-clusters and nucleation. *Physical Chemistry Chemical Physics*, *12*, 786.
- Chan, T.-L., Ciobanu, C. V., Chuang, F.-C., Lu, N., Wang, C.-Z., & Ho, K.-M. (2006). Magic structures of H-passivated (110) silicon nanowires. *Nano Letters*, *6*, 277.
- Chen, Y. P., Xie, Y. E., & Yan, X. H. (2008). Electron transport of L-shaped graphene nanoribbons. *Journal of Applied Physics*, *103*, 063711.
- Chopra, N. G., Luyken, R. J., Cherrey, K., Crespi, V. H., Cohen, M. L., Louie, S. G., & Zettl, A. (1995). Boron-nitride nanotubes. *Science*, *269*, 966.
- Corso, M., Auwarter, W., Muntwiler, M., Tamai, A., Greber, T., & Osterwalder, J. (2004). Boron nitride nanomesh. *Science*, *303*, 217.
- Darby, S., Mortimer-Jones, T. V., Johnston, R. L., & Roberts, C. (2002). Theoretical study of Cu-Au nanoalloy clusters using a genetic algorithm. *Journal of Chemical Physics*, *116*, 1536.
- Dasgupta, N. P., Sun, J., Liu, C., Brittan, S., Andrews, S. C., Lim, J., Gao, H., Yan, R., & Yang, P. (2014). 25th anniversary article: Semiconductor nanowires – Synthesis, characterization, and applications. *Advanced Materials*, *26*, 2137.
- de Heer, W. A. (1993). The physics of simple metal clusters: Experimental aspects and simple models. *Reviews of Modern Physics*, *65*, 611.
- de Oliveira, I. S. S., & Miwa, R. H. (2009). Boron and nitrogen impurities in SiC nanowires. *Physical Review B*, *79*, 085427.
- Doye, J. P. K. (2003). Identifying structural patterns in disordered metal clusters. *Physical Review B*, *68*, 195418.
- Doye, J. P. K. (2006). Lead clusters: Different potentials, different structures. *Computational Materials Science*, *35*, 227.
- Doye, J. P. K., & Wales, D. J. (1999). Structural transitions and global minima of sodium chloride clusters. *Physical Review B*, *59*, 2292.
- Du, J., Shen, N., Zhu, L., & Wang, J. (2010). Emergence of noncollinear magnetic ordering in bimetallic $\text{CO}_{6-n}\text{Mn}_n$ clusters. *Journal of Physics D*, *43*, 015006.
- Dugan, N., & Erkoc, S. (2008). Stability analysis of graphene nanoribbons by molecular dynamics simulations. *Physica Status Solidi B*, *245*, 695.
- Dugan, N., & Erkoc, S. (2009). Genetic algorithms in application to the geometry optimization of nanoparticles. *Algorithms*, *2*, 410.
- Enoki, T., & Takai, K. (2009). The edge state of nanographene and the magnetism of the edge-state spins. *Solid State Communications*, *149*, 1144.
- Enyashin, A. N., & Seifert, G. (2005). Structure, stability and electronic properties of TiO_2 nanostructures. *Physica Status Solidi B*, *242*, 1361.
- Erkoc, S. (1997). Empirical many-body potential energy function used in computer simulations of condensed matter properties. *Physics Reports*, *278*, 79.
- Erkoc, S. (2001a). Empirical potential energy functions used in the simulations of materials properties. In D. Stauffer (Ed.), *Annual review of computational physics* (Vol. IX, pp. 1–103). Singapore: World Scientific.
- Erkoc, S. (2001b). Structural and electronic properties of single-wall BN nanotubes. *Journal of Molecular Structure (THEOCHEM)*, *542*, 89.
- Erkoc, S. (2003). Structural and electronic properties of “benzorods”. *Journal of Molecular Structure (THEOCHEM)*, *639*, 157.
- Erkoc, S. (2004a). Does tubular structure of carbon form only from graphene sheet? *Physica E*, *25*, 69.
- Erkoc, S. (2004b). Semi-empirical SCF-MO calculations for the structural and electronic properties of single-wall InP nanotubes. *Journal of Molecular Structure (THEOCHEM)*, *676*, 109.
- Erkoc, S., & Kokten, H. (2005). Structural and electronic properties of single-wall ZnO nanotubes. *Physica E*, *28*, 162.
- Erkoc, S., Leblebicioglu, K., & Halici, U. (2003). Application of genetic algorithms to geometry optimization of microclusters: A comparative study of empirical potential energy functions for silicon. *Materials and Manufacturing Processes*, *18*, 329.

- Erkoc, S., Malcioglu, O. B., & Tasci, E. (2004). Structural and electronic properties of single-wall GaN nanotubes: Semi-empirical SCF-MO calculations. *Journal of Molecular Structure (THEOCHEM)*, *674*, 1.
- Ezawa, M. (2006). Peculiar width dependence of the electronic properties of carbon nanoribbons. *Physical Review B*, *73*, 045432.
- Fa, W., Luo, C., & Dong, J. (2005). Structure-dependent ferroelectricity of niobium clusters (Nb_N , $N = 2-52$). *Physical Review B*, *71*, 245415.
- Fa, W., Luo, C., & Dong, J. (2006). Coexistence of ferroelectricity and ferromagnetism in tantalum clusters. *Journal of Chemical Physics*, *125*, 114305.
- Farrow, M. R., Chow, Y., & Woodley, S. M. (2014). Structure prediction of nanoclusters; a direct or a pre-screened search on the DFT energy landscape? *Physical Chemistry Chemical Physics*, *16*, 21119.
- Finnis, M. W., & Sinclair, J. E. (1984). A simple empirical n-body potential for transition metals. *Philosophical Magazine A*, *50*, 45.
- Freitag, M. (2008). Nanoelectronics goes flat out. *Nature Nanotechnology*, *3*, 455.
- Gafner, Y. Y., Gafner, S. L., & Entel, P. (2004). Formation of an icosahedral structure during crystallization of nickel nanoclusters. *Physics of the Solid State*, *46*, 1327.
- Gao, G., & Kang, H. S. (2008). First-principles study of silicon nitride nanotubes. *Physical Review B*, *78*, 165425.
- Geim, A. K., & Novoselov, K. S. (2007). The rise of graphene. *Nature Materials*, *6*, 183.
- Ghosh, P., Kahaly, M. U., & Waghmare, U. V. (2007). Atomic and electronic structures, elastic properties, and optical conductivity of bulk Te and Te nanowires: A first-principles study. *Physical Review B*, *75*, 245437.
- Gorjizadeh, N., Farajian, A. A., Esfarjani, K., & Kawazoe, Y. (2008). Spin and band-gap engineering in doped graphene nanoribbons. *Physical Review B*, *78*, 155427.
- Gunlycke, D., Li, J., Mintmire, J. W., & White, C. T. (2007). Altering low-bias transport in zigzag-edge graphene nanostrips with edge chemistry. *Applied Physics Letters*, *91*, 112108.
- Haberland, H. (1994). *Clusters of atoms and molecules*. Berlin: Springer.
- Halicoglu, T., & Bauschlicher, C. W. (1988). Physics of microclusters. *Reports on Progress in Physics*, *51*, 883.
- Han, M. Y., Ozyilmaz, B., Zhang, Y. B., & Kim, P. (2007). Energy band-gap engineering of graphene nanoribbons. *Physical Review Letters*, *98*, 206805.
- Hartke, B. (2002). Structural transitions in clusters. *Angewandte Chemie (International Ed. in English)*, *41*, 1468.
- Hestenes, M. R., & Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, *49*, 409.
- Hsu, P. J., & Lai, S. K. (2006). Structures of bimetallic clusters. *Journal of Chemical Physics*, *124*, 044711.
- Hu, J., Liu, X. W., & Pan, B. C. (2008). A study of the size-dependent elastic properties of ZnO nanowires and nanotubes. *Nanotechnology*, *19*, 285710.
- Huda, M. N., & Kleinman, L. (2006). h-BN monolayer adsorption on the Ni(111) surface: A density functional study. *Physical Review B*, *74*, 075418.
- Hyun, J. K., Zhang, S., & Lauthon, L. J. (2013). Nanowire heterostructures. *Annual Review of Materials Research*, *43*, 451.
- Ince, A., & Erkoc, S. (2011). Silicene nanoribbons: Molecular-dynamics simulations. *Computational Materials Science*, *50*, 865.
- Jarrold, M. F., & Constant, V. A. (1991). Silicon cluster ions: Evidence for a structural transition. *Physical Review Letters*, *67*, 2994.
- Jayasekera, T., & Mintmire, J. W. (2007). Transport in multiterminal graphene nanodevices. *Nanotechnology*, *18*, 424033.
- Jensen, P. J., & Pastor, G. M. (2003). Low-energy properties of two-dimensional magnetic nanostructures: Interparticle interactions and disorder effects. *New Journal of Physics*, *5*, 68.

- Jia, X., Hofmann, M., Meunier, V., Sumpster, B. G., Campos-Delgado, J., Romo-Herrera, J. M., Son, H., Hsieh, Y.-P., Reina, A., Kong, J., Terrones, M., & Dresselhaus, M. S. (2009). Controlled formation of sharp zigzag and armchair edges in graphitic nanoribbons. *Science*, 323, 1701.
- Johnston, R. L. (2002). *Atomic and molecular clusters*. London: Taylor & Francis.
- Kabir, M., Mookerjee, A., & Kanhere, D. G. (2006). Structure, electronic properties, and magnetic transition in manganese clusters. *Physical Review B*, 73, 224439.
- Kagimura, R., Nunes, R. W., & Chacham, H. (2007). Surface dangling-bond states and band Lineups in Hydrogen-Terminated Si, Ge, and Ge/Si nanowires. *Physical Review Letters*, 98, 026801.
- Kempa, T. J., Day, R. W., Kim, S.-K., Park, H.-G., & Lieber, C. M. (2013). Semiconductor nanowires: A platform for exploring limits and concepts for nano-enabled solar cells. *Energy and Environmental Science*, 6, 719.
- Kilic, M. E., & Erkoç, S. (2013a). Molecular dynamics simulations of ZnO nanostructures under strain: II – Nanorods. *Journal of Computational and Theoretical Nanoscience*, 10, 112.
- Kilic, M. E., & Erkoç, S. (2013b). Molecular dynamics simulations of ZnO nanostructures under strain: I – Nanoribbons. *Journal of Computational and Theoretical Nanoscience*, 10, 104.
- Kim, D.-H., & Rogers, J. A. (2008). Stretchable electronics: Materials strategies and devices. *Advanced Materials*, 20, 4887.
- Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220, 671.
- Knickerbein, M. B. (2001). Experimental observation of superparamagnetism in manganese clusters. *Physical Review Letters*, 86, 5255.
- Kohler, C., Seifer, G., & Frauheim, T. (2006). Magnetism and the potential energy hypersurfaces of Fe₅₃ to Fe₅₇. *Computational Materials Science*, 35, 297.
- Krainyukova, N. V. (2006). The crystal structure problem in noble gas nanoclusters. *Thin Solid Films*, 515, 1658.
- Kulkarni, A. J., Zhou, M., & Ke, F. J. (2005). Orientation and size dependence of the elastic properties of zinc oxide nanobelts. *Nanotechnology*, 16, 2749.
- Kumar, V., & Kawazoe, Y. (2002). Icosahedral growth, magnetic behavior, and adsorbate-induced metal-nonmetal transition in palladium clusters. *Physical Review B*, 66, 144413.
- Landau, L. D., & Lifshitz, E. M. (1980). *Statistical physics, part I, sections 137 and 138*. Oxford: Pergamon.
- Lau, K. C., Pati, R., Pandey, R., & Pineda, A. C. (2006). First-principles study of the stability and electronic properties of sheets and nanotubes of elemental boron. *Chemical Physics Letters*, 418, 549.
- Law, M., Goldberger, J., & Yang, P. (2004). Semiconductor nanowires and nanotubes. *Annual Review of Materials Research*, 34, 83.
- Lee, B., & Rudd, R. E. (2007). First-principles calculation of mechanical properties of Si_i001_i nanowires and comparison to nanomechanical theory. *Physical Review B*, 75, 195328.
- Lennard-Jones, J. E. (1931). Cohesion. *Proceedings of the Physical Society*, 43, 461.
- Li, F., Zhu, Z., Zhao, M., & Xia, Y. (2008a). Ab initio calculations on the magnetic properties of hydrogenated boron nitride nanotubes. *Journal of Physical Chemistry C*, 112, 16231.
- Li, Z. Q., Henriksen, E. A., Jiang, Z., Hao, Z., Martin, M. C., Kim, P., Stormer, H. L., & Basov, D. N. (2008b). Dirac charge dynamics in graphene by infrared spectroscopy. *Nature Physics*, 4, 532.
- Liang, H., & Upmanyu, M. (2006). Axial-strain-induced torsion in single-walled carbon nanotubes. *Physical Review Letters*, 96, 165501.
- Lin, Y.-M., Jenkins, K. A., Valdes-Garcia, A., Small, J. P., Farmer, D. B., & Avouris, P. (2009). Operation of graphene transistors at gigahertz frequencies. *Nano Letters*, 9, 422.
- Liu, X., Bauer, M., Bertagnolli, H., Roduner, E., van Slagere, J., & Philli, F. (2006). Structure and magnetization of small monodisperse platinum clusters. *Physical Review Letters*, 97, 253401.
- Liu, J. F., Wright, A. R., Zhang, C., & Ma, Z. S. (2008). Strong terahertz conductance of graphene nanoribbons under a magnetic field. *Applied Physics Letters*, 93, 041106.

- Ma, S., & Wang, G. (2006). Structures of medium size germanium clusters. *Journal of Molecular Structure (THEOCHEM)*, 767, 75.
- Mackay, A. L. (1962). A dense non-crystallographic packing of equal spheres. *Acta Crystallographica*, 15, 916.
- Maiti, A. (2008). Multiscale modeling with carbon nanotubes. *Microelectronics Journal*, 39, 208.
- Malcioğlu, O. B., & Erkoc, S. (2004). Stability of C₆₀ chains: Molecular dynamics simulations. *Journal of Molecular Graphics and Modelling*, 23, 367.
- Martin, T. P. (1996). Shells of atom. *Physics Reports*, 273, 199.
- Mermin, N. D. (1968). Crystalline order in two dimensions. *Physical Review*, 176, 250.
- Meyer, J. C., Geim, A. K., Katsnelson, M. I., Novoselov, K. S., Booth, T. J., & Roth, S. (2007). The structure of suspended graphene sheets. *Nature*, 446, 60.
- Michaelian, K., Rendon, N., & Garzon, I. L. (1999). Structure and energetics of Ni, Ag, and Au nanoclusters. *Physical Review B*, 60, 2000.
- Migas, D. B., & Borisenko, V. E. (2008). Effects of oxygen, fluorine, and hydroxyl passivation on electronic properties of {001}-oriented silicon nanowires. *Journal of Applied Physics*, 104, 024314.
- Moraga, G. G. (1993). *Cluster chemistry*. Berlin: Springer.
- Morse, M. D. (1986). Clusters of transition-metal atoms. *Chemical Reviews*, 86, 1049.
- Nakabayashi, J., Yamamoto, D., & Kurihara, S. (2009). Band-selective filter in a zigzag graphene nanoribbon. *Physical Review Letters*, 102, 066803.
- Nava, P., Seierka, M., & Ahlrichs, R. (2003). Density functional study of palladium clusters. *Physical Chemistry Chemical Physics*, 5, 3372.
- Neto, A. H. C., Guinea, F., Peres, N. M. R., Novoselov, K. S., & Geim, A. K. (2009). The electronic properties of graphene. *Reviews of Modern Physics*, 81, 109.
- Niu, T., & Li, A. (2015). From two-dimensional materials to heterostructures. *Progress in Surface Science*, 90, 21.
- Novoselov, K. S., Geim, A. K., Morozov, S. V., Jiang, D., Zhang, Y., Dubonos, S. V., Grigorieva, I. V., & Firsov, A. A. (2004). Electric field effect in atomically thin carbon films. *Science*, 306, 666.
- Noya, E. G., & Doye, J. P. K. (2006). Structural transitions in the 309-atom magic number Lennard-Jones cluster. *Journal of Physical Chemistry*, 124, 104503.
- Oymak, H., & Erkoc, S. (2004). Structural and energetic features of Al_nTi_nNi_n (n = 1–16) nanoparticles: Molecular-dynamics simulations. *Modelling and Simulation in Materials Science and Engineering*, 12, 109.
- Oymak, H., & Erkoc, S. (2012). Group 12 elements and their small clusters: Electric dipole polarizability of Zn, Cd, and Hg, Zn₂ dimer and higher Zn_n microclusters, and neutral, cationic, and anionic zinc oxide molecules (ZnO, ZnO⁺ and ZnO⁻). *International Journal of Modern Physics B*, 26, 1230003.
- Ozdamar, B., & Erkoc, S. (2012). Molecular-dynamics simulations of sawtooth-like graphene nanoribbons under strain. *Advanced Science Engineering and Medicine*, 4, 459.
- Ozdamar, B., & Erkoc, S. (2013). Structural properties of silicon nanorods under strain: Molecular dynamics simulations. *Journal of Computational and Theoretical Nanoscience*, 10, 1.
- Ozgur, U., Alivov, Y. I., Liu, C., Teke, A., Reshchikov, M. A., Dogan, S., Avrutin, V., Cho, S. J., & Morkoc, H. (2005). A comprehensive review of ZnO materials and devices. *Journal of Applied Physics*, 98, 041301.
- Pan, Z. W., Dai, Z. R., & Wang, Z. L. (2001). Nanobelts of semiconducting oxides. *Science*, 291, 1947.
- Pathak, S., Shenoy, V. B., & Baskaran, G. (2008). Possibility of high T_c superconductivity in doped graphene, arXiv:0809.0244v1 [cond-mat.supr-con] 1 Sep 2008.
- Pawluk, T., Hirata, Y., & Wang, L. (2005). Studies of iridium nanoparticles using density functional theory calculations. *Journal of Physical Chemistry B*, 109, 20817.
- Payne, F. W., Jiang, W., & Bloomfield, L. A. (2006). Magnetism and magnetic isomers in free chromium clusters. *Physical Review Letters*, 97, 193401.

- Pedersen, T. G., Flindt, C., Pedersen, J., Jauho, A.-P., Mortensen, N. A., & Pedersen, K. (2008a). Optical properties of graphene antidot lattices. *Physical Review B*, *77*, 245431.
- Pedersen, T. G., Flindt, C., Pedersen, J., Mortensen, N. A., Jauho, A.-P., & Pedersen, K. (2008b). Graphene antidot lattices: Designed defects and spin qubits. *Physical Review Letters*, *100*, 136804.
- Pekoz, R., & Erkoç, S. (2008). Quantum chemical treatment of Li/Li⁺ doped defected carbon nanocapsules. *Physica E*, *40*, 2752.
- Peres, N., Castro, A., & Guinea, F. (2006a). Conductance quantization in mesoscopic graphene. *Physical Review B*, *73*, 195411.
- Peres, N. M. R., Guinea, F., & Neto, A. H. C. (2006b). Electronic properties of two-dimensional carbon. *Annals of Physics*, *321*, 1559.
- Pisani, L., Chan, J. A., Montanari, B., & Harrison, N. M. (2007). Electronic structure and magnetic properties of graphitic ribbons. *Physical Review B*, *75*, 064418.
- Qi, J., Shi, D., Zhao, J., & Jiang, X. (2008). Stable structures and electronic properties of the oriented Bi Nanowires and Nanotubes from first-principle calculations. *Journal of Physical Chemistry C*, *112*, 10745.
- Rathi, S. J., & Ray, A. K. (2008). On the existence and stability of single walled SiGe nanotubes. *Chemical Physics Letters*, *466*, 79.
- Rey, C., Gallego, L. J., Garcia-Rodeja, J., Alonso, J. A., & Iniguez, M. P. (1993). Molecular-dynamics study of the binding energy and melting of transition-metal clusters. *Physical Review B*, *48*, 8253.
- Rodriguez-Lopez, J. L., Aguilera-Granja, F., Michaelian, K., & Vega, A. (2003). Structure and magnetism of cobalt clusters. *Physical Review B*, *67*, 174413.
- Rosales, L., Pacheco, M., Barticevic, Z., Latge, A., & Orellana, P. A. (2008). Transport properties of graphene nanoribbons with side-attached organic molecules. *Nanotechnology*, *19*, 065402.
- Rudd, R. E., & Lee, B. (2008). Mechanics of silicon nanowires: Size-dependent elasticity from first principles. *Molecular Simulation*, *34*, 1.
- Rurali, R. (2005). Electronic and structural properties of silicon carbide nanowires. *Physical Review B*, *71*, 205405.
- Rurali, R., & Lorente, N. (2005). Metallic and semimetallic silicon {100} nanowires. *Physical Review Letters*, *94*, 026805.
- Rusanov, A. I. (2014). Thermodynamics of graphene. *Surface Science Reports*, *69*, 296.
- Sadowski, T., & Ramprasad, R. (2007). Stability and electronic structure of CdSe nanorods from first principles. *Physical Review B*, *76*, 235310.
- Sasaki, K. I., Murakami, S., & Saito, R. (2006). Gauge field for edge state in graphene. *Journal of the Physical Society of Japan*, *75*, 074713.
- Sevincli, H., Topsakal, M., Durgun, E., & Ciraci, S. (2008). Electronic and magnetic properties of 3d transition-metal atom adsorbed graphene and graphene nanoribbons. *Physical Review B*, *77*, 195434.
- Shevlin, S. A., & Guo, Z. X. (2006). Transition-metal-doping-enhanced hydrogen storage in boron nitride systems. *Applied Physics Letters*, *89*, 153104.
- Shvartsburg, A. A., & Jarrold, M. F. (1999). Tin clusters adopt prolate geometries. *Physical Review A*, *60*, 1235.
- Shvartsburg, A. A., & Jarrold, M. F. (2000). Transition from covalent to metallic behavior in group-14 clusters. *Chemical Physics Letters*, *317*, 615.
- Sieck, A., Frauenheim, T., & Jackson, K. A. (2003). Shape transition of medium-sized neutral silicon clusters. *Physica Status Solidi B*, *240*, 537.
- Singh, R., & Kroll, P. (2009). Magnetism in graphene due to single-atom defects: Dependence on the concentration and packing geometry of defects. *Journal of Physics. Condensed Matter*, *21*, 196002.
- Singh, V., Daeha Joung, D., Zhai, L., Soumen Das, S., Khondaker, S. I., & Seal, S. (2011). Graphene based materials: Past, present and future. *Progress in Materials Science*, *56*, 1178.

- Son, Y. W., Cohen, M. L., & Louie, S. G. (2006). Half-metallic graphene nanoribbons. *Nature*, *444*, 347.
- Stillinger, F. H., & Weber, T. A. (1985). Computer simulation of local order in condensed phases of silicon. *Physical Review B*, *31*, 5262.
- Sun, X. H., Li, C. P., Wong, W. K., Wong, N. B., Lee, C. S., Lee, S. T., & Teo, B. K. (2002). Formation of silicon carbide nanotubes and nanowires via reaction of silicon (from disproportionation of silicon monoxide) with carbon nanotubes. *Journal of the American Chemical Society*, *124*, 14464.
- Tang, Q., & Zhou, Z. (2013). Graphene-analogous low-dimensional materials. *Progress in Materials Science*, *58*, 1244.
- Tang, C., Yan, W., Zheng, Y., Li, G., & Li, L. (2008). Dirac equation description on the electronic states and magnetic properties of a square graphene quantum dot, arXiv:0811.4312v1 [cond-mat.str-el] 26 Nov 2008.
- Tiago, M. L., Zhou, Y., Alemany, M. M. G., Saad, Y., & Chelikowski, J. R. (2006). Evolution of magnetism in iron from the atom to the bulk. *Physical Review Letters*, *97*, 147201.
- Topsakal, M., Akturk, E., Sevincli, H., & Ciraci, S. (2008). First-principles approach to monitoring the band gap and magnetic state of a graphene nanoribbon via its vacancies. *Physical Review B*, *78*, 235435.
- Ustunel, H., & Erkoç, S. (2007). Structural properties and stability of nanoclusters. *Journal of Computational and Theoretical Nanoscience*, *4*, 928.
- Ustunel, H., Roundy, D., & Arias, T. A. (2005). Modelling a suspended nanotube oscillator. *Nano Letters*, *5*, 523.
- Venkataramanan, N. J., Khazaei, M., Sahara, R., Mizuseki, H., & Kawazoe, Y. (2009). First-principles study of hydrogen storage over Ni and Rh doped BN sheets. *Chemical Physics*, *359*, 173.
- Vvedensky, D. D. (2004). Multiscale modelling of nanostructures. *Journal of Physics. Condensed Matter*, *16*, R1537.
- Wales, D. J., & Doye, J. P. K. (1997). Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *Journal of Physical Chemistry A*, *101*, 5111.
- Wang, Z. L. (2004). Functional oxide nanobelts: Materials, properties and potential applications in nanosystems and biotechnology. *Annual Review of Physical Chemistry*, *55*, 159.
- Wang, Z. L. (2009). ZnO nanowire and nanobelt platform for nanotechnology. *Materials Science and Engineering R*, *64*, 33.
- Wang, G., & Li, X. (2008). Predicting Young's modulus of nanowires from first-principles calculations on their surface and bulk materials. *Journal of Applied Physics*, *104*, 113517.
- Wang, Z. F., Li, Q., Zheng, H., Ren, H., Su, H., Shi, Q. W., & Chen, J. (2007). Tuning the electronic structure of graphene nanoribbons through chemical edge modification: A theoretical study. *Physical Review B*, *75*, 113406.
- Wang, W. L., Meng, S., & Kaxiras, E. (2008). Graphene nanoflakes with large spin. *Nano Letters*, *8*, 241.
- Weltner, W. J., & Van Zee, R. J. (1984). Transition metal molecules. *Annual Review of Physical Chemistry*, *35*, 291.
- Wright, A. R., Wang, G. X., Xu, W., Zeng, Z., & Zhang, C. (2009). The spin-orbit interaction enhanced terahertz absorption in graphene around the K point. *Microelectronics Journal*, *40*, 857.
- Wu, X., & Zeng, X. C. (2008). Sawtooth-like graphene nanoribbon. *Nano Research*, *1*, 40.
- Wu, K. L., Lai, S. K., & Lin, S. D. (2005). Finite temperature properties for zinc nanoclusters. *Molecular Simulation*, *31*, 399.
- Wu, X., Pei, Y., & Zeng, X. C. (2009). B₂C graphene, nanotubes, and nanoribbons. *Nano Letters*. doi:10.1021/nl803758s.
- Xie, Y. E., Chen, Y. P., Sun, L. Z., Zhang, K. W., & Zhong, J. X. (2009). The effect of corner form on electron transport of L-shaped graphene nanoribbons. *Physica B*, *404*, 1771.

- Yagli, H., & Erkoç, S. (2014). Structural properties of copper, silver and gold nanorods under strain: Molecular dynamics simulations. *Journal of Defence Science*, 13, 59.
- Yang, R., & Wang, Z. L. (2006). Springs, rings, and spirals of rutile-structured tin oxide nanobelts. *Journal of the American Chemical Society*, 128, 1466.
- Yang, M., Jackson, K. A., Koehler, C., Frauenheim, T., & Jellinek, J. (2006). Structure and shape variations in intermediate-size cor clusters. *Journal of Chemical Physics*, 124, 024308.
- Yao, Y. H., Gu, X., Ji, M., Gong, X. G., & Wang, D.-S. (2007). Structures and magnetic moments of Ni_n (n = 10–60) clusters. *Physics Letters A*, 360, 629.
- Zhang, X. W., & Yang, G. W. (2009). Novel band structures and transport properties from graphene nanoribbons with armchair edges. *Journal of Physical Chemistry C*, 113, 4662.
- Zhang, W., Ran, X., Zhao, H., & Wang, L. (2004). The nonmetallicity of molybdenum clusters. *Journal of Chemical Physics*, 121, 7717.
- Zhang, C., Chen, L., & Ma, Z. S. (2008a). Orientation dependence of the optical spectra in graphene at high frequencies. *Physical Review B*, 77, 241402.
- Zhang, Y. Y., Wang, C. M., & Tan, V. B. C. (2008b). Examining the effects of wall numbers on buckling behavior and mechanical properties of multiwalled carbon nanotubes via molecular dynamics simulations. *Journal of Applied Physics*, 103, 053505.
- Zhao, M., Miao, Y., Huang, Y., Kewei Xu, K., & Ma, F. (2014). The effect of H adsorption on the electronic and magnetic states in the hybrid structure of graphene and BN. *Computational Materials Science*, 93, 50.
- Zhou, J., & Dong, J. (2008). Radial breathing-like mode of wide carbon nanoribbon. *Physics Letters A*. doi:10.1016/j.physleta.2008.10.059.

Part IV
Biomolecules

Manoj K. Shukla

Roberto Cammi and Jacopo Tomasi

Contents

Introduction: Quantum Cluster Theory and Solvation	1518
The Coupled-Cluster Theory for the Polarizable Continuum Model	1523
The PCM-CC Reference State and the PCM-CC Energy Functional	1524
The PCM Hartree-Fock Reference State	1524
The Coupled-Cluster PCM Free-Energy Functional	1526
The PTDE and the PTE Coupled-Cluster Schemes	1526
The Coupled-Cluster PTDE Scheme	1527
The Coupled-Cluster PTE Scheme	1528
The Explicit PCM-CCSD Equations	1529
PCM-CCSD Analytical Gradients	1531
PCM-CCSD Analytical Gradients with Relaxed MO	1533
The Equation-of-Motion Coupled-Cluster Theory for the Polarizable Continuum Model ..	1534
The Excited-State PCM-CC Wave Functions	1535
PCM-EOM-CCSD Analytical Gradients	1538
The PCM-EOM-CCSD Equations for the Z Amplitudes	1538
The Analytical Gradients of the PCM-EOM-CCSD Free-Energy Functional ΔG_K^{EOM} ..	1539
The Response Functions Theory for Molecular Solutes	1541
The Variational Time-Dependent Theory for the Polarizable Continuum Model	1541
The Response Functions of Molecular Solutes	1544
The Coupled-Cluster Response Functions	1545
Conclusions	1550
Appendix A: The Solute-Solvent PCM Operator	1551
Bibliography	1552

R. Cammi (✉)

Dipartimento di Chimica, Università di Parma, Parma, Italy

e-mail: roberto.cammi@unipr.it

J. Tomasi

Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Pisa, Italy

e-mail: tomasi@dcc.unipi.it

Abstract

Recent extensions of the coupled-cluster (CC) theory for molecular solute described within the polarizable continuum model (PCM) were summarized. The advances covered in this review regard (i) the analytical gradients for the PCM-CC theory at the single and double excitation level, (ii) the analytical gradients for the PCM-EOM-CC theory at the single and double excitation level for the descriptions of the excited-state properties of molecular solute, and (iii) the coupled-cluster theory for the linear and quadratic molecular response functions of molecular solutes. These computational advances can be profitably used to study molecular processes in condensed phase, where both the accuracy of the QM descriptions and the influence of the environment play a critical role.

Introduction: Quantum Cluster Theory and Solvation

In this paper, we present some aspects of our recent work on the extension of computing methods connected to the polarizable continuum model (PCM). Our persevering attention to PCM can be in some sense immediately justified, having us proposed this method years ago and having continued to develop it in the years. In another sense, a justification is required and may be expressed under the form of an answer to a question: why to pass to implement PCM with sophisticated procedures as those given by the coupled-cluster (CC) theory

Coupled-cluster is the top level at which quantum mechanical (QM) calculations on molecules can be nowadays performed (Bartlett and Musiał 2007); PCM is characterized by a drastic simplification of the material system by replacing the degrees of freedom of the solvent molecules with a two body integral operator. Apparently the two methods are operating at very different scales of accuracy, but actually they can be profitably coupled. To show it, we pass to examine some features of PCM. This rapid examination of some technical aspects of the methods used in PCM will be also useful as an introductory section to our exposition.

The methodology on which PCM is based is that of an effective Hamiltonian expressed in terms of interaction integral operators depending on the solution of a nonlinear Schrödinger equation. The method is powerful and versatile, allowing applications of very different nature. A good number of such applications have been elaborated in the years by us and by other groups, and an undetermined, but large, number of new applications looks promising. Many among them are going beyond the field of dilute solutions; others address the description of solvent effects under special conditions, a few regard higher levels of the QM theory.

We are here considering an application of this last type. The features of PCM we have to examine are so limited to the developments of the QM *ab initio* version of the procedures (other versions regard semiempirical and semiclassical descriptions).

PCM was born as a method to describe electrostatic solvation effects on the electronic distribution of a molecule (Miertuš et al. 1981). The method was conceived for *ab initio* QM calculations, using a technique to solve the electrostatic problem never used before in molecular sciences, a technique based on the use of

the concept of apparent charges. The integral interaction operator is thus reduced to the Coulomb interaction with a charge distribution spread on the surface of the solute and represented by a set of point charges.

Solute-solvent interactions are not limited to the electrostatic one. The non-electrostatic terms have been introduced in PCM with the aid of a phenomenological (and effective) Hamiltonian (Bonaccorsi et al. 1982). Each term of the interaction has a different physical meaning but a similar formal expression, given by an integral operator with kernel $Q_X(r, r')$, where X stays for one of the contribution to the interaction energy: electrostatic, repulsion, dispersion, cavity formation. The effective Hamiltonian mixes quantum and classical components. In particular, the medium which appears in the kernel of the $Q_X(r, r')$ operators has to be precisely defined, in agreement with the physical conditions established for the experiments the computational model has to describe.

The formal elaboration of the physical components of the model has to be followed by the elaboration of the corresponding mathematical procedures. In this step, several “concepts” of quantum and classical origin have to be introduced: examples are the molecular surface and the notion of charging process.

We shall use here the concept of charging process as a guide in our exposition. Charging processes are of general use in physics, but rarely mentioned in molecular quantum mechanics; more explicit is its use in statistical thermodynamics to which our model is connected.

The basic task in molecular QM, that of computing energy and charge distribution of an isolated molecule at a clamped nuclear geometry, corresponds to a charging process. The analogous process in solution with an effective continuum Hamiltonian presents however important differences. The medium defining the liquid portion of the material systems is changed (polarized) by charging processes. A portion of the energy gained in a reversible charging process is so lost. The remaining portion of the energy is the free energy of the system, measured from a reference state composed by the unperturbed medium and by the appropriate assembly of noninteracting nuclei and electrons giving origin to the solute molecule. This process has a classical counterpart and can be derived on QM grounds without invoking physical arguments (a derivation is based on the variational theorem with an effective Hamiltonian linearly dependent on the solution) (Tomasi and Persico 1994).

The picture of the charging process can be used for the phenomenological partition we have introduced a few sentences above. This is a practiced used in statistical thermodynamics.

The contributions can be described in an ordered sequence of charging processes; at the end of each process, the medium is modified (polarized), and the so polarized medium is the starting point for the following charging process. The selected order of charging influences the result, because couplings in this procedure are neglected and they change according the ordering, but with an appropriate selection of the sequence, the results of the solvation energy are well reproduced with a computational cost decidedly lower than that of the full calculation (Tomasi et al. 2003).

The option of decoupling the non-electrostatic part from the main calculation has also another practical advantage that we shall exploit in the following of this paper, namely, of provisionally discard the non-electrostatic terms during the elaboration of a new electrostatic procedure. The non-electrostatic contributions to the energy are numerically large, but feebly coupled with the electrostatic one. We stress, however, that decoupling and simplified descriptions of single components are not strictly necessary: *ab initio* procedures computing everything in a unique calculation exists (Curutchet et al. 2006).

We have so far examined some points of the original PCM procedure. Other elements of interest for the present paper, drawn from the numerous procedures implemented and tested later (Tomasi et al. 2005), will be here considered. These points regard the improvements in the description of the solute's electronic wave function and the description of excited states, including nonequilibrium solvent effects.

Improvements in the quantum description of a solute mostly regard the electron correlation effects. There are, in principle, no restrictions to introduce in the formal framework of our effective Hamiltonian any version of QM methods describing electron correlation. Calculations including a description of correlation at a low level are nowadays routinely performed, but for calculations at high levels, the convenience of keeping low computational cost becomes a necessity. There are features in the effective Hamiltonians for continuum solvation promising substantial reductions in some specific versions of post Hartree-Fock calculations. In the present paper, we shall make use of a decoupling of the QM charging process elaborated years ago to better study the behavior of the many-body perturbation scheme in continuum solution and later adopted to simplify calculations in the series of MPn calculations (Olivares del Valle and Tomasi 1991). The perturbation series may be expanded on the electronic configurations of the molecular solute's electron distribution computed in *vacuo* and then used for the calculation of solvent reaction potential and solvation-free energy; this approximation is called PTD, being a version of the perturbation theory based on the electronic density. Another scheme is based on the expansion of the molecular solute's electron distribution computed with PCM at the Hartree-Fock level and used to compute solvent reaction potential and energy; this approximation is called PTE. The comparison of the results obtained with the two schemes gave insights on the relative importance of the decoupled electron correlation and solvation potential effects on the solute. A more complete procedure couples the two effects into a single iterative computational scheme, with the correlated electronic density used to make the solvent field self-consistent; the procedure is called PTED. We shall examine in a following section the PTE and PTDE version for the coupled-cluster method.

In this paper, we shall also consider the first basic aspects of the dynamical, i.e., nonequilibrium, behavior of liquid media. The subject embraces a very large variety of phenomena, requiring different formulations of the continuum models. We cite here an aspect, relatively simple, related to a phenomenon occurring in a span of time relatively short with respect to the characteristic relaxation times (CRT)

of the solvent: the vertical electronic transitions in solutes (Kim and Hynes 1990; Basilevsky and Chudinov 1990; Marcus 1992).

Electronic transitions are very fast, and the solvent molecules have no time to rearrange themselves during this phenomenon, saving for the electronic component of the solvent polarization which has a CRT of the same order of magnitude of the vertical electronic transition. This effect means that the energy of the vertical transition (a free energy, we remark) has a component due to the solvent interaction limited to changes in the fast electronic polarization. The effects of the electronic transition continue in the time, but these do not regard the vertical process and have to be described with a different formulation of the continuum method.

There is to remark, for the sake of completeness, that there are two theoretical models, respectively, named Marcus and Pekar partitions, which describe this simple process, differing in the intermediate stages of the elaboration of the problem, but arriving at the same result when the correct expression for the nonequilibrium free-energy functional is used. The PCM versions of the two models have been presented in papers Aguilar et al. (1993) and Cammi and Tomasi (1995b). For a comment, see also Aguilar (2001). A detailed comparison of the formulas of two models has been reported in Sect. 5.1 of Tomasi et al. (2005). The Pekar partition will be used in the following sections of this paper; the Marcus partition is in this case less convenient.

Quantum models in condensed systems often present alternative computational routes merging into a unique form when the effect of the medium is neglected. In coding the formulas for problems, a great attention must be paid to the correct inclusion of the corrections to satisfy the free-energy requirements.

This last remark also holds for the topic we shall consider now: the calculation of the electronic excitation energies. Two approaches are generally used to get the excitation energies: a version based on the CI approach or a version based on the response theory. Both approaches have in condensed media problems not present in the corresponding calculations in vacuo.

The roots of a CI calculation give in vacuo the energies of the various states; in solution these values have to be corrected for the effect of the solvent electronic polarization. Also the corresponding description of the so defined electronic states has to be modified (with an iterative procedure), because mutual polarization effects occur at the same time. This operation of polarization has to be separately performed for each electronic state, being the effective Hamiltonian of each excited state different from the others. The method is said “state specific” (SS), and it constitutes the main way for the study of excited-state potential surfaces.

The second approach is based on a Hamiltonian with explicit time dependence, provided with an opportune form of the of the time-dependent variation principle (the Frenkel principle is generally used Cammi et al. 1996). From this starting point, linear and nonlinear response functions are derived.

The linear response functions (LR) nowadays are amply used both in gas phase and in solution for the characterization of excited states and of molecular properties. The complete equivalence between SS and LR results for the excitation energies was

universally accepted, but recently it was shown (Kongsted et al. 2002; Cammi et al. 2005; Corni et al. 2005) that this equivalence is valid only in vacuo and that LR results can be seriously in error in solution. A computational strategy to reduce this error was proposed (Caricato et al. 2006; Improta et al. 2006), efficient enough to allow the exploitation of the less expensive LR scheme.

It is convenient to mention here another aspect of the continuum solvation theory that will be used in the following pages: the calculation of analytical derivatives.

It is essential for computational methods addressing chemical problems the availability of analytic expressions for derivatives. The variety of derivatives appearing in molecular calculations is quite large, but we limit ourselves to consider derivatives of the energy with respect to nuclear coordinates.

The derivatives in solution have important additional features with respect to the derivatives in vacuo. In solution it is compelling to compute derivatives taking into account partial derivation with respect to cavity parameters. An accurate description of the cavity shape is compulsory: examples of bad results due to inaccurate descriptions of the cavity are present in every field of application – from the determination of the molecular structure to the equilibrium among conformers to spectroscopic properties. Spherical cavities have to be used only as a first-order approximation; more subtle are the errors with cavities badly defined or treated in a simplified way during the calculations. The problems of accurate handling of the derivatives of the cavity parameters become more complex and computing expensive with the increase of the order of the derivative and of the computational level. The problem of the numerical stability of the second derivatives obliged us in the past to change the definition of the cavity elements. The analytical expressions of the first and second derivatives with respect to nuclear coordinates have been elaborated for the present version, but we shall use in the second part of this paper the first derivatives only.

The methodological and computational bases of PCM have been elaborated more extensively than shown in this exposition, but what said would be sufficient to show the care necessary to develop a reliable version of CC for PCM. Nothing has been said about the numerical accuracy of PCM results. The definition of the accuracy for a method allowing numerous different applications is a problem; it is almost compulsory to separately consider different classes of phenomena. The largest number of applications of solvent methods regards the solvation energy. Actually there is no need for this property of an accurate description of the electronic distribution, and in fact many methods are based on semiclassical approximations. There are however difficult cases in which there is the need of an accurate description of geometry and electronic charge distribution: in these cases, *ab initio* methods of sufficient accuracy are necessary. There are several methods satisfying this criterion, and all perform well. A recent comparison (Cramer and Truhlar 2008; Klamt et al. 2009) gives some indications, even if the really difficult cases were not much abundant in these comparisons. Another point to remark is that the best fitting with experimental data was obtained using “calibrated parameters.” We have not talked in the preceding pages of empirical calibration of parameters over experimental data, because we have not used this procedure in the elaboration of

PCM methods (and we are not using it in this paper). Calibration of parameters is necessary for explicitly declared semiempirical methods, but it is risky for methods aspiring to generality because ad hoc parameterizations are limited to one (or few) property and performed over a single specific definition of the computational parameters (QM method, basis set, etc.). Anyway, versions of PCM parameterized for solvation energies exist, and they behave very well.

More interesting is the consideration of accuracy for more complex problems of spectroscopic type, which actually constitute the direct motivation of this work. PCM is among the best methods in this field, with features not present on other methods, but the need of having description of the electronic structure of the chromophore more accurate than those currently used in these studies is in our opinion quite manifest. There has been a preceding attempt, several years ago, of using a CC wave function modified by continuum solvent effects, just for the calculation of molecular properties. The proposal was made by Mikkelsen and coworkers (Christiansen and Mikkelsen 1999) using a spherical cavity with a multipolar expansion of the solvent reaction potential. This attempt was bold and interesting, but the continuum model was forcibly simple, with the perspective of giving results, when applied to more complex chromophores, in which a gain in the description of the solute is accompanied by a poorer description of the solvent effect. Time was not yet ripe for efficient applications of the approach to systems of real spectroscopic interest. This is the answer to the question put at the beginning of this introduction. There is the need for some specific scopes of molecular calculations more accurate than those currently available, and the more convenient way to arrive at this level of accuracy apparently is given by an accurate (in the formulation and in the execution) *ab initio* continuum code coupled to an efficient procedure of the coupled-cluster family. The first principles of molecular dynamics at a high level of quantum description are not feasible. The use of computer simulations at a lower level accompanied by a final calculation at the CC level is a risky procedure, requiring extensive validation.

In this introduction, we have not examined the details of the coupled-cluster procedures because we have used in the implementation standard aspects of it, on which particularly comments were not necessary (Bartlett and Musiał 2007). Details and comments will be necessary when an extension of the present study shall be published.

The Coupled-Cluster Theory for the Polarizable Continuum Model

The coupled-cluster (CC) theory for the PCM model exploits the formalism of the coupled-cluster energy functional. As for the case of isolated molecules, this formalism allows to recast the coupled-cluster method for PCM into a pseudo-variational problem. However, the physical status of this energy functional has an important difference with respect to isolated molecules as it corresponds to the

free energy of the whole solute-solvent system (Christiansen and Mikkelsen 1999; Cammi 2009), as we have already said in the Introduction.

In this section, we review both the formal derivation of the PCM-CC equations and corresponding explicit equations at the single and double excitation level (PCM-CCSD).

The PCM-CC Reference State and the PCM-CC Energy Functional

The coupled-cluster wave function of the molecular solute is written as the exponential ansatz:

$$|CC \rangle = e^T |HF \rangle \quad (1)$$

where $|HF \rangle$ is the reference state given by the Hartree-Fock ground state of the solvated molecule and the cluster operator T is given as a sum of all possible excitation operators over the N electrons:

$$T = T_1 + T_2 + \dots + T_N$$

$$T_n = \frac{1}{(n!)^2} \sum_{aij\dots} t_{ij\dots}^{ab\dots} a_a^\dagger a_i a_b^\dagger a_j \dots \quad (2)$$

weighted by the amplitude t_i^a, t_{ij}^{ab} , etc. The excitation operators are here represented as products of second-quantization electron creation (a_i^\dagger, a_b^\dagger) and annihilation operators (a_i, a_b). As usual, indexes (i, j, k, \dots) and (a, b, c, \dots) denote, respectively, occupied and vacant spin orbitals MO, while (p, q, r, \dots) denote general spin orbitals.

The PCM Hartree-Fock Reference State

The Hartree-Fock reference state $|HF \rangle$ is obtained by the stationarity condition of PCM Hartree-Fock free-energy functional \mathcal{G}_{HF} (Cammi and Tomasi 1995a):

$$\mathcal{G}_{HF} = \langle HF | H^o + 1/2 \langle HF | \hat{Q} | HF \rangle \cdot \mathbf{V} | HF \rangle \quad (3)$$

where H^o is the Hamiltonian of the isolated solute and $\langle HF | \hat{Q} | HF \rangle \cdot \hat{\mathbf{V}}$ is the solute-solvent interaction term, being $\langle HF | \hat{Q} | HF \rangle$ the solvent polarization charges computed with the Hartree-Fock state reference state $|HF \rangle$ and $\hat{\mathbf{V}}$ the molecular electrostatic potential operator of the solute (see Appendix A); the one-half factor in front of the solute-solvent potential term is due the energy loss during the process of polarization of the dielectric representing the solvent, as already said in the introduction.

In an N -electron system with spin orbitals expanded over a set of atomic orbitals (AO), $\{\chi_\mu, \chi_\nu, \dots\}$, \mathcal{G}^{HF} may be written as

$$\mathcal{G}^{HF} = \sum_{\mu\nu} P_{\mu\nu}^{HF} (h_{\mu\nu} + j_{\mu\nu}) + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} P_{\mu\nu}^{HF} P_{\lambda\sigma}^{HF} [(\mu\lambda||\nu\sigma) + \mathcal{B}_{\mu\nu,\lambda\sigma}] + \tilde{V}_{NN} \quad (4)$$

where $h_{\mu\nu}$ are the matrix elements, in the AO basis, of the one-electron core operator, $(\mu\lambda||\nu\sigma)$ are the antisymmetrized combination of regular two-electron repulsion integrals (ERIs), and $P_{\mu\nu}^{HF}$ indicates the elements of the Hartree-Fock density matrix. The one-electron matrix elements $j_{\mu\nu}$ and the pseudo two-electron integrals $\mathcal{B}_{\mu\nu,\lambda\sigma}$ describe the solute-solvent interactions within the PCM-Fock operator. The one-particle AO integrals $j_{\mu\nu}$ and the pseudo two-electrons integrals $\mathcal{B}_{\mu\nu,\lambda\sigma}$ represent, respectively, the interactions with the nuclear and with the electronic components of the ASC charges. The solvent integrals $j_{\mu\nu}$, $\mathcal{B}_{\mu\nu,\lambda\sigma}$ may be expressed in the following form

$$j_{\mu\nu} = \mathbf{v}_{\mu\nu} \cdot \mathbf{q}_{Nuc} \quad (5)$$

$$\mathcal{B}_{\mu\nu,\lambda\sigma} = \mathbf{v}_{\mu\nu} \cdot \mathbf{q}_{\lambda\sigma} \quad (6)$$

where $\mathbf{v}_{\mu\nu}$ is a vector collecting the AO integrals of the electrostatic potential operator evaluated at the positions \mathbf{s}_k of the ASC charges, \mathbf{q}_{Nuc} is a vector collecting the ASC charges produced by the nuclear charge distribution, and $\mathbf{q}_{\lambda\sigma}$ is a vector collecting the apparent charges produced by the elementary charge distribution $\chi_\lambda^*(\mathbf{r})\chi_\sigma(\mathbf{r})$. The apparent surface charges may be obtained in terms of the electrostatic potential of the solute charge distributions by using the integral equation formalism of the PCM (Cancès et al. 1997).

The last term of Eq. (4), \tilde{V}_{NN} represents the interaction between the nuclei and the nuclear component of the apparent surface charges, $\tilde{V}_{NN} = 1/2 \mathbf{v}_{Nuc} \cdot \mathbf{q}_{Nuc}$.

Requiring that \mathcal{G}_{HF} be stationary with respect to the variation of MO expansion coefficients, we obtain the PCM-HF equations:

$$\sum_{\nu} (f_{\mu\nu}^{PCM} - \epsilon_p S_{\mu\nu}) c_{\nu p} = 0 \quad (7)$$

where $S_{\mu\nu}$ and $f_{\mu\nu}^{PCM}$ are, respectively, the matrix elements of the overlap matrix and of the PCM-Fock matrix, in the AO basis, and ϵ_p and $c_{\nu p}$ are, respectively, the orbital energy and the expansion coefficients of the p MO. The PCM-Fock matrix elements are given by:

$$f_{\mu\nu}^{PCM} = (h_{\mu\nu} + j_{\mu\nu}) + G_{\mu\nu}(\mathbf{P}^{HF}) + X_{\mu\nu}(\mathbf{P}^{HF}) \quad (8)$$

where $G_{\mu\nu}(\mathbf{P}^{HF})$ are the matrix elements of the effective Coulomb-exchange two-electron operator, while $X_{\mu\nu}(\mathbf{P}^{HF})$ are the matrix elements of the solvent operator

representing an effective Coulomb two-electron mediated by the polarization of the solvent.

The Coupled-Cluster PCM Free-Energy Functional

To define the PCM-CC free-energy functional, we introduce the bra coupled-cluster wave function:

$$\langle \tilde{C}C | = \langle HF | (1 + \Lambda) e^{-T} \quad (9)$$

where Λ is a de-excitation operator,

$$\Lambda = \Lambda_1 + \Lambda_2 + \dots \Lambda_n = \frac{1}{(n!)^2} \sum_{ijkabc\dots} \lambda_{abc\dots}^{ijk\dots} a_i^\dagger a_a^\dagger a_j^\dagger a_b^\dagger a_k^\dagger a_c \dots \quad (10)$$

$\lambda_a^i, \lambda_{ab}^{ij}, \dots$ are the de-excitation amplitudes.

In terms of the bra and ket wave functions, the coupled-cluster free-energy functional (Christiansen and Mikkelsen 1999; Cammi 2009) may be written as

$$G_{CC}(\Lambda, T) = \langle HF | (1 + \Lambda) e^{-T} H_0 e^T | HF \rangle + \frac{1}{2} \bar{\mathbf{Q}}(T, \Lambda) \cdot \bar{\mathbf{V}}(T, \Lambda) \quad (11)$$

where $\langle HF | (1 + \Lambda) e^{-T} H_0 e^T | HF \rangle$ is the energy functional for the isolated molecules and $\bar{\mathbf{V}}$ and $\bar{\mathbf{Q}}$ are, respectively, the coupled-cluster expectation value of the PCM operators $\hat{\mathbf{Q}}$ and $\hat{\mathbf{V}}$ (Bartlett and Musiał 2007):

$$\bar{\mathbf{Q}}(T, \Lambda) = \langle HF | (1 + \Lambda) e^{-T} \mathbf{Q} e^T | HF \rangle \quad (12)$$

$$\bar{\mathbf{V}}(T, \Lambda) = \langle HF | (1 + \Lambda) e^{-T} \mathbf{V} e^T | HF \rangle \quad (13)$$

We note that arguments of the expectation values (12) and (13) denote an explicit functional dependence on the T and Λ amplitudes. These amplitudes, which define completely the coupled-cluster wave function of the molecular solute, can be determined imposing the stationary conditions on the energy functional $G_{CC}(\Lambda, T)$.

The PTDE and the PTE Coupled-Cluster Schemes

To develop a coupled-cluster theory within the PTE and PTDE schemes described in the Introduction, the free-energy functional $G_{CC}(\Lambda, T)$ of Eq. (11) is partitioned into a Hartree-Fock component $\mathcal{G}_{\mathcal{HF}}$ and into a coupled-cluster component $\Delta G_{CC}(\Lambda, T)$:

$$G_{CC}(\Lambda, T) = \Delta G_{CC}(\Lambda, T) + \mathcal{G}_{\mathcal{HF}} \quad (14)$$

where $\mathcal{G}_{\mathcal{HF}}$ is the HF contribution, defined in Eq. (3), and $\Delta G_{CC}(\Lambda T)$ denotes the correlative CC contribution:

$$\Delta G_{CC}(\Lambda, T) = \langle HF | (1 + \Lambda) e^{-T} \left[H(0)_N + \frac{1}{2} \bar{\mathbf{Q}}_N(\Lambda, T) \cdot \mathbf{V}_N \right] e^T | HF \rangle \quad (15)$$

The partition of Eq. (14) is based on the so-called normal ordered operators $H(0)_N$, \mathbf{V}_N and \mathbf{Q}_N (Cammi 2009). Specifically, $\bar{\mathbf{Q}}_N(\Lambda, T) \cdot \mathbf{V}_N$ is the component of the solvent reaction potential due to the correlation CC electronic density, and $H(0)_N$ is the normal ordered form of Hamiltonian of the solute in presence of the frozen Hartree-Fock reaction field $\bar{\mathbf{Q}}_{HF}$.

The Coupled-Cluster PTDE Scheme

Within the PTDE scheme, the coupled-cluster electronic density is used to make the solvent reaction field self-consistent. This corresponds to make the free-energy functional $\Delta G_{CC}(\Lambda, T)$ stationary with respect to the Λ and the T amplitudes.

The stationarity with respect to the Λ amplitudes gives the equation for the self-consistent T amplitudes:

$$\frac{\partial \Delta G_{CC}}{\partial \lambda_{ab\dots}^{ij\dots}} = \langle HF | \{i^\dagger a j^\dagger b \dots\} e^{-T} H(\Lambda, T)_N e^T | HF \rangle = 0 \quad (16)$$

In turn, the stationarity with respect to the T amplitudes gives the equation for the self-consistent Λ parameters:

$$\frac{\partial \Delta G_{CC}}{\partial t_{ij\dots}^{ab\dots}} = \langle HF | (1 + \Lambda) e^{-T} [H(\Lambda, T)_N, \{a^\dagger i a^\dagger j \dots\}] e^T | HF \rangle = 0 \quad (17)$$

In Eqs. (16) and (17), $H(\Lambda, T)_N$ is the Hamiltonian of the molecular solute in the PTDE scheme:

$$H(\Lambda, T)_N = H(0)_N + \bar{\mathbf{Q}}_N(\Lambda, T) \cdot \mathbf{V}_N \quad (18)$$

where $H(0)_N$ is the normal ordered form of Hamiltonian of the solute in presence of the frozen HF reaction field and $\bar{\mathbf{Q}}_N(\Lambda, T) \cdot \hat{\mathbf{V}}_N$ is the coupled-cluster component of the solvent reaction potential, as described above.

The equation for the T amplitudes has a clear physical meaning: it corresponds to projection in the manifold spanned by all the orthonormal excitations to $|HF\rangle$ of the coupled-cluster Schrodinger equation for the molecular solute:

$$[H(0)_N + \bar{\mathbf{Q}}_N(\Lambda, T) \cdot \mathbf{V}_N] e^T | HF \rangle = \Delta E_{CC} e^T | HF \rangle \quad (19)$$

where ΔE_{CC} is given by

$$\Delta E_{CC} = \langle HF | (1 + \Lambda) e^{-T} H(\Lambda, T)_N e^T | HF \rangle \quad (20)$$

The coupled-cluster eigenvalue ΔE_{CC} differs from the free-energy functional $\Delta G_{CC}(\Lambda, T)$ by the work spent during the charging process of the coupled-cluster reaction field, i.e.,

$$\Delta G_{CC} = \Delta E_{CC} - \frac{1}{2} \bar{\mathbf{Q}}_N(\Lambda, T) \cdot \bar{\mathbf{V}}_N(\Lambda, T) \quad (21)$$

where $\bar{\mathbf{V}}_N(\Lambda, T)$ is the coupled-cluster expectation value of the electrostatic potential.

Being the Hamiltonian $H(\Lambda, T)_N$ dependent on both T and Λ parameters, the corresponding PTDE equations (16) and (17) are coupled, and they must be solved in an iterative and self-consistent way.

The Coupled-Cluster PTE Scheme

In the PTE coupled-cluster approximation, the electronic distribution is computed with PCM at the Hartree-Fock level. This approximation is easily obtained from the above PTDE equations by neglecting all the contributions related to the coupled-cluster solvent operator $\bar{\mathbf{Q}}_N(\Lambda, T) \cdot \mathbf{V}_N$. Specifically, the PTE Hamiltonian is the Hamiltonian $H(0)_N$, which contains the fixed reaction potential of the solute at the HF level, and the free-energy functional is given by

$$\Delta G_{CC}^{PTE} = \Delta E_{CC}^{PTE} = \langle HF | (1 + \Lambda) e^{-T} H(0)_N e^T | HF \rangle \quad (22)$$

The stationary condition of ΔG_{CC}^{PTE} leads to the following T equations:

$$\frac{\partial \Delta G_{CC}^{PTE}}{\partial \lambda_{ab\dots}^{ij\dots}} = \langle HF | \{i^\dagger a j^\dagger b \dots\} e^{-T} H(0)_N e^T | HF \rangle = 0 \quad (23)$$

and of the Λ equations

$$\frac{\Delta \partial G_{CC}}{\partial t_{ij\dots}^{ab\dots}} = \langle HF | (1 + \Lambda) e^{-T} H(0)_N, \{a^\dagger i a^\dagger j \dots\} e^T | HF \rangle = 0 \quad (24)$$

Within the PTE approximation, the PCM coupled-cluster equations (16) and (17) have the same structure of the coupled-cluster equations of the isolated molecules and the T , and Λ amplitudes equations are not more coupled. As a consequence, we don't need to solve the Λ equations to compute the energy functional ΔG_{CC}^{PTE} .

In the next section, we describe how the basic PTDE and PTE equations lead to explicit expressions suitable for the implementation in quantum chemistry computational codes.

The Explicit PCM-CCSD Equations

The explicit form of the PCM-CC equations (16), (17), (18), (19), (20), (21), (22), (23) and (24) has been recently presented (Cammi 2009) for the coupled-cluster single and double (CCSD) level, which restricts the T and Λ operators to excitation operators T_1, T_2 and to the de-excitation operators Λ_1, Λ_2 , respectively.

The key entities are the similarity-transformed PTE Hamiltonian $\mathcal{H}(0) = e^{-T} H(0)_N e^T$ and the similarity-transformed molecular electrostatic potential operators $\mathcal{V} = e^{-T} \widehat{\mathbf{V}}_N e^T$. Both operators can be expressed as a terminated Baker-Campbell-Hausdorff expansions. Specifically, $\mathcal{H}(0)$ terminates at the fourfold commutator, because it has at most two-particle interactions:

$$\begin{aligned} \mathcal{H}(0) &= \widehat{H(0)}_N + [\widehat{H(0)}_N, T] + \frac{1}{2} [[\widehat{H(0)}_N, T]T] \\ &+ \frac{1}{3!} [[[\widehat{H(0)}_N, T]T]T] + \frac{1}{4!} [[[[\widehat{H(0)}_N, T], T]T]T] = (H(0)_N \exp(T))_C \end{aligned} \quad (25)$$

while \mathcal{V} terminates at the twofold commutator as it has at most one-particle interactions:

$$\mathcal{V} = \widehat{\mathbf{V}}_N + [\widehat{\mathbf{V}}_N, T] + \frac{1}{2} [[\widehat{\mathbf{V}}_N, T]T] = (\mathbf{V}_N \exp(T))_C \quad (26)$$

In Eqs. (25) and (26), the subscript c indicates that only the so-called ‘‘connected’’ terms are considered, when the operators are expressed in terms of normal ordered product of second-quantization creation and annihilation operators (Bartlett and Musiał 2007).

The similarly transformed coupled-cluster PTDE-PCM Hamiltonian can then be written in the connected form

$$\mathcal{H}(\Lambda, T) = (H(0)_N \exp(T))_C + (\mathbf{V}_N \exp(T))_C \quad (27)$$

and the explicit PCM-CC equations may be obtained by substitution into Eqs. (16), (17), (18), (19), and (20). The corresponding integrals may be obtained by an algebraic approach, which exploits the Wick theorem (Bartlett and Musiał 2007) for normal product operators with $H(T, \Lambda)_N$ written in the second quantized form

$$H(T, \Lambda)_N = \sum_{p,q} \left(f_{p,q}^{PCM} + \bar{\mathbf{Q}}_N(\Lambda, T) \mathbf{v}_{pq} \right) \{p^\dagger q\} + \frac{1}{4} \sum_{p,q,r,s} \langle pq || rs \rangle \{p^\dagger q^\dagger sr\} \quad (28)$$

The explicit equation for the correlation energy ΔE_{CC} is obtained as

$$\Delta E_{CC} = \langle HF | \hat{H}(T, \Lambda)_N \left(T_2 + T_1 + \frac{1}{2} \right) T_1^2 | HF \rangle \quad (29)$$

with the corresponding algebraic form given by

$$\Delta E_{CC} = \bar{\mathbf{Q}}_N(\Lambda, T) \cdot \mathbf{v}_{ia} t_i^a + \frac{1}{4} \langle ij || ab \rangle \tau_{ij}^{ab} \quad (30)$$

where $\tau_{ij}^{ab} = t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a$.

The explicit expression for the T_1, T_2 equations are given by

$$\begin{aligned} 0 = & \langle_i^a | \hat{H}(T, \Lambda)_N \left(1 + T_2 + T_1 + T_1 T_2 + \frac{1}{2} T_1^2 + \frac{1}{3!} T_1^3 \right) | HF \rangle_c \\ = & \langle_i^a | f_N^{PCM} \left(1 + T_1 + T_2 + \frac{1}{2!} T_1^2 \right) \\ & + \bar{\mathbf{Q}}_N(\Lambda, T) \cdot \mathbf{V}_N \left(1 + T_1 + T_2 + \frac{1}{2!} T_1^2 \right) \\ & W_N \left(T_1 + T_2 + \frac{1}{2!} T_1^2 + T_1 T_2 + \frac{1}{3!} T_1^3 \right) | HF \rangle \end{aligned} \quad (31)$$

$$\begin{aligned} 0 = & \langle_{ij}^{ab} | \hat{H}(T, \Lambda)_N \left(1 + T_2 + T_1 + T_1 T_2 + \frac{1}{2} T_1^2 + \frac{1}{2!} T_2^2 \right. \\ & \left. + \frac{1}{2} T_1^2 T_2 + \frac{1}{3!} T_1^3 + \frac{1}{4!} T_1^4 \right) | HF \rangle_c \\ = & \langle_{ij}^{ab} | f_N^{PCM} (T_2 + T_1 T_2) + \bar{\mathbf{Q}}_N(\Lambda, T) \cdot \mathbf{V}_N (T_2 + T_1 T_2) \\ & + W_N \left(T_2 + \frac{1}{2!} T_1^2 + T_1 T_2 + \frac{1}{3!} T_1^3 + \frac{1}{2!} T_2^2 + \frac{1}{2!} T_1^2 T_2 + \frac{1}{4!} T_1^4 \right) | HF \rangle_c \end{aligned} \quad (32)$$

where $\langle_{ij}^{ab\dots} | = \langle HF | \{i^\dagger a j^\dagger b \dots\}$.

The explicit Λ_1, Λ_2 equations are obtained as:

$$0 = \langle HF | \hat{H}(T, \Lambda)_N e^T |_i^a \rangle_c + \langle HF | (\Lambda_1 + \Lambda_2) (\hat{H}(T, \Lambda)_N e^T)_c |_i^a \rangle_c \quad (33)$$

and

$$\begin{aligned} 0 = & \langle HF | \hat{H}(T, \Lambda)_N e^T |_{ij}^{ab} \rangle_c + \langle HF | (\Lambda_1 + \Lambda_2) (\hat{H}(T, \Lambda)_N e^T)_c |_{ij}^{ab} \rangle_c \\ & + \sum_{c=a,b;k=i,j} \langle HF | \hat{H}(T, \Lambda)_N e^T |_k^c \rangle_c \langle_k^c | \Lambda_1 + \Lambda_2 |_{ij}^{ab} \rangle \end{aligned} \quad (34)$$

The algebraic form for the T_1, T_2 and Λ_1, Λ_2 equations is reported elsewhere (Cammi 2009).

The above PTDE PCM-CC explicit equations can be easily implemented in any computational quantum code. However, the PTDE coupled-cluster procedure is different with respect to a calculation on isolate molecules, as anticipate above. The electron correlation component of the reaction field $\bar{\mathbf{Q}}_N(\Lambda, T) \cdot \mathbf{v}_{me}$ enters, as an

effective one-body contribution, both in the T and in the Λ equations, which must be solved simultaneously by using a recursive procedure. The coupling between the T, Λ equations has obvious computational consequences, with an increase of computational cost with respect to a coupled-cluster calculation for isolated molecules. In fact, in this case, it is not necessary to solve the Λ equation to obtain the CC energy and wave function.

The explicit equations for the PCM-CC-PTE approximation can be easily obtained from the PTDE equations by neglecting all terms involving the coupled-cluster components of the solvent reaction potential. This makes the computational cost of the PTE approximation similar to a coupled-cluster calculation for isolated molecules.

PCM-CCSD Analytical Gradients

The variational properties of the PCM-CC energy functional can be exploited to obtain analytical expressions of the energy derivative. For reasons of length, in this section, the exposition will be limited to the analytical first derivatives only. The interested reader may find the analytical second derivative expressions in Cammi (2009).

The differentiation with respect to a perturbational parameter α , of the stationary PCM-CC functional ΔG_{CC} , leads to an expression which avoids the first derivative of the T, Λ amplitudes. The first derivative $\partial \Delta G_{CC} / \partial \alpha = \Delta G_{CC}^\alpha$ can be expressed in the following form:

$$\begin{aligned} \Delta G_{CC}^\alpha = & \langle HF | (1 + \Lambda) | e^{-T} [H(0)_N^\alpha e^T | HF \rangle \\ & + \frac{1}{2} \langle HF | (1 + \Lambda) | e^{-T} \mathbf{Q}_N^\alpha e^T | HF \rangle \cdot \bar{\mathbf{V}}_N(\Lambda, T) \\ & + \frac{1}{2} \bar{\mathbf{Q}}_N(\Lambda, T) \cdot \langle HF | (1 + \Lambda) | e^{-T} \mathbf{V}_N^\alpha e^T | HF \rangle \end{aligned} \quad (35)$$

where the upperscript α of the various normal ordered operators denotes the total derivative of their second-quantization form.

The differentiated Hamiltonian $H(0)_N^\alpha$ in the second quantized form is given by

$$H(0)_N^\alpha = \sum_{pq} f_{pq}^{PCM,\alpha} \{p^\dagger q\} + \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle^\alpha \{p^\dagger q^\dagger sr\} \quad (36)$$

where $f_{p,q}^{PCM,\alpha}$ and $\langle pq || rs \rangle^\alpha$ denote, respectively, the total first derivatives of the PCM-Fock matrix elements and of the usual two-electron integrals, both in the MO basis.

The derivate matrix elements $f_{p,q}^{PCM,\alpha}$ may be written in terms of the derivatives of the constituting MO integrals

$$f_{pq}^{PCM,\alpha} = h_{pq}^\alpha + j_{pq}^\alpha + \sum_j^{occ} (\langle pj || qj \rangle^\alpha + \mathcal{B}_{jj,pq}^\alpha) \quad (37)$$

where the derivatives of the solvent integrals j_{pq}^α and $\mathcal{B}_{jj,pq}^\alpha$ are given by

$$j_{pq}^\alpha = \mathbf{v}_{pq}^\alpha \cdot \mathbf{q}_{Nuc} + \mathbf{v}_{pq} \cdot \mathbf{q}_{Nuc}^\alpha \quad (38)$$

$$\mathcal{B}_{pq,rs}^\alpha = \mathbf{v}_{\mu\nu}^\alpha \cdot \mathbf{q}_{rs} + \mathbf{v}_{\mu\nu} \cdot \mathbf{q}_{rs}^\alpha \quad (39)$$

where $\mathbf{q}_{p,q}^\alpha$, $\mathbf{v}_{p,q}^\alpha$ are, respectively, the differentiated apparent charge integrals and the differentiated electrostatic potential integrals (Cammi and Tomasi 1994).

The differentiated operators \mathbf{Q}_N^α and \mathbf{V}_N^α of Eq. (35) are given by:

$$\mathbf{Q}_N^\alpha = \sum_{pq} \mathbf{q}_{pq}^\alpha \{p^\dagger q\} \quad (40)$$

$$\mathbf{V}_N^\alpha = \sum_{pq} \mathbf{v}_{pq}^\alpha \{p^\dagger q\}. \quad (41)$$

Introducing Eqs. (36), (37), (38), (39), (40), and (41), the analytical gradients ΔG_{CC}^α can be written in the following explicit form:

$$\begin{aligned} \Delta G_{CC}^\alpha = & \sum_{ab} f_{ab}^{PCM,\alpha} \gamma_{ab}^{CC-\text{resp}} + \sum_{ij} f_{ij}^{PCM,\alpha} \gamma_{ij}^{CC-\text{resp}} \quad (42) \\ & + \sum_{ai} f_{ai}^{PCM,\alpha} \gamma_{ai}^{CC-\text{resp}} + \sum_{ia} f_{ia}^{PCM,\alpha} \gamma_{ia}^{CC-\text{resp}} \\ & + \frac{1}{2} \sum_{pqrs} \mathcal{B}_{pq,rs}^\alpha \gamma_{pq}^{CC-\text{resp}} \gamma_{rs}^{CC-\text{resp}} \\ & + \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle^\alpha \Gamma_{rspq} \end{aligned}$$

where Γ_{rspq} are elements of the effective two-particle density matrix Γ (Bartlett and Musiał 2007).

Equation (42) is the most general form of ΔG_{CC}^α from which we can obtain explicit computational expressions. To proceed further, as in the case of isolated molecules, we have to consider two alternative forms of the PCM-CCSD analytical derivatives: the so-called nonrelaxed MO form, which neglects the orbital relaxation effects, and the so-called relaxed MO form, which includes these effects (Gauss 1999). Which of the two forms must be used is a controversial issue in the CC property calculations of isolated molecules. The use of unrelaxed derivatives has been advocated by Koch and Jørgensen (1990) for the calculation of electrical properties. On the other, the use of the orbital relaxation effect is mandatory in all cases where perturbation-dependent basis functions are employed. For geometrical derivatives, only the MO-relaxed derivatives provide a correct description of the

potential energy surface. In the following subsection, we present the orbital relaxed form, the gradients ΔG_{CC}^α . The interested reader may find the corresponding formulation for the unrelaxed MO gradients in Cammi (2009).

PCM-CCSD Analytical Gradients with Relaxed MO

When the MO are allowed to relax under the perturbation, the occupied-virtual block of the derivative of the PCM-Fock matrix, $f_{a,i}^{PCM,\alpha}$, vanishes by virtue of the Brillouin's theorem. Therefore, the contribution to the gradients ΔG_{CC}^α of Eq. (42) due to the derivative PCM-Fock matrix consists only of an occupied-occupied and virtual-virtual blocks:

$$\begin{aligned} \Delta G_{CC}^\alpha &= \sum_{ab} f_{ab}^{PCM,\alpha} \gamma_{ab}^{CC-\text{resp}} + \sum_{ij} f_{ij}^{PCM,\alpha} \gamma_{ij}^{CC-\text{resp}} \\ &+ \frac{1}{2} \sum_{pqrs} (\mathbf{q}_{pq}^\alpha \cdot \mathbf{v}_{rs} + \mathbf{q}_{pq} \cdot \mathbf{v}_{rs}^\alpha) \gamma_{pq}^{CC-\text{resp}} \gamma_{rs}^{CC-\text{resp}} \\ &+ \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle^\alpha \Gamma_{rspq} \end{aligned} \quad (43)$$

For the evaluation of the PCM-CC gradient, the first derivatives of the PCM-Fock matrix and of the other one- and two-electron MO integrals are needed. These derivatives lead to terms involving the skeleton derivatives of the MO integrals, i.e., integrals which involve the derivative of the corresponding AO integrals and also lead to terms involving the derivative of the MO coefficients. Then, the gradients ΔG_{CC}^α may be written as:

$$\begin{aligned} \Delta G_{CC}^\alpha &= \sum_{ab} \gamma_{ab}^{CC-\text{resp}} f_{ab}^{PCM,[\alpha]} + \sum_{ij} \gamma_{ij}^{CC-\text{resp}} f_{ij}^{PCM,[\alpha]} \\ &+ \frac{1}{2} \sum_{rs} \gamma_{rs}^{CC-\text{resp}} \left[\mathbf{q}_{rs}^{[\alpha]} \cdot \bar{\mathbf{v}}_N + \bar{\mathbf{Q}}_N \cdot \mathbf{v}_{rs}^{[\alpha]} \right] \\ &+ \sum_{pqrs} \Gamma_{pq,rs} \langle pq || rs \rangle^{[\alpha]} + \sum_{pq} I'_{pq} U_{pq}^\alpha \end{aligned} \quad (44)$$

where the upperscript $[\alpha]$ denotes skeleton derivatives of the MO integrals; U_{mi}^a are the derivatives of the MO coefficients $c_{\mu p}$, in the MO basis; and I'_{pq} are auxiliary matrix elements whose expressions are given in Cammi (2009).

The explicit evaluation of the U_{mi}^a derivatives may be avoided by using the orthonormality constraints of the MO and by using the interchange (Z -vector) method of Handy and Schaefer properly extended to the PCM framework (Cammi et al. 1999).

The resulting PCM-CCSD gradients may be written as

$$\begin{aligned} \Delta G_{CC}^{\alpha} = & \sum_{ab} \gamma_{ab}^{CC-\text{resp}} f_{ab}^{PCM,[\alpha]} + \sum_{ij} \gamma_{ij}^{CC-\text{resp}} f_{ij}^{PCM,[\alpha]} + \sum_{ij} \gamma_{ai}^{MO-\text{resp}} f_{ai}^{PCM,[\alpha]} \\ & + \frac{1}{2} \sum_{rs} \gamma_{rs}^{CC-\text{resp}} \left[\mathbf{q}_{rs}^{[\alpha]} \cdot \bar{\mathbf{v}}_N + \bar{\mathbf{Q}}_N \cdot \mathbf{v}_{rs}^{[\alpha]} \right] + \sum_{pqrs} \Gamma_{pq,rs} < pq || rs >^{[\alpha]} \\ & + \sum_{ij} I'_{ij} S_{ij}^{[\alpha]} + \sum_{ab} I_{ab} S_{ab}^{[\alpha]} - 2 \sum_{ai} I'_{ia} S_{ai}^{[\alpha]} \end{aligned} \quad (45)$$

where the matrix elements $\gamma_{ai}^{MO-\text{resp}}$ are obtained as solution of the PCM-Z vector equation and $I'_{ia} = I_{ia} - \gamma_{ai}^{MO-\text{resp}} f_{ii}^{PCM}$ and I'_{ij} are given by

$$I'_{ij} = I_{ij} - \sum_{em} \gamma_{em}^{MO-\text{resp}} (< ei || mj > + < im || je > + 2B_{em,ij})$$

Finally, the expression of the PCM-CCSD gradients can be reverted from the MO to the AO representation, for an efficient computational implementation. The corresponding AO form is

$$\begin{aligned} \Delta G_{CC}^{\alpha} = & \sum_{\mu\nu} \gamma_{\mu\nu}^{CC-MO} (h_{\mu\nu}^{\alpha} + j_{\mu\nu}^{\alpha}) + \sum_{\mu\nu} I'_{\mu\nu} S_{\mu\nu}^{\alpha} \\ & + \sum_{\mu\nu\rho\sigma} \left(\gamma_{\mu\nu}^{CC-MO} P_{\sigma\rho}^{HF} + \frac{1}{2} \gamma_{\mu\nu}^{CC-\text{resp}} \gamma_{\sigma\rho}^{CC-\text{resp}} \right) B_{\mu\nu\sigma\rho}^{\alpha} \\ & + \sum_{\mu\nu\sigma\rho} \Gamma'_{\mu\nu,\sigma\rho} < \mu\sigma || \rho\sigma >^{\alpha} \end{aligned} \quad (46)$$

where $\Gamma'_{\mu\nu,\sigma\rho} = \Gamma_{\mu\nu,\sigma\rho} + \gamma_{\mu\nu}^{CC-MO} P_{\sigma\rho}^{HF}$.

Equation (46) refers to the PCM-CCSD analytical gradients within the PTDE scheme, as it contains terms involving coupled-cluster contribution to the solvent reaction field. By neglecting these contributions, we obtain the corresponding analytical gradients within the PTE-approximated scheme.

The Equation-of-Motion Coupled-Cluster Theory for the Polarizable Continuum Model

In the previous sections, we have shown the coupled-cluster theory to the description of the ground state properties of molecular solutes. In this section, we present the coupled-cluster theory for the description of the excited states of solvated molecules.

We have already anticipated that the polarizable continuum model supports two approaches for the description of the excited states of molecular solutes, a version

based on the CI approach (state-specific SS) and a version based on the response theory, and that these approaches have in condensed media problems not present in the corresponding calculations in vacuo.

Each approach has advantages and disadvantages. The LR approach is computationally more convenient, as it gives the whole spectrum of the excited states of interest in a single calculation, but is physically biased. In fact, in the LR approach, the solute-solvent interaction contains a term related to the one-particle transition densities of the solute connecting the reference state adopted in the LR calculation, which usually corresponds to the electronic ground state, to the excited electronic state. The SS approach is computationally more expensive, as it requires a separate calculation for each of excited states of interest, but is physically unbiased. In fact, in the SS approach, the solute-solvent interaction is determined by the effective one-particle electron density of the excited states.

An excited-state CC method for continuum solvation models has been developed by Mikkelsen and co-worker (1999), based on the LR approach. In this section, we present an excited-state CC state-specific (SS) method for the PCM. As already discussed in the Introduction, we will limit ourselves to a formulation of the excited-state PCM-CC to describe vertical absorption processes (the emission processes require a slightly different formulation) and in particular of absorption processes from the electronic ground state to a generic excited state. For this processes, we can give two versions: a first version in which the solvent is in equilibrium polarization with both states (state 0 and state K) and a second version in which the polarization of the K -th state is in nonequilibrium for the reasons discussed in the introduction. The formulation that we here report refers to the equilibrium formulation. For the nonequilibrium formulation, there are two versions: the formulation corresponding to the version II (Pekar) can be directly obtained from the equilibrium case by substituting the polarization charges of the solvent that appear in Eq. (52) for the equilibrium case, with those corresponding to the nonequilibrium case (Tomasi et al. 2005) and the version I of the nonequilibrium case requiring a slightly more elaborated formulation.

The Excited-State PCM-CC Wave Functions

In the excited-state PCM-CC approximation, the excited electronic states are represented by a linear (CI-like) expansion buildup on the coupled-cluster wave function for the ground state. The corresponding expansion coefficients are then determined by solving a nonlinear eigenvalue problem. The eigenvalue equations are determined by a variational procedure involving the PCM-CC free-energy functional.

The EOM-CC ket wave function (Nakatsuji and Hirao 1978; Stanton and Bartlett 1993) for the K -th state is defined as

$$|\Psi_K\rangle = \mathcal{R}_K e^T |HF\rangle \quad (47)$$

where $e^T|HF\rangle$ is the coupled-cluster state obtained by solving the PCM-PTE equation (see Eq. (23) in section “The PTDE and the PTE Coupled-Cluster Schemes”) and \mathcal{R}_K is a quasi-particle excitation operator

$$\begin{aligned}\mathcal{R}_K &= \mathcal{R}_{K,0} + \mathcal{R}_{K,1} + \mathcal{R}_{K,2} + \dots \\ \mathcal{R}_{K,n} &= \frac{1}{n!^2} \sum_{ijkl\dots abc\dots} r_{ijk\dots}^{abc\dots}(K) a_a^\dagger a_i a_b^\dagger a_j a_c^\dagger a_k \dots\end{aligned}\quad (48)$$

The corresponding EOM-CC bra wave function is given by

$$\langle \tilde{\Psi}_K | = \langle HF | \mathcal{L}_K e^{-T} \quad (49)$$

where \mathcal{L}_K is a de-excitation operator

$$\begin{aligned}\mathcal{L}_K &= \mathcal{L}_{K,0} + \mathcal{L}_{K,1} + \mathcal{L}_{K,2} + \dots \\ \mathcal{L}_{K,n} &= \frac{1}{n!^2} \sum_{ijkl\dots abc\dots} l_{ijk\dots}^{abc\dots}(K) a_i^\dagger a_a a_j^\dagger a_b a_k^\dagger a_c \dots\end{aligned}\quad (50)$$

The set ket and bra wave functions \mathcal{L}_K and \mathcal{R}_K satisfy the property of bi-orthogonality many-body systems

$$\langle \tilde{\Psi}_K | \Psi_L \rangle = \langle \mathcal{L}_K | \mathcal{R}_L \rangle = \delta_{KL} \quad (51)$$

The variational PCM-EOM-CC free-energy functional for the state of interest, ΔG_K^{EOM} , may be written as

$$\begin{aligned}\Delta G_K^{EOM} &= \langle HF | \mathcal{L}_K e^{-T} H_N(0) e^T \mathcal{R}_K | HF \rangle / \langle \mathcal{L}_K | \mathcal{R}_K \rangle \\ &+ \frac{1}{2} \bar{\mathbf{Q}}_K^{EOM} \cdot \bar{\mathbf{V}}_K^{EOM} \\ &+ \langle HF | \mathcal{Z} e^{-T} H_N(0) \exp(T) | HF \rangle\end{aligned}\quad (52)$$

Here the first term on the right side represents the EOM-CC energy of the state on interest in presence of the fixed HF reaction potential, while the second term $\bar{\mathbf{Q}}_K^{EOM} \cdot \bar{\mathbf{V}}_K^{EOM}$ is the EOM-CC solute-solvent interaction contribution. Specifically, $\bar{\mathbf{Q}}_K^{EOM}$ and $\bar{\mathbf{V}}_K^{EOM}$ are the EOM-CC expectation value, respectively, of the polarization charges and of the electrostatic potential for the K -th state:

$$\begin{aligned}\bar{\mathbf{Q}}_K^{EOM} &= \langle HF | \mathcal{L}_K e^{-T} \hat{\mathbf{V}}_N e^T \mathcal{R}_K | HF \rangle / \langle \mathcal{L}_K | \mathcal{R}_K \rangle \\ \bar{\mathbf{V}}_K^{EOM} &= \langle HF | \mathcal{L}_K e^{-T} \hat{\mathbf{Q}}_N e^T \mathcal{R}_K | HF \rangle / \langle \mathcal{L}_K | \mathcal{R}_K \rangle\end{aligned}$$

The last term of ΔG_K^{EOM} in Eq. (52) introduces the constraint for the ground state coupled-cluster wave function and contains the de-excitation operator \mathcal{Z} given by

$$\mathcal{Z} = \mathcal{Z}_1 + \mathcal{Z}_2 + \dots \quad (53)$$

$$\mathcal{Z}_n = \frac{1}{n!^2} \sum_{ijkl\dots abc\dots} \xi_{ijk\dots}^{abc\dots} a_i^\dagger a_a a_j^\dagger a_b a_k^\dagger a_c \dots$$

The presence of the coupled-cluster constraint term is only relevant for the evaluation of analytical gradient of the PCM-EOM-CC energy functional.

The PCM-EOM-CC Eigenvalue Equations

Imposing that ΔG_{PCM}^{EOM} be stationary with respect to the \mathcal{R}_K and \mathcal{L}_K amplitudes, we obtain a right-hand and a left-hand eigenvalue equation:

$$\mathcal{H}_K^{EOM} \mathcal{R}_K |HF\rangle = \Delta E_K^{EOM} \mathcal{R}_K |HF\rangle \quad (54)$$

$$\langle HF | \mathcal{L}_K \Delta E_K^{EOM} = \langle HF | \mathcal{L}_K \mathcal{H}_K^{EOM} \quad (55)$$

where \mathcal{H}_K^{EOM} is the similarity-transformed state-specific PCM-EOM-CC Hamiltonian:

$$\mathcal{H}_K^{EOM} = e^{-T} H_N^{EOM}(K) e^T = H_N^{EOM}(K) \quad (56)$$

with

$$H_N^{EOM}(K) = H_N(0) + \bar{\mathbf{Q}}_K^{EOM} \cdot \mathbf{V}_N \quad (57)$$

where the first term of state-specific Hamiltonian describes the solute in presence of the fixed HF polarization charges, while the second term, $\bar{\mathbf{Q}}_K^{EOM} \cdot \mathbf{V}_N$, represents the interaction of the solute with the polarization charges produced by the solute in the excited-state K -th.

The left and right eigenvalues and the EOM-CC energy can be obtained from the matrix representing the non-Hermitian Hamiltonian \mathcal{H}_K^{EOM} in a suitable functional space. We limit ourselves to consider the case of the functional space associated to the coupled-cluster reference function at the CCSD level.

The EOM-CCSD and the electronic states are associated with the diagonal representation of \mathcal{H}_K^{EOM} in the subspace $|\mathbf{p}\rangle$, spanned by the $|HF\rangle$ determinant and by its single $|i^a\rangle$ and double $|ij^{ab}\rangle$ excited determinants (i.e., $|\mathbf{p}\rangle = |HF\rangle \oplus |i^a\rangle \oplus |ij^{ab}\rangle$). Equations (54) and (55) can be converted to a non-Hermitian CI-like eigenvalue problem

$$\mathbf{H}_K^{\bar{EOM}} \mathbf{R}_K = \Delta E_K^{EOM} \mathbf{R}_K \quad (58)$$

$$\mathbf{L}_L \mathbf{H}_K^{\bar{EOM}} = \mathbf{L}_K \Delta E_K^{EOM} \quad (59)$$

where \mathbf{R}_k and \mathbf{L}_k represent the vectors of coefficients for the chosen excited state and \mathbf{H}_K^{EOM} is the PCM-EOM matrix Hamiltonian. The elements of matrix \mathbf{H}_K^{EOM} can be evaluated by standard methods. Details are given elsewhere (Cammi et al. 2010; Cammi 2010).

The manifold of the eigenvectors of matrix \mathbf{H}_K^{EOM} can be determined using a non-Hermitian modification of Davidson's method (Hirao and Nakatsuji 1982). From this manifold, we can extract the \mathbf{R}_K and \mathbf{L}_K eigenvectors and the corresponding eigenvalue ΔE_K^{EOM} . The remaining eigenvectors of the manifold are not of interest for the topics treated in this chapter.

Finally, the free-energy functional value for the state of interest may be obtained in terms of the EOM-CC eigenvalue ΔE_K^{EOM} as

$$\Delta G_K^{EOM} = \Delta E_K^{EOM} - \frac{1}{2} \mathbf{v}_K^{EOM} \cdot \mathbf{Q}_K^{EOM} \quad (60)$$

We have so completed the first step of the PCM-EOM-CC procedure, leading to the complete determination of the EOM-CC eigenfunctions and free energy for the target state K . We note that in this first step we have only exploited the stationarity condition of ΔG_K^{EOM} with respect to the left and right eigenvectors. As already said, the stationary conditions on the \mathcal{Z} and T amplitude are not necessary at this stage, but they are instead necessary to obtain analytical derivatives of the ΔG_K^{EOM} functional.

PCM-EOM-CCSD Analytical Gradients

Analytical first derivative of the free-energy PCM-EOM functional ΔG^{EOM} may be easily obtained by introducing the stationarity with respect to the \mathcal{Z} and the T amplitudes. This avoids the evaluation of the first derivative of these parameters with respect to the external perturbations.

The PCM-EOM-CCSD Equations for the \mathcal{Z} Amplitudes

The stationary condition for the \mathcal{Z} amplitudes leads to the PTE coupled-cluster equations for the reference ground state, which are assumed to be satisfied (see Eq. (23)); the stationary condition for the T amplitudes gives in turn the equations for the \mathcal{Z} amplitudes.

The derivative of ΔG^{EOM} with respect to the T amplitudes may be written as

$$\begin{aligned} \frac{\partial \Delta G_K^{EOM}}{\partial t_\tau} = & \langle HF | \mathcal{L} [\mathcal{H}_K^{EOM}, \hat{t}] \mathcal{R} | HF \rangle \\ & + \langle HF | \mathcal{Z} [\mathcal{H}(0), \hat{t}] | HF \rangle = 0 \end{aligned} \quad (61)$$

where $\hat{\tau}$ is the excitation operator associated to the amplitude coefficient t_τ , \mathcal{H}_K^{EOM} is the similarity-transformed EOM Hamiltonian, and $\mathcal{H}(0)$ is the PTE similarity-transformed Hamiltonian of Eq. (25); the square brackets denote a commutator operator.

Equation (61) can be rewritten as an explicit linear system of equation for the \mathcal{Z} amplitudes:

$$\frac{\partial \Delta G_{PCM}^{EOM}}{\partial t_\tau} = \langle HF | \mathcal{Z} | \mathbf{g} \rangle \langle \mathbf{g} | \bar{\mathcal{H}}(0) | \mathbf{g} \rangle - \langle HF | \Xi | \mathbf{g} \rangle = 0 \quad (62)$$

or

$$\langle HF | \mathcal{Z} | \mathbf{g} \rangle = - \langle HF | \Xi | \mathbf{g} \rangle \langle \mathbf{g} | \bar{\bar{H}}(0) | \mathbf{g} \rangle^{-1} \quad (63)$$

where $|\mathbf{g}\rangle$ denotes the subspace of the single and double excited Slater determinants, $|\mathbf{g}\rangle = |i^a\rangle \oplus |ij^{ab}\rangle$, $\bar{\bar{H}}(0)$ is defined as $\bar{\bar{H}}(0) = \mathcal{H}(0) - \langle 0 | \bar{H}(0) | 0 \rangle$, and Ξ denotes the de-excitation operator

$$\Xi = \Xi_{11} + \Xi_2 + \dots \quad (64)$$

$$\Xi_n = \frac{1}{n!2} \sum_{ijkl\dots abc\dots} \xi_{ijk\dots}^{abc\dots} a_i^\dagger a_a a_j^\dagger a_b a_k^\dagger a_c \dots \quad (65)$$

whose amplitudes Ξ_n are given as matrix elements of the operator $\mathcal{R}\mathcal{H}_K^{EOM} | \mathbf{q} \rangle \langle \mathbf{q} | \mathcal{R}$ between $\langle HF |$ and $|\mathbf{g}\rangle = |i^a\rangle \oplus |ij^{ab}\rangle$

$$\langle HF | \Xi | \mathbf{g} \rangle = \langle HF | \mathcal{R}\mathcal{H}_K^{EOM} | \mathbf{q} \rangle \langle \mathbf{q} | \mathcal{R} | \mathbf{g} \rangle \quad (66)$$

The explicit equation of the amplitudes ξ_n contains explicit PCM contributions which are given elsewhere (Cammi et al. 2010; Cammi 2010).

The Analytical Gradients of the PCM-EOM-CCSD Free-Energy Functional ΔG_K^{EOM}

The derivatives of free-energy functional ΔG_K^{EOM} may then written as

$$\begin{aligned} \frac{\partial \Delta G_K^{EOM}}{\partial \alpha} &= \langle HF | (\mathcal{L}_K + \mathcal{Z}) e^{-T} H_N(0) e^T \mathcal{R}_K | HF \rangle \\ &+ \frac{1}{2} \langle HF | \mathcal{L}_K e^{-T} \mathbf{V}_N^\alpha e^T \mathcal{R}_K | HF \rangle \cdot \mathbf{Q}_K^{EOM} \\ &+ \frac{1}{2} \langle HF | \mathcal{L}_K e^{-T} \mathbf{Q}_N^\alpha e^T \mathcal{R}_K | HF \rangle \cdot \mathbf{V}_K^{EOM} \end{aligned} \quad (67)$$

This expression can be recasted in a form that involves contractions between effective one- and two-particle EOM-CCSD density matrices and differentiated one- and two-electron MO integrals as

$$\begin{aligned} \frac{\partial \Delta G_K^{EOM}}{\partial \alpha} &= \sum_{pq} f_{pq}^{PCM,\alpha} \gamma_{pq}^{EOM}(K) \\ &+ \frac{1}{2} \sum_{pqrs} \mathcal{B}_{pq,rs}^\alpha \gamma_{pq}^{EOM-NR} \gamma_{rs}^{EOM-NR}(K) \\ &+ \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle^\alpha \Gamma_{EOM}(K) \end{aligned} \quad (68)$$

where the Einstein summation convention is followed. The density matrices, γ^{EOM-NR} , γ^{EOM} , and Γ^{EOM} , are defined by Stanton (1993):

$$\begin{aligned} \gamma_{pq}^{EOM-NR}(K) &= \langle HF | \mathcal{L}[p^\dagger q \exp(T)]_c \mathcal{R} | HF \rangle \\ \gamma_{pq}^{EOM}(K) &= \gamma^{EOM-NR} + \langle HF | \mathcal{Z}[p^\dagger q \exp(T)]_c | HF \rangle \\ \Gamma_{pqrs}^{EOM}(K) &= \langle HF | \mathcal{L}[p^\dagger q^\dagger s r e^T]_c \mathcal{R} | HF \rangle + \langle HF | \mathcal{Z}[p^\dagger q^\dagger s r e^T]_c | HF \rangle \end{aligned}$$

where the subscript c denotes a limitation to connected diagrams. The effective density γ^{EOM-NR} is the usual reduced one-particle density (expectation value of $p^\dagger q$), while the effective density matrices γ^{EOM} and Γ^{EOM} contain also correction terms that involve the ξ amplitude to account for the response of T to the perturbation.

The PCM-EOM-CCSD analytical gradients (52) have the same form of PCM-CCSD analytical gradients (42). Then, following the same procedure described for the PCM-CCSD gradients, it is easy to show that the differentiation of the MO integrals in Eq. (52) leads to terms involving the derivatives of the MO coefficients. In turn, the perturbative MO coefficients can be avoided using exploiting the PCM Z-vector technique (Cammi et al. 1999) to obtain a perturbation-independent one-particle $\gamma_{\mu\nu}^{MO-resp}$. The corresponding expression of the PCM-EOM analytical gradients in the AO basis is given by

$$\begin{aligned} \frac{\partial \Delta G_K^{EOM}}{\partial \alpha} &= \sum_{\mu\nu} \gamma_{\mu\nu}^{EOM-MO}(K) \left(h_{\mu\nu}^\alpha + j_{\mu\nu}^\alpha \right) + \sum_{\mu\nu} I'_{\mu\nu} S_{\mu\nu}^\alpha \\ &+ \sum_{\mu\nu\rho\sigma} \left(\gamma_{\mu\nu}^{EOM-MO} P_{\sigma\rho}^{HF} + \frac{1}{2} \gamma_{\mu\nu}^{EOM-NR} \gamma_{\sigma\rho}^{EOM-NR} \right) \mathcal{B}_{\mu\nu,\sigma\rho}^\alpha \\ &+ \sum_{\mu\nu\sigma\rho} \Gamma'_{\mu\nu,\sigma\rho} \langle \mu\sigma || \rho\sigma \rangle^\alpha \end{aligned} \quad (69)$$

where the additional one-particle density matrix $\gamma_{\mu\nu}^{EOM-MO}$ is defined as $\gamma_{\mu\nu}^{EOM-MO} = \gamma_{\mu\nu}^{EOM} + \gamma_{\mu\nu}^{MO-resp}$.

The equation of the PCM-EOM-CC analytical gradients contains additional one-electron MO derivative integrals with respect to the corresponding gradients for isolated molecules (Stanton 1993). As these solvation terms can be evaluated with a small effort, the PCM-EOM-CC analytical gradients can be performed for all the molecular systems for which are feasible the EOM-CCSD gradient calculations in gas phase.

The Response Functions Theory for Molecular Solutes

The response functions theory for molecules in the PCM solvation method (Cammi 2013) is an extension of the response theory for isolated molecules (Olsen and Jørgensen 1995; Helgaker et al. 2012), including this latter as a limit case. The extension from isolated to solvated molecules involves new aspects not present in the first case. These aspects are mainly related to the coupling between the interactions of the molecular solute both with the medium (i.e., the solvent) and with an external field which is in general time-dependent. A first aspect concerns the nonlinearity in the time-dependent QM problem. More specifically, the time-dependent QM problem requires a proper extension of the time-dependent QM variation principle. A second aspect regards the time dependence in the solute-solvent responsive interaction, which must be described in a nonequilibrium solvation scheme (Marcus 1956). Finally, a third aspect regards the problem of the connection of the response functions of the molecular solutes with the corresponding macroscopic counterparts (i.e., the macroscopic susceptibilities) measured in the experiments.¹ For reasons of available space, in the following, we will focus on the first two of these aspects (nonlinearity and nonequilibrium), while for the third aspect (macroscopic susceptibilities), we refer the interested reader to Cammi et al. (2000a) and Cammi and Mennucci (2007).

The Variational Time-Dependent Theory for the Polarizable Continuum Model

Let us consider the time-dependent Schrödinger equation for a molecular solute

$$i \frac{\partial}{\partial t} |\bar{\Psi}(t)\rangle = H |\bar{\Psi}(t)\rangle \quad (70)$$

¹The problem, historically known as *local field problem*, is due to the fact that the molecular solutes are locally subjected to perturbing fields which are different to that measured by the experimenter.

where H is the effective time-dependent Hamiltonian

$$H = H^0 + \bar{\mathbf{Q}}(\Psi; t) \cdot \mathbf{V} + V'(t) \quad (71)$$

where H^0 is the Hamiltonian of the isolated molecule, $\bar{\mathbf{Q}}(\Psi; t) \cdot \mathbf{V}$ is the potential energy term representing the solute-solvent electrostatic interaction, and $V'(t)$ is a general time-dependent external perturbation (We assume that $V'(t)$ is applied adiabatically so that it vanishes at $t = -\infty$).

In Eq. (71), $\mathbf{Q}(\Psi; t)$ represents the polarization charges induced by the solute on the boundary of the cavity hosting the solute itself within the dielectric medium representing the solvent, the dot represents a vectorial inner product, and \hat{V} is a vectorial operator representing the electrostatic potential of the solute at the boundary cavity. The solvent polarization charges collected in $\mathbf{Q}(\Psi; t)$ depend parametrically on time, as they are determined by the first-order density matrix $\Psi^*(t)\Psi(t)$ of the molecular solute; the most general form of $\bar{\mathbf{Q}}(\Psi; t)$ must take into account the nonequilibrium solvation effects related to the intrinsic dynamics of solvent polarization.

In the presence of a general time-dependent perturbation $V'(t)$, the time-dependent polarization charges $\mathbf{Q}(\Psi; t)$ may be formally defined as:

$$\bar{\mathbf{Q}}(\Psi; t) = \int_{-\infty}^t \langle \Psi(t') | \mathbf{Q}(t - t') | \Psi(t') \rangle dt' \quad (72)$$

where $\mathbf{Q}(t - t')$ is an apparent charge operator nonlocal in time, describing a general nonequilibrium solvation regime. In a nonequilibrium solvation regime, the description of polarization of the medium cannot be given in terms of a single value of the dielectric permittivity of the medium (as in the case of the equilibrium solvation), but it requires in principle knowledge of the whole spectrum of the frequency-dependent dielectric permittivity $\epsilon(\omega)$ (Marcus 1956).²

A convenient approach to face both the nonlinear QM problem (70) and the nonequilibrium solvation problem (72) is given by the time-dependent quasi-free-energy formalism.

The Time-Dependent Quasi-Free-Energy and Its Variational Principle

It has been shown (Cammi and Tomasi 1996) that the time-dependent nonlinear Schrödinger (70) can be obtained from the Hamilton principle when a suitable QM Lagrangian density is defined and that this Lagrangian density implies an extension of the time-dependent Frenkel's variational principle (Frenkel 1934) to time-dependent nonlinear Hamiltonians (70).

²When the external perturbation $V(t')$ is limited to static components ($\omega = 0$) and to components in the field of the optical frequencies, the description of the nonequilibrium polarization of the medium can be given in terms of a two-valued dielectric permittivity, ϵ_0 and ϵ_∞ corresponding, respectively, to the static dielectric permittivity and to the dielectric permittivity at optical frequencies.

Let us consider the time-dependent wave function $\bar{\Psi}(t) >$ expressed in the phase-isolated form

$$|\bar{\Psi}(t) > = e^{-iF(t)}|\Psi(t) > \quad (73)$$

where $F(t)$ is a function of time and $\Psi(t)$ is the so-called regular wave function, which depends only parametrically on time and which reduces in the unperturbed limit (i.e., $t = -\infty$) to the time-independent wave function describing a stationary state of the molecular solute.

The phase $F(t)$ of Eq. (73) is determined by substitution into Eq. (70):

$$\dot{F}(t) = \langle \Psi(t) | H - i \frac{\partial}{\partial t} | \Psi(t) \rangle$$

while the regular wave function $\Psi(t)$ satisfies a generalized time-dependent Frenkel variational principles (Frenkel 1934):

$$\delta \mathcal{G}(t) + i \frac{\partial}{\partial t} \langle \Psi | \delta \Psi \rangle = 0 \quad (74)$$

where $\mathcal{G}(t)$ is the time-dependent quasi-free-energy functional

$$\mathcal{G}(t) = \langle \Psi(t) | H^o + \frac{1}{2} \bar{\mathbf{Q}}(\Psi; t) \cdot \mathbf{V} - i \frac{\partial}{\partial t} | \Psi(t) \rangle \quad (75)$$

Let us now consider the case of a periodic time-dependent perturbation:

$$V(t + T; \epsilon) = V(t; \epsilon) \quad (76)$$

with period T , frequency $\omega = \frac{2\pi}{T}$, and perturbation strength ϵ . In this case, the phase-isolated form of the wave function acquires the same periodicity T :

$$|\Psi(t + T) > = |\Psi(t) > . \quad (77)$$

and by a time integration over the characteristic period T ³, the variational condition (74) for the regular wave function (77) reduces to:

$$\delta \{ \mathcal{G}(t) \}_T = \delta \{ \langle \Psi(t) | H^o + \frac{1}{2} \bar{\mathbf{Q}}(\Psi; t) \cdot \mathbf{V} - i \frac{\partial}{\partial t} | \Psi(t) \rangle \}_T = 0 \quad (78)$$

being $\{ \mathcal{G}(t) \}_T$ time-averaged quasi-free-energy functional:

$$\{ \mathcal{G}(t) \}_T = \frac{1}{T} \int_{-T/2}^{T/2} \langle \Psi(t) | H^o + \frac{1}{2} \bar{\mathbf{Q}}(\Psi; t) \cdot \mathbf{V} - i \frac{\partial}{\partial t} | \Psi(t) \rangle dt \quad (79)$$

³The time-average over a period T for a general time-dependent function $g(t)$ is defined as $\{g(t)\}_T = \frac{1}{T} \int_{-T/2}^{T/2} g(t) dt$.

The time-averaged quasi-free-energy $\{\mathcal{G}(t)\}_T$ (79) satisfies a generalization of the time-dependent Hellmann-Feynman theorem. Specifically, if the Hamiltonian (71) contains an external perturbation $\epsilon V(\omega)$ with amplitude ϵ and periodicity $T = 2\pi/\omega$, we have:

$$\begin{aligned} \frac{d\{\mathcal{G}(t)\}_T}{d\epsilon} &= \left\{ \langle \Psi(t) | \frac{\partial H}{\partial \epsilon} | \Psi(t) \rangle \right\}_T \\ &= \left\{ \langle \Psi(t) | V(\omega) | \Psi(t) \rangle \right\}_T \end{aligned} \quad (80)$$

where $\langle \Psi(t) | V(\omega) | \Psi(t) \rangle$ corresponds to the time-dependent expectation value of observable associated with the perturbing operator $V(\omega)$ and the time-average $\{\dots\}_T$ is related to its Fourier components having frequency ω_X .

The time-dependent stationary conditions (78) and the corresponding Hellmann-Feynman theorem (80) are the basic equations for the determination of the response functions of the molecular solutes in the presence of periodic external perturbations.

The Response Functions of Molecular Solutes

The molecular response functions of molecular solutes are defined in relation to a general time-dependent perturbation $V'(t)$ with several components (L) each having multiple harmonic frequencies (K_i):

$$V'(t) = \sum_{i=1}^L \sum_{j=-K_i}^{K_i} e^{-i\omega_j t} \epsilon_{X_i}(\omega_j) X_i \quad (81)$$

where X_i is the i -th perturbation operator, having amplitude $\epsilon_{X_i}(\omega_i)$ in correspondence to the frequency ω_i .

The response functions are defined as the coefficients of a Fourier-perturbation expansion of the time-dependent expectation values of a generic observable Y of the molecular system with respect to the perturbation $V(t)$ (Olsen and Jørgensen 1995):

$$\begin{aligned} \langle \Psi(t) | Y | \Psi(t) \rangle &= \langle Y \rangle_0 + \sum_{i,j}^{L,K_i} e^{-i\omega_j t} \langle \langle Y; X_i \rangle \rangle_{\omega_j} \epsilon_{X_i}(\omega_j) + \\ &+ \frac{1}{2} \sum_{i,j,k,l}^{L,M,K_i,K_j} e^{-i(\omega_k + \omega_l)t} \langle \langle Y; X_i, X_j \rangle \rangle_{\omega_k + \omega_l} \epsilon_{X_i}(\omega_k) \epsilon_{X_j}(\omega_l) + \dots \end{aligned} \quad (82)$$

where $\langle \langle Y; X_i \rangle \rangle_{\omega_{X_i}}$ represents the linear response function describing the contribution of $Y(t)$ of first order in the perturbation X_i having frequency ω_{X_i} and $\langle \langle Y; X_i, X_j \rangle \rangle_{\omega_{X_i} + \omega_{X_j}}$ denotes the quadratic response function describing the contribution of $Y(t)$ quadratic in the perturbations X_i, X_j with frequencies $\omega_{X_i}, \omega_{X_j}$ and in a similar way are defined the higher-order response functions $\langle \langle Y; X_i, X_j, X_k, \dots \rangle \rangle_{\omega_{X_i} + \omega_{X_j} + \omega_{X_k} \dots}$ (see Table 1).

Table 1 Selected response functions $\langle\langle X; Y, Z, \dots \rangle\rangle_{\omega_Y, \omega_Z, \dots}$ of molecular solutes. The operators are μ_α , Cartesian components of the electric dipole operator, $\Theta_{\alpha, \beta}$, element of the electric quadrupolar tensor, m_α Cartesian component of the magnetic dipole operator

Response function	Property
$\langle\langle \mu_\alpha; \mu_\beta \rangle\rangle_\omega$ (Cammi et al. 1996)	Dipole polarizability $\alpha_{\alpha, \beta}(-\omega; \omega)$
$i \langle\langle m_\alpha; \mu_\beta \rangle\rangle_\omega$ (Mennucci et al. 2002; Caricato 2013b)	Optical rotation tensor $G_{\alpha, \beta}^i(\omega)$
$\langle\langle \mu_\alpha; \Theta_\beta \rangle\rangle_\omega$ (Rizzo et al. 2007)	Quadrupolar polarizability tensor $\alpha_{\alpha, \beta}(-\omega; \omega)$
$\langle\langle \mu_\alpha; \mu_\beta, \mu_\gamma \rangle\rangle_\omega$ (Cammi et al. 1996)	First polarizability $\beta_{\alpha, \beta, \gamma}(-\omega_\alpha; \omega_\beta, \omega_\gamma)$
$\langle\langle \mu_\alpha; \mu_\beta, \mu_\gamma, \mu_\delta \rangle\rangle_{\omega_\alpha, \omega_\beta, \omega_\delta}$ (Cammi et al. 1996)	Second polarizability $\beta_{\alpha, \beta, \gamma}(-\omega_\alpha; \omega_\beta, \omega_\gamma)$
$-\langle\langle \mu_\alpha; \mu_\beta, m_\gamma, m_\delta \rangle\rangle_{\omega, 0, 0}$	
$+\langle\langle \mu_\alpha; \mu_\beta, m_\gamma, m_\delta \rangle\rangle_{\omega, 0, 0}$ (Cappelli et al. 2003)	Frequency dependent mixed electric magnetic hyper-magnetizability $\eta_{\alpha, \beta, \gamma, \delta}(-\omega; \omega, 0, 0)$

The molecular response functions (82) may be determined by means of the Hellmann-Feynman theorem (80) when the time-averaged free-energy functional (79) is expanded in orders of the perturbation $V'(t)$ and the variational condition (78) is satisfied at the various order.

The response function theory for molecular solutes within PCM has been developed for linear and nonlinear response functions and at several QM levels, including the variational wave function methods (SCF, MCSCF) and the density functional theory (Cammi et al. 1996; Cammi and Mennucci 1999; Tomasi et al. 1999; Cammi et al. 2000b; Cossi and Barone 2001; Mennucci et al. 2002; Cammi et al. 2003; Cappelli et al. 2003, 2005; Frediani et al. 2005; Jansik et al. 2006; Rizzo et al. 2007; Zhao et al. 2007; Ferrighi et al. 2007; Lin et al. 2008) and more recently at the coupled-cluster level (Cammi 2012a, b; Caricato 2013b). In the following, we will present the general aspects of the theory and its specific formulation at the coupled-cluster level.

The Coupled-Cluster Response Functions

The time-dependent coupled-cluster wave function may be expressed as (Christiansen et al. 1998)

$$|CC(t)\rangle = e^{-iF(t)} e^{T(t)} |HF\rangle \quad (83)$$

where the reference state $|HF\rangle$ is the fixed, time-independent Hartree-Fock ground state of the solvated molecules⁴, the phase factor $F(t)$ is a function of time, and $T(t)$ is the time-dependent cluster excitation operator.

⁴(here, we consider the molecular orbital (MO) *unrelaxed* approach in which the reference state does not depend on the perturbation $V'(t)$ (Christiansen et al. 1998))

The coupled-cluster phase factor $F(t)$ is determined as:

$$\dot{F}(t) = \langle HF | (1 + \Lambda) e^{-T} \left(H(0)_N + \bar{\mathbf{Q}}_N(t) \cdot \mathbf{V}_N + V(t) - i \frac{\partial}{\partial t} \right) e^T | HF \rangle$$

Here $\Lambda(t)$ is the time-dependent coupled-cluster de-excitation operator, and $\bar{\mathbf{Q}}_N(t)$ is the vector collecting the time-dependent apparent charges:

$$\bar{\mathbf{Q}}_N(t) = \int_{-\infty}^t \langle HF | (1 + \Lambda(t')) e^{-T(t')} \mathbf{Q}_N(t, t') e^{T(t')} | HF \rangle dt' \quad (84)$$

where $\mathbf{Q}_N(t, t')$ is an apparent charge operator nonlocal with respect to the time (see Eq. (72) and Λ is a time-dependent de-excitation operator.

The time-dependent coupled-cluster $T(t)$ and $\Lambda(t)$ amplitudes are determined using the quasi-free-energy functional

$$\begin{aligned} \{\Delta G_{CC}(t)\}_T = & \frac{1}{T} \int_{-T/2}^{T/2} \langle HF | (1 + \Lambda(t)) e^{-T(t)} [H(0)_N \\ & + \frac{1}{2} \bar{\mathbf{Q}}_N(t) \cdot \mathbf{V}_N + V(t) - i \frac{\partial}{\partial t}] e^T(t) | HF \rangle dt \quad (85) \end{aligned}$$

(where the time integration is over a common multiple of periods, T , such that $V(t + T) = V(t)$ Christiansen et al. 1998), with the variational condition

$$\delta\{\Delta G_{CC}(t)\}_T = 0 \quad (86)$$

The variational Eq. (86) is solved using an additional perturbative approximation in which the coupled-cluster T and Λ amplitudes are expanded with respect to the components of the perturbation $V'(t)$:

$$T(t) = T^{(0)} + T^{(1)} + T^{(2)} \dots \quad (87a)$$

$$\Lambda(t) = \Lambda^{(0)} + \Lambda^{(1)} + \Lambda^{(2)} \dots \quad (87b)$$

with:

$$T^{(1)} = \sum_i \exp(-i\omega_i t) \sum_x \epsilon_X(\omega_i) T^X(\omega_i) \quad (88a)$$

$$T^{(1)} = \frac{1}{2} \sum_{i,j} \exp(-i(\omega_i + \omega_j)t) \sum_{X,Y} \epsilon_X(\omega_i) \epsilon_Y(\omega_j) T^{XY}(\omega_i, \omega_j) \dots$$

$$\Lambda^{(1)} = \Lambda^{(0)} + \sum_i \exp(-i\omega_i t) \sum_x \epsilon_X(\omega_i) \Lambda^X(\omega_i) \quad (88b)$$

$$\Lambda^{(2)} = \frac{1}{2} \sum_{i,j} \exp(-i(\omega_i + \omega_j)t) \sum_{X,Y} \epsilon_X(\omega_i) \epsilon_Y(\omega_j) \Lambda^{XY}(\omega_i, \omega_j) \dots$$

The expansions (88a), (88b) lead to an expansion of the time-averaged of the quasi-free-energy functional (85):

$$\{\Delta G(t)_{CC}\}_T = \{\Delta G_{CC}^{(0)}\}_T + \{\Delta G_{CC}^{(1)}\}_T + \{\Delta G_{CC}^{(2)}\}_T + \dots \quad (89)$$

where the n -th order term $\{\Delta G_{CC}^{(n)}\}_T$ is defined as

$$\{\Delta G_{CC}^{(n)}\}_T = \frac{1}{n!} \sum_{i,j,\dots} \sum_{x,y,z} \epsilon_X(\omega_i) \epsilon_Y(\omega_j) \dots \times G^{XY\dots}(\omega_i, \omega_j, \dots) \quad (90)$$

where the Fourier components $G^{XY\dots}(\omega_i, \omega_j, \dots)$ are defined as

$$G^{XY\dots}(\omega_i, \omega_j, \dots) = \frac{d^n \{\Delta G^{(n)}\}_T}{d\epsilon_i d\epsilon_j \dots} \quad (91)$$

with $\omega_i + \omega_j + \dots = 0$.

The perturbation expansion of Eq. (89) is then introduced into the time-average variational criterion (86), and separating by orders, we arrive to a set of variational criteria on the Fourier components $G^{XY\dots}(\omega_i, \omega_j, \dots)$

$$\delta G^{XY\dots}(\omega_i, \omega_j, \dots) = 0 \quad (92)$$

that determine the Fourier coefficients $T^{XY\dots}(\omega_i, \omega_j, \dots)$ and $\Lambda^{XY\dots}(\omega_i, \omega_j, \dots)$ for the T and Λ amplitudes (88a) and (88b).

When the variational conditions (92) are satisfied, to the Fourier components $\tilde{G}^{XY\dots}(\omega_i, \omega_j, \dots)$ correspond to the response functions of the molecular solutes:

$$\begin{aligned} \langle\langle X; Y \rangle\rangle_{\omega_i} &= \tilde{G}^{XY}(\omega_i, \omega_j) \\ \langle\langle X; Y, Z \rangle\rangle_{\omega_i} &= \tilde{G}^{XYZ}(\omega_i, \omega_j, \omega_k) \\ \vdots &= \vdots \\ \langle\langle X; Y, Z, \dots \rangle\rangle_{\omega_i} &= \tilde{G}^{XYZ\dots}(\omega_i, \omega_j, \dots) \end{aligned} \quad (93)$$

where the tilde denotes stationarity. The response functions (93) satisfy a generalized $(2n + 1)$ rule with respect to the perturbed T and Λ amplitudes.

In the following subsection, we give a detailed account of the explicit form of the coupled-cluster response equations from which the linear and quadratic response functions can be obtained.

Coupled-Cluster Response Equations

The first-order equations for the coupled-cluster Fourier amplitudes $\Lambda^X(\pm\omega)$ $T^X(\pm\omega)$ are determined from the stationarity of the second-order time-averaged quasi-free-energy $G^{XY}(-\omega, \omega)$:

$$\begin{aligned}
G^{XY}(-\omega, \omega) = & P(XY) \langle HF | (1 + \Lambda) e^{-T} | [X, T^Y(\omega)] e^T | HF \rangle \\
& + \langle HF | (1 + \Lambda) e^{-T} [[H_N, T^X(-\omega)], T^Y(\omega)] e^T | HF \rangle \\
& - \omega P(XY) \langle HF | \Lambda^X(-\omega) T^Y(\omega) | HF \rangle \\
& + P(XY) \langle HF | \Lambda^X(-\omega) e^{-T} [Y + [H_N, T^Y(\omega)] e^T | HF \rangle \\
& \frac{1}{2} P(XY) \bar{\mathbf{Q}}_N^X(-\omega) \cdot \bar{\mathbf{V}}_N^Y(\omega)
\end{aligned} \quad (94)$$

Here, $H_N = H_N(0) + \bar{\mathbf{Q}}_N^{(0)} \cdot \mathbf{V}$ is the unperturbed Hamiltonian of the molecular solute; $\bar{\mathbf{Q}}_N^X(-\omega)$ and $\bar{\mathbf{V}}_N^Y(\omega)$ are given by

$$\begin{aligned}
\bar{\mathbf{Q}}_N^X(-\omega) = & \langle HF | (1 + \Lambda) e^{-T} [\mathbf{Q}_N^{|\omega|}, T^X(-\omega)] e^T | HF \rangle \\
& + \langle HF | \Lambda^X(-\omega) e^{-T} \mathbf{Q}_N^{|\omega|} e^T | HF \rangle
\end{aligned} \quad (95a)$$

$$\begin{aligned}
\bar{\mathbf{V}}_N^Y(\omega) = & \langle HF | (1 + \Lambda) e^{-T} [\mathbf{V}_N, T^Y(\omega)] e^T | HF \rangle \\
& + \langle HF | \Lambda^Y(\omega) e^{-T} \mathbf{V}_N e^T | HF \rangle
\end{aligned} \quad (95b)$$

In Eq. (95a) the polarization charges operator $\mathbf{Q}_N^{|\omega|}$ is defined as

$$\mathbf{Q}_N^{|\omega|} = \mathbf{T}(\omega) \cdot \mathbf{V}_N \quad (96)$$

where $\mathbf{T}(\omega)$ is the frequency-dependent PCM response matrix evaluated with the dielectric permittivity $\epsilon(\omega)$ of the solvent at the frequency ω .

- The equations for the first-order $T^X(\pm\omega)$ amplitudes are obtained from the stationary condition $\delta G^{XY}(-\omega, \omega) = 0$ with respect to the first-order $\Lambda^X(\pm\omega)$ amplitudes:

$$\begin{aligned}
0 = & \langle HF | \tau_p^\dagger e^{-T} \tilde{Y}(\omega) e^T | HF \rangle \\
& + \langle HF | \tau_p^\dagger [e^{-T} H_N e^T, T^Y(\omega)] - \omega T^Y(\omega) | HF \rangle
\end{aligned} \quad (97)$$

where $\tilde{Y}(\omega)$ is an effective perturbing operator:

$$\tilde{Y}(\omega) = Y + \bar{\mathbf{Q}}_N^Y(\omega) \cdot \mathbf{V}_N \quad (98)$$

Note that this operator indicates that to the direct perturbation with the external field, represented by the operator Y , it is added an indirect source of perturbation due to the effect of the direct perturbation itself on the solute-solvent interaction $\bar{\mathbf{Q}}_N^Y(\omega) \cdot \mathbf{V}_N$.

- The equations for the first-order $\Lambda^X(\pm\omega)$ amplitudes are obtained from the stationary condition $\delta G^{XY}(-\omega, \omega) = 0$ with respect to the first-order $T^X(\pm\omega)$ amplitudes:

$$\begin{aligned}
0 = & \langle HF | (1 + \Lambda) e^{-T} [\tilde{Y}(\omega), \tau_p] e^T | HF \rangle \\
& + \langle HF | (1 + \Lambda) e^{-T} [[H_N, \tau_p], T^Y(\omega)] e^T | HF \rangle \\
& - \omega \langle HF | \Lambda^Y(\omega) \tau_p | HF \rangle + \langle HF | \Lambda^Y(\omega) e^{-T} [H_N, \tau_p] e^T | HF \rangle
\end{aligned} \quad (99)$$

Coupled-Cluster Linear and Quadratic Response Functions

- The linear response functions $\langle\langle X; Y \rangle\rangle_\omega$ can then be obtained from Eq. (93) by introducing the stationary condition (97) in to the second-order quasi-free-energy (94)

$$\begin{aligned}
\langle\langle X; Y \rangle\rangle_\omega &= \tilde{G}^{XY} \\
&= \frac{1}{2}C^{\pm\omega} \{ P(XY) \langle HF | (1 + \Lambda) e^{-T} [[X, T^Y(\omega)] e^T | HF \rangle \\
&\quad + \langle HF | (1 + \Lambda) e^{-T} [[H_N, T^X(-\omega)], T^Y(\omega)] e^T | HF \rangle \\
&\quad + \frac{1}{2} P(XY) \langle HF | (1 + \Lambda) e^{-T} [\mathbf{Q}_N^{|\omega|}, T^X(-\omega)] e^T | HF \rangle \\
&\quad \cdot \langle HF | (1 + \Lambda) e^{-T} [\mathbf{V}_N, T^Y(\omega)] e^T | HF \rangle \\
&\quad + \frac{1}{2} P(XY) \langle HF | \Lambda^X(-\omega) \mathbf{Q}_N^{|\omega|} e^T | HF \rangle \cdot \langle HF | \Lambda^Y(\omega) \mathbf{V}_N e^T | HF \rangle \}
\end{aligned} \tag{100}$$

where $C^{\pm\omega}$ is a symmetrization operator with respect to sign change of the frequencies to ensure that only the real part of the response function is retained and $P(x, y)$ is a permutation operator between x and y .

Alternatively, introducing the stationary condition (99) into the second-order quasi-free-energy (94), we can write the linear response functions as

$$\begin{aligned}
\langle\langle X; Y \rangle\rangle_\omega &= \frac{1}{2}C^{\pm\omega} \{ \langle HF | (1 + \Lambda) e^{-T} [X, T^Y(\omega)] e^T | HF \rangle \\
&\quad + \langle HF | \Lambda^Y(\omega) e^{-T} X e^T | HF \rangle \}
\end{aligned} \tag{101}$$

- The quadratic response functions $\langle\langle X; Y, Z \rangle\rangle_{\omega_Y, \omega_Z}$ can be obtained in agreement with the $(2n + 1)$ rule as

$$\begin{aligned}
\langle\langle X; Y, Z \rangle\rangle_{\omega_Y, \omega_Z} &= \tilde{G}^{XYZ} \\
&= \frac{1}{2}C^{\pm\omega} \{ P^3(XYZ) \langle HF | (1 + \Lambda) e^{-T} [[X, T^Y(\omega_Y), T^Z(\omega_Z)] e^T | HF \rangle \\
&\quad + P^6(XYZ) \langle HF | \Lambda^X(\omega_X) e^{-T} [Y, T^Z(\omega_Z)] e^T | HF \rangle \\
&\quad + \langle HF | (1 + \Lambda) e^{-T} [[H_N, T^X(\omega_X)], T^Y(\omega_Y), T^Z(\omega_Z)] e^T | HF \rangle \\
&\quad + P^3(XYZ) \langle HF | \Lambda^X(\omega_X) e^{-T} [[H_N, T^Y(\omega_Y)], T^Z(\omega_Z)] e^T | HF \rangle \\
&\quad + \frac{1}{2} P^3(XYZ) \left(\tilde{\mathbf{Q}}_N^{XY}(\omega_X + \omega_Y) \cdot \tilde{\mathbf{V}}_N^Z(\omega_Z) + \tilde{\mathbf{Q}}_N^X(\omega_X) \cdot \tilde{\mathbf{V}}_N^{YZ}(\omega_Y + \omega_Z) \right) \}
\end{aligned} \tag{102}$$

where $P^3(xyz)$ and $P^6(xyz)$ denote, respectively, the cyclic permutation operator and the full permutation operator of their arguments, while $\tilde{\mathbf{Q}}_N^{XY}(\omega_X + \omega_Y)$ ⁵ and $\tilde{\mathbf{V}}_N^{YZ}(\omega_Y + \omega_Z)$ are defined as

$$\begin{aligned} \mathbf{Q}_N^{\tilde{X}Y}(\omega_X + \omega_Y) &= \langle HF | (1 + \Lambda) e^{-T} [[\mathbf{Q}_N^{|\omega_X + \omega_Y|}, T^X(\omega_X), T^Y(\omega_Y)] e^T | HF \rangle \\ &\quad + P(XY) \langle HF | \Lambda^X(\omega_X) e^{-T} [\mathbf{Q}_N^{|\omega_Z|}, T^Y(\omega_Y)] e^T | HF \rangle \\ \mathbf{V}_N^{\tilde{Y}Z}(\omega_Y + \omega_Z) &= \langle HF | (1 + \Lambda) e^{-T} [[\mathbf{V}_N, T^X(\omega_X), T^Y(\omega_Y)] e^T | HF \rangle \\ &\quad + P(XY) \langle HF | \Lambda^X(\omega_X) e^{-T} [\mathbf{V}_N, T^Y(\omega_Y)] e^T | HF \rangle \end{aligned} \quad (103)$$

Conclusions

We have presented a short review on recent progresses toward the extension of the QM coupled-cluster level for the study of the electronic structure and properties of molecular solutes with the Polarizable Continuum Model framework (Cammi 2009). Specifically, we have presented (i) the detailed expression for the evaluation of the analytical gradients for the PCM-CC theory at the single and double excitation level, for the ground states; (ii) the expressions of the analytical gradients for the PCM-EOM-CC theory at the single and double excitation level for the descriptions of the excited-state properties of molecular solute; and (iii) the expressions of the linear and quadratic PCM-CC response functions.

Several implementations of these PCM-CC theories have been presented in the last years within the Gaussian suite of programs (Frisch et al. 2009b; Frisch et al. 2009a). Caricato et al. have presented implementations of the PCM-CC analytical gradients, of the PCM-EOM-CC gradients, and of the linear response functions PCM-CC (Caricato et al. 2010, 2011; Caricato 2011; Caricato and Scalmani 2011; Caricato 2012a, b, 2013b, a, 2014). Fukuda et al. have presented various implementations of the PCM-CC and PCM-EOM-CC methods and of the corresponding analytical derivatives within the framework of SAC/SACCI methods (Cammi et al. 2010; Fukuda et al. 2011, 2012; Fukuda and Ehara 2013; Fukuda et al. 2014; Fukuda and Ehara 2014; Fukuda et al. 2015).

We hope that these computational progresses can be profitably used to study molecular processes in condensed phase, where both the accuracy of the QM descriptions and the influence of the environment play a critical role, as in photoionization processes, electronic transitions, and charge transfer reactions.

⁵The polarization charges operator $\mathbf{Q}_N^{|\omega_X + \omega_Y|}$ is evaluated from Eq. (96) using the dielectric permittivity $\epsilon(|\omega_X + \omega_Y|)$ of the solvent at the frequency $|\omega_X + \omega_Y|$. Therefore, Eq. (103) is able to describe the nonequilibrium solvation effects in the quadratic response function describing general second-order molecular processes (Cammi et al. 1996).

Appendix A: The Solute-Solvent PCM Operator

In the Polarizable Continuum Model for salvation, the molecular solute is hosted into a cavity of a polarizable dielectric medium representing the solvent. The cavity is accurately modeled on the shape of the molecular solute (Miertuš et al. 1981), and the dielectric medium is characterized by the dielectric permittivity ϵ of the bulk solvent. The physics of the model is very simple. The solute charge distribution polarizes the dielectric medium which in turn acts back on the solute, in a process mutual polarization that continues until self-consistence is reached. The polarization of the solvent is represented by an apparent charge distribution (ASC) spread on the cavity surface. In computational practice, the ASC is discretized to a set of point charges, and the solute-solvent interaction is expressed as in terms of the interaction between these and the charge distribution of the molecular solute.

In this framework, the PCM solute-solvent interaction operator (see Eq. (3)) can be defined in terms of a molecular electrostatic operators \mathbf{V} and of a charge operator \mathbf{Q} describing the PCM solute-solvent interaction operator (Cammi et al. 2002, 2005). \mathbf{V} is a vector collecting the molecular electrostatic potential operator, evaluated at the positions of the ASC charges:

$$(\mathbf{V})_j = v_{el}(\mathbf{s}_j) + v_{nuc}(\mathbf{s}_j) \quad (104)$$

where $v_{el}(\mathbf{s}_j)$ and $v_{nuc}(\mathbf{s}_j)$ are the electrostatic potential of the electrons and of the nuclei, respectively, evaluated at the position \mathbf{s}_j of the j -th ASC point charge. In particular,

$$v_{el}(\mathbf{s}_j) = \sum_i^N \frac{-1}{|\mathbf{r}_i - \mathbf{s}_j|} \quad (105)$$

where N is the number of electrons of the molecular solute and \mathbf{r}_i is the vector position of the i -th electron. A similar expression holds for the nuclei contribution $\hat{v}_{nuc}(\mathbf{s}_j)$.

The apparent charges vector operator \mathbf{Q} can be formally defined in a different way, depending on the different version of the BEM. In the integral equation formalism (IEF) version of the PCM (Cancès et al. 1997), which is here adopted, \mathbf{Q} may be expressed as a linear transformation of the vector operator \mathbf{V} :

$$(\mathbf{Q})_j = \sum_l^{NTS} T_{jl}(\hat{\mathbf{V}})_l \quad (106)$$

where T_{jl} are elements of the BEM square matrix. The matrix is dimensioned as the ASC charges, and it depends on the geometry of surface cavity and on the dielectric permittivity of the medium. The expectation value of the $\hat{\mathbf{Q}}$ operator ($\langle \Psi | \hat{\mathbf{Q}} | \Psi \rangle$) is a vector collecting the actual ASC charges induced by the molecular solute described by the wave function Ψ .

The solute-solvent interaction operator ($\langle \Psi | \hat{\mathbf{Q}} | \Psi \rangle \cdot \hat{\mathbf{V}}$ of Eq. (2) of the text is then the inner product between the expectation values of the ASC, ($\langle \Psi | \hat{\mathbf{Q}} | \Psi \rangle$), and the molecular electrostatic operator \mathbf{V} :

$$\langle \Psi | \mathbf{Q} | \Psi \rangle \cdot \mathbf{V} = \sum_j^{NTS} \langle \Psi | (\hat{\mathbf{Q}})_j | \Psi \rangle (\hat{\mathbf{V}})_j \quad (107)$$

where $\langle \Psi | (\mathbf{Q})_j | \Psi \rangle$ denotes the expectation value of the j -th apparent surface charge.

Bibliography

- Aguilar, M. (2001). Separation of the Electric Polarization into Fast and Slow Components: A Comparison of Two Partition Schemes. *Journal of Physical Chemistry A*, *105*, 10393.
- Aguilar, M. A., Olivares del Valle, F. J., & Tomasi, J. (1993). Nonequilibrium solvation: An ab initio quantum mechanical method in the continuum cavity model approximation. *Journal of Chemical Physics*, *98*, 7375.
- Bartlett, R. J., & Musiał, M. (2007). Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics*, *79*, 291; Helgaker, T., Jørgensen, P., & Olsen, J. (2000). *Molecular electronic structure theory*. New York: Wiley; Naktsuji, H. (1997). SAC-CI method: theoretical aspects and some applications. In J. Lecsczynski (Ed.) *Computational chemistry: review of current trends* (Vol. 2, pp. 62–124). Singapore: World Scientific.
- Basilevsky, M. V., & Chudinov, G. E. (1990). Application of generalized transition state theory for calculation of the rate constant of a chemical reaction with charge transfer in a polar solvent. *Chemical Physics*, *144*, 155.
- Bonaccorsi, R., Ghio, C., & Tomasi, J. (1982). The effect of the solvent on electronic transition and other properties of molecular solutes. In R. Carbó, (Ed.), *Current aspects of quantum chemistry* (p. 407). Amsterdam: Elsevier.
- Cammi, R. (2009). Quantum cluster theory for the polarizable continuum model. I. The CCSD level with analytical first and second derivatives. *Journal of Chemical Physics*, *131*, 164104.
- Cammi, R. (2010). Coupled-Cluster Theories for the Polarizable Continuum Model. II. Analytical Gradients for Excited States of Molecular Solutes by the Equation of Motion Coupled-Cluster Method, *International Journal of Quantum Chemistry*, *110*, 3040.
- Cammi, R. (2012a). Coupled-cluster theory for the polarizable continuum model. III. A response theory for molecules in solution. *International Journal of Quantum Chemistry*, *112*, 2547.
- Cammi, R. (2012b). Recent Advances in the Coupled-Cluster Analytical Derivatives Theory for Molecules in Solution Described With the Polarizable Continuum Model (PCM). *Advances in Quantum Chemistry*, *64*, 1.
- Cammi, R. (2013). *Molecular response functions for the polarizable continuum model*. Cham: Springer.
- Cammi, R., & Mennucci, B. (1999). Linear response theory for the polarizable continuum model. *Journal of Chemical Physics*, *110*, 9877.
- Cammi, R., & Mennucci, B. (2007). Macroscopic Nonlinear Optical Properties from Cavity Models. In B. Mennucci & R. Cammi (Ed.) *Continuum solvation models in chemical physics* (p. 238). Chichester: Wiley.
- Cammi, R., & Tomasi, J. (1994). Analytical derivatives for molecular solutes. I. Hartree-Fock energy first derivatives with respect to external parameters in the polarizable continuum model. *Journal of Chemical Physics*, *100*, 7495.

- Cammi, R., & Tomasi, J. (1995a). Remarks on the use of the apparent surface charges (ASC) methods in solvation problems: Iterative versus matrix-inversion procedures and the renormalization of the apparent charges. *Journal of Computational Chemistry*, *16*, 1449.
- Cammi, R., & Tomasi, J. (1995b). Nonequilibrium solvation theory for the polarizable continuum model: A new formulation at the SCF level with application to the case of the frequency-dependent linear electric response function. *International Journal of Quantum Chemistry: Quantum Chemistry Symposium*, *29*, 465.
- Cammi, R., & Tomasi, J. (1996). Time-dependent variational principle for nonlinear Hamiltonians and its application to molecules in the liquid phase. *International Journal of Quantum Chemistry*, *60*, 297.
- Cammi, R., Cossi, M., Mennucci, B., & Tomasi, J. (1996). Analytical Hartree–Fock calculation of the dynamical polarizabilities α , β , and γ of molecules in solution. *Journal of Chemical Physics*, *105*, 10556.
- Cammi, R., Mennucci, B., & Tomasi, J. (1999). Second-Order Møller–Plesset Analytical Derivatives for the Polarizable Continuum Model Using the Relaxed Density Approach. *Journal of Physical Chemistry A*, *103*, 9100.
- Cammi, R., Mennucci, B., & Tomasi, J. (2000a). An Attempt To Bridge the Gap between Computation and Experiment for Nonlinear Optical Properties: Macroscopic Susceptibilities in Solution. *Journal of Physical Chemistry A*, *104*, 4690.
- Cammi, R., Mennucci, B., & Tomasi, J. (2000b). Fast Evaluation of Geometries and Properties of Excited Molecules in Solution: A Tamm-Dancoff Model with Application to 4-Dimethylaminobenzonitrile. *Journal of Physical Chemistry A*, *104*, 5631.
- Cammi, R., Corni, S., Mennucci, B., Tomasi, J., Ruud, K., & Mikkelsen, K. V. (2002). A second-order, quadratically convergent multiconfigurational self-consistent field polarizable continuum model for equilibrium and nonequilibrium solvation. *Journal of Chemical Physics*, *117*, 13.
- Cammi, R., Frediani, L., Mennucci, B., & Rudd, K. (2003). Multiconfigurational self-consistent field linear response for the polarizable continuum model: Theory and application to ground and excited-state polarizabilities of para-nitroaniline in solution. *Journal of Chemical Physics*, *119*, 5818.
- Cammi, R., Corni, S., Mennucci, B., & Tomasi, J. (2005). Electronic excitation energies of molecules in solution: State specific and linear response methods for nonequilibrium continuum solvation models. *Journal of Chemical Physics*, *122*, 101513.
- Cammi, R., Fukuda, R., Ehara, M., & Nakatsuji, H. (2010). Symmetry-adapted cluster and symmetry-adapted cluster-configuration interaction method in the polarizable continuum model: Theory of the solvent effect on the electronic excitation of molecules in solution. *Journal of Chemical Physics*, *133*, 024104.
- Cancès, E., Mennucci, B., & Tomasi, J. (1997). A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *Journal of Chemical Physics*, *107*, 3092.
- Cappelli, C., Mennucci, B., Tomasi, J., Cammi, R., Rizzo, A., Rikken, G.L.J.A., Mathevet, R., & Rizzo, C. (2003). The Cotton–Mouton effect of furan and its homologues in the gas phase, for the pure liquids and in solution. *Journal of Chemical Physics*, *118*, 10712.
- Cappelli, C., Mennucci, B., Cammi, R., & Rizzo, A. (2005). Quantum Mechanical Polarizable Continuum Model Approach to the Kerr Effect of Pure Liquid. *Journal of Physical Chemistry B*, *109*, 18706.
- Caricato, M. (2011). Improving upon the reference reaction field approximation at no cost. *Journal of Chemical Physics*, *135*, 074113
- Caricato, M. (2012a). Absorption and Emission Spectra of Solvated Molecules with the EOM–CCSD–PCM Method. *Journal of Chemical Theory and Computation*, *8*, 4494.
- Caricato, M. (2012b). Exploring Potential Energy Surfaces of Electronic Excited States in Solution with the EOM–CCSD–PCM Method. *Journal of Chemical Theory and Computation*, *8*, 5081.

- Caricato, M. (2013a). A comparison between state-specific and linear-response formalisms for the calculation of vertical electronic transition energy in solution with the CCSD-PCM method. *Journal of Chemical Physics*, *139*, 044116.
- Caricato, M. (2013b). Implementation of the CCSD-PCM linear response function for frequency dependent properties in solution: Application to polarizability and specific rotation. *Journal of Chemical Physics*, *139*, 114103.
- Caricato, M. (2014). A corrected-linear response formalism for the calculation of electronic excitation energies of solvated molecules with the CCSD-PCM method. *Chemical Theory and Computation*, *10*(10), 99.
- Caricato, M., & Scalmani, G. (2011). On the Importance of the Orbital Relaxation in Ground-State Coupled Cluster Calculations in Solution with the Polarizable Continuum Model of Solvation. *Journal of Chemical Theory and Computation*, *7*, 4012.
- Caricato, M., Mennucci, B., Tomasi, J., Ingrosso, F., Cammi, R., Corni, S., & Scalmani, G. (2006). Formation and relaxation of excited states in solution: A new time dependent polarizable continuum model based on time dependent density functional theory. *Journal of Chemical Physics*, *125*, 124520.
- Caricato, M., Scalmani, G., Trucks, G. W., & Frisch, M. J. (2010). Coupled Cluster Calculations in Solution with the Polarizable Continuum Model of Solvation. *Journal of Physical Chemistry Letters*, *1*, 2369.
- Caricato, M., Scalmani, G., & Frisch, M. J. (2011). Brueckner doubles coupled cluster method with the polarizable continuum model of solvation. *Journal of Chemical Physics*, *134*, 244113.
- Christiansen, O., & Mikkelsen, K. V. (1999). A coupled-cluster solvent reaction field method. *Journal of Chemical Physics*, *110*, 1365; Christiansen, O., & Mikkelsen, K. V. (1999). Coupled cluster response theory for solvated molecules in equilibrium and nonequilibrium solvation. *Journal of Chemical Physics*, *110*, 8348.
- Christiansen, O., Jørgensen, P., & Hättig, C. (1998). Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy. *International Journal of Quantum Chemistry*, *68*, 1.
- Corni, S., Cammi, R., Mennucci, B., & Tomasi, J. (2005). Electronic excitation energies of molecules in solution within continuum solvation models: Investigating the discrepancy between state-specific and linear-response methods. *Journal of Chemical Physics*, *123*, 134512.
- Cossi, M., & Barone, V. (2001). Time-dependent density functional theory for molecules in liquid solutions. *Journal of Chemical Physics*, *115*, 4701.
- Cramer, C. J., & Truhlar, D. G. (2008). A Universal Approach to Solvation Modeling. *Accounts of Chemical Research*, *41*, 760.
- Curutchet, C., Orozco, M., Luque, F. J., Mennucci, B., & Tomasi, J. (2006). Dispersion and repulsion contributions to the solvation free energy: Comparison of quantum mechanical and classical approaches in the polarizable continuum model. *Journal of Computational Chemistry*, *27*, 1769.
- Ferrighi, L., Frediani, L., Fossgaard, E., & Ruud, K. (2007). Two-photon absorption of [2.2]paracyclophane derivatives in solution: A theoretical investigation. *Journal of Chemical Physics*, *127*, 244103.
- Frediani, L., rinkevicius, Z., & Ågren, H. (2005). Two-photon absorption in solution by means of time-dependent density-functional theory and the polarizable continuum model. *Journal of Chemical Physics*, *122*, 244104.
- Frenkel, J. (1934). *Wave mechanics: advanced general theory*. Oxford: Clarendon Press.
- Frisch, M. J., et al. (2009a) *Gaussian development version gdvH35*. Pittsburg: Gaussian Inc.
- Frisch, M. J., et al. (2009b). *Gaussian 09*. Pittsburg: Gaussian Inc.
- Fukuda, R., & Ehara, M. (2013). Mechanisms for Solvatochromic Shifts of Free-Base Porphine Studied with Polarizable Continuum Models and Explicit Solute-Solvent Interactions. *Journal of Chemical Theory and Computation*, *9*, 470.
- Fukuda, R., & Ehara, M. (2014). An efficient computational scheme for electronic excitation spectra of molecules in solution using the symmetry-adapted cluster-configuration interaction

- method: The accuracy of excitation energies and intuitive charge-transfer indices. *Journal of Chemical Physics*, *141*, 154104.
- Fukuda, R., Ehara, M., & Cammi, R. (2011). Nonequilibrium solvation for vertical photoemission and photoabsorption processes using the symmetry-adapted cluster-configuration interaction method in the polarizable continuum model. *Journal of Chemical Physics*, *134*, 104109.
- Fukuda, R., Chidthong, R., Ehara, M., & Cammi, R. (2012). Optical absorption and fluorescence of PRODAN in solution: Quantum chemical study based on the symmetry-adapted cluster-configuration interaction method. *Chemical Physics Letters*, *552*, 53.
- Fukuda, R., Ehara, M., & Cammi, R. (2014). Electronic excitation spectra of molecules in solution calculated using the symmetry-adapted cluster-configuration interaction method in the polarizable continuum model with perturbative approach. *Journal of Chemical Physics*, *140*, 064104.
- Fukuda, R., Ehara, M., & Cammi, R. (2015). Modeling Molecular Systems at Extreme Pressure by an Extension of the Polarizable Continuum Model (PCM) Based on the Symmetry-Adapted Cluster-Configuration Interaction (SAC-CI) Method: Confined Electronic Excited States of Furan as a Test Case. *Journal of Chemical Theory and Computation*, *11*, 2063.
- Gauss, J. (1999). Coupled-cluster theory. In P. v. R. Schleyer (Ed.) *Encyclopedia of computational chemistry* (Vol. I, pp. 617–636). New York: Wiley.
- Helgaker, T., Coriani, S., Jørgensen, P., Kristensen, K., Olsen, J., & Ruud, K. (2012). Recent Advances in Wave Function-Based Methods of Molecular-Property Calculations. *Chemical Reviews*, *112*, 543.
- Hirao, K., & Nakatsuji, H. (1982). A generalization of the Davidson's method to large nonsymmetric eigenvalue problems. *Journal of Computational Physics*, *45*, 246.
- Improta, R., Barone, V., Scalmani, G., & Frisch, M. J. (2006). A state-specific polarizable continuum model time dependent density functional theory method for excited state calculations in solution. *Journal of Chemical Physics*, *126*, 054103.
- Jansík, B., Rizzo, A., Frediani, L., Ruud, K., & Coriani, S. (2006). Combined density functional/polarizable continuum model study of magnetochiral birefringence: Can theory and experiment be brought to agreement?. *Journal of Chemical Physics*, *125*, 234105.
- Kim, H. J., & Hynes, J. T. (1990). Equilibrium and nonequilibrium solvation and solute electronic structure. I. Formulation. *Journal of Chemical Physics*, *93*, 5194.
- Klamt, A., Mennucci, B., Tomasi, J., Barone, V., Curutchet, C., Orozco, M., & Luque, F. J. (2009). On the Performance of Continuum Solvation Methods. A Comment on "Universal Approaches to Solvation Modeling" *Accounts of Chemical Research*, *42*, 489
- Koch, H., & Jørgensen, P. (1990). Coupled cluster response functions. *Journal of Chemical Physics*, *93*, 3333.
- Kongsted, J., Oesetd, A., Mikkelsen, K. V., & Christiansen, O. (2002). The QM/MM approach for wavefunctions, energies and response functions within self-consistent field and coupled cluster theories. *Molecular Physics*, *100*, 1813.
- Lin, N., Ferrighi, L., Zhao, X., Ruud, K., Rizzo, A., & Luo, Y. (2008). Solvent Effects on the Three-Photon Absorption of a Symmetric Charge-Transfer Molecule. *Journal of Physical Chemistry B*, *112*, 4703.
- (a) Marcus, R. A. (1956). Electrostatic Free Energy and Other Properties of States Having Nonequilibrium Polarization. *Journal of chemical physics*, *24*, 979; (b) Newton, M. D., & Friedman, H. L. (1988). Green function theory of charge transfer processes in solution. *Journal of Chemical Physics*, *88*, 4460; (c) Truhlar, D. G., Schenter, G. K., & Garrett, B. C. (1993). Inclusion of nonequilibrium continuum solvation effects in variational transition state theory. *Journal of Chemical Physics*, *98*, 5756.
- Marcus, R. A. (1992). Schroedinger equation for strongly interacting electron-transfer systems. *Journal of Physical Chemistry*, *96*, 1753.
- Mennucci, B., Tomasi, J., Cammi, R., Cheesman, J. R., Frisch, M. J., Davlin, F. J., Gabriel, S., & Stephens, P. J. (2002). Polarizable Continuum Model (PCM) Calculations of Solvent Effects on Optical Rotations of Chiral Molecules. *Journal of Physical Chemistry A*, *106*, 6102.

- Miertuš, S., Scrocco, E., & Tomasi, J. (1981). Electrostatic interaction of a solute with a continuum. A direct utilization of ab-initio molecular potentials for the prevision of solvent effect. *Chemical Physics*, *55*, 117.
- Nakatsuji, H., & Hirao, K. (1978). Cluster expansion of the wavefunction. Pseudo-orbital theory based on the SAC expansion and its application to the spin density of open-shell systems. *Journal of Chemical Physics*, *68*, 14279.
- (a) Olivares del Valle, F. J., & Tomasi, J. (1991). Electron correlation and solvation effects. I. Basic formulation and preliminary attempt to include the electron correlation in the quantum mechanical polarizable continuum model so as to study solvation phenomena *Chemical Physics*, *150*, 139; (b) Aguilar, M., Olivares del Valle, F. J., & Tomasi, J. (1991). Electron correlation and solvation effects. II. The description of the vibrational properties of a water molecule in a dielectric given by the application of the polarizable continuum model with inclusion of correlation effects, *Chemical Physics*, *150*, 151.
- Olsen, J., & Jørgensen, P. (1995). Time Dependent Response Theory with Applications to Self-Consistent Field and Multiconfigurational Self-Consistent Filed Wave Functions. In D. Yarkony (Ed.), *Modern electronic structure theory* (Vol. 2, p. 857). Singapore: World Scientific.
- Rizzo, A., Frediani, L., & Ruud, K. (2007). An ab initio investigation of the Buckingham birefringence of furan, thiophene, and selenophene in cyclohexane solution. *Journal of Chemical Physics*, *127*, 164321.
- Stanton, J. F. (1993). Many body methods for excited state potential energy surfaces. I. General theory of energy gradients for the equation of motion coupled cluster method. *Journal of Chemical Physics*, *99*, 8840.
- Stanton, J. F., & Bartlett, R. J. (1993). The equation of motion coupled cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *Journal of Chemical Physics*, *98*, 7029.
- Tomasi, J., & Persico, M. (1994). Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. *Chemical Reviews*, *94*, 2027.
- Tomasi, J., Cammi, R., & Mennucci, B. (1999). Medium effects on the properties of chemical systems: An overview of recent formulations in the polarizable continuum model (PCM). *International Journal of Quantum Chemistry*, *75*, 783.
- Tomasi, J., Mennucci, B., & Cammi, R. (2003). Molecules in solution. In S. Wilson (Ed.) *Handbook of molecular physics and quantum chemistry* (Vol. 3, pp. 299–328). New York: Wiley.
- Tomasi, J., Mennucci, B., & Cammi, R. (2005). Quantum Mechanical Continuum Solvation Models. *Chemical Reviews*, *105*, 2999.
- Zhao, K., Ferrighi, L., Frediani, L., Wang, C.-K., & Luo, Y. (2007). Solvent effects on two-photon absorption of dialkylamino substituted distyrylbenzene chromophore. *Journal of Chemical Physics*, *126*, 204509.

Spin-Orbit Coupling in Enzymatic Reactions and the Role of Spin in Biochemistry

35

Boris F. Minaev, Hans Ågren, and V. O. Minaeva

Contents

Introduction	1558
O ₂ Interaction with Heme, FAD, and Oxidases	1559
Definition of Spin and the Angular Momentum Conservation	1561
Molecular Oxygen Structure and Spectra	1563
Spin-Prohibition of Dioxygen Reactions	1565
Cytochrome Oxidases and Related Heme-Containing Enzymes	1566
Dioxygen Reaction with Glucose Oxidase	1569
Dioxygen Binding to Heme	1571
Spin-Orbit Coupling in O ₂ -Heme Interaction	1574
The Role of Spin in Biochemistry	1578
External Magnetic Field Effects in Biochemistry	1579
Spin Neuroscience	1581
Conclusions	1582
Bibliography	1584

B.F. Minaev (✉)

Organic Chemistry, Bogdam Khmel'nitskii National University, Cherkassy, Ukraine

Theoretical Chemistry, School of Biotechnology, Royal Institute of Technology, Stockholm, Sweden

e-mail: bfmin@rambler.ru

H. Ågren

Division of Theoretical Chemistry and Biology, School of Biotechnology, Royal Institute of Technology, Stockholm, Sweden

Organic Chemistry, Bogdam Khmel'nitskii National University, Cherkassy, Ukraine

e-mail: agren@theochem.kth.se

V.O. Minaeva

Organic Chemistry, Bogdam Khmel'nitskii National University, Cherkassy, Ukraine

e-mail: minaeva@cdu.edu.ua

Abstract

We review the general concept of nonadiabatic quantum spin transitions in biochemistry. A few important examples are highlighted to illustrate the concept: the role of spin effects in oxidases, cytochromes, in dioxygen binding to heme, in photosynthesis, and in tentative models of consciousness. The most thoroughly studied of these effects are connected with dioxygen activation by enzymes. Discussion on the mechanisms of overcoming spin prohibitions in dioxygen reactions with flavin-dependent oxygenases and with hemoglobin and myoglobin is presented in some detail. We consider spin-orbit coupling (SOC) between the starting triplet state from the entrance channel of the O₂ binding to glucose oxidase, to ferrous heme, and the final singlet open-shell state in these intermediates. Both triplet (T) and singlet (S) states in these examples are dominated by the radical-pair structures $D^+ - O_2^-$ induced by charge transfer; the peculiarities of their orbital configurations are essential for the SOC analysis. An account of specific SOC in the open π_g -shell of dioxygen helps to explain the probability of T-S transitions in the active site near the transition state. Simulated potential energy surface cross-sections along the reaction coordinates for these multiplets, calculated by density functional theory, agree with the notion of a relatively strong SOC induced inside the oxygen moiety by an orbital angular momentum change in the π_g -shell during the T-S transition. The SOC model explains well the efficient spin inversion during the O₂ binding with heme and glucose oxidase, which constitutes a key mechanism for understanding metabolism. Other examples of nontrivial roles of spin effects in biochemistry are briefly discussed.

Introduction

Many enzymes, for example, those that contain copper or iron-heme active sites (cytochromes, peroxidases), are paramagnetic and undergo spin transitions depending on the change of the metal site oxidation (Franzen 2002; Jensen and Ryde 2004; Prabhakar et al. 2003; Shaik et al. 2002; Shikama 2006; Sigfridsson and Ryde 2002; Stryer 1995). Oxygen molecules contain two unpaired valence electrons (Sawyer 1991) and are therefore also paramagnetic because of the nonzero total electron spin. The complete oxidation of organic materials by dioxygen in order to transfer them into CO₂ and water is highly exothermic (Sawyer 1991), but the kinetic barriers preclude, fortunately, this spontaneous combustion of living matter at normal conditions in the open air because of the spin selection rules (Prabhakar et al. 2003). Like mass and charge, spin is an integral property of the electron, however, spin reveals itself through the relativistic quantum equation as a pure quantum process associated with the structure of space-time, thus it is more fundamental. The role of this fundamental quantum feature of the electron in biochemistry could be extremely important, especially with respect to the mystery of the brain, the nature of consciousness (Penrose 1994), and oxygen activation by enzymes (Franzen 2002; Jensen and Ryde 2004; Prabhakar et al. 2003; Sawyer 1991; Shaik et al. 2002; Shikama 2006; Sigfridsson and Ryde 2002; Stryer 1995).

When O₂ reaches a red blood cell in the alveoli of the lung, where the gas exchange of carbon dioxide and oxygen takes place, the hemoglobin binds O₂ in such a strange way that the total electron spin of the whole heme-O₂ system is not conserved (Sigfridsson and Ryde 2002). In ordinary chemistry the spin conservation rule is usually fulfilled (Sawyer 1991; Shikama 2006; Stryer 1995), which is in agreement with the general conservation law of physics (Sawyer 1991). But the change of the electron spin is a rather common case in biochemistry of enzymes, in the respiration cycle and in other processes connected with dioxygen consumption (Blomberg et al. 2005; Friedman and Campbell 1987; Lane 2002; Minaev 2007; Minaev et al. 2007; Silva and Ramos 2008; Strickland and Harwey 2007). The spin change has to be compensated by an orbital angular momentum change in order to follow the general conservation law of the total angular momentum. The present review will show that this requirement is fulfilled in a number of enzymatic processes that play key roles in biochemistry; this is accomplished by an analysis of spin-orbit coupling perturbations.

With the appearance of photosynthesis on the primitive earth, a dramatic change in life occurred when living cells adapted themselves from anaerobic life to the aerobic form (Shikama 2006). Anaerobic glycolysis represented a successful attempt to extract some of the chemical energy from glucose. In contrast, the complete oxidation of glucose to CO₂ and H₂O by utilizing oxygen molecules was the most significant and crucial advancement in cellular metabolism. It thus allowed organisms to enhance their ability to utilize the chemical energy of sugar for a most effective consumption. In light of the high efficiency of energy conservation, this dramatic change from anaerobic to aerobic systems was designated as an “oxygen revolution” (Lane 2002; Penrose 1994; Shikama 2006; Stryer 1995). The most essential feature of this revolution was the involvement of spin-catalysis and quantum spin effects in ordinary chemistry (Minaev 2007). In fact, both processes, photosynthesis and respiration, are spin-forbidden; from one side we have diamagnetic reactants and the paramagnetic species from the other side $6\text{CO}_2 + 6\text{H}_2\text{O} + \Delta E \xrightleftharpoons{\text{chlorophyll}} \text{C}_6\text{H}_{12}\text{O}_6 + 6\text{O}_2$. In order to synthesize O₂ with chlorophyll from green plants beneath the sun and to use oxygen for glucose oxidation, nature has designed very smart chemical manipulations that allow spin prohibition in dioxygen activation to be overcome (Minaev 2007). With this revolution a large number of important enzymes that are based on spin-catalysis principles (Blomberg et al. 2000a; Jensen et al. 2005; Lane 2002; Minaev 2002, 2004, 2007, 2010; Minaev et al. 2007; Petrich et al. 1988; Prabhakar et al. 2002, 2004; Silva and Ramos 2008) have been developed and spread (Blomberg et al. 2000b; Klinman 2001; Minaev and Lunell 1993; Sheldon 1993).

O₂ Interaction with Heme, FAD, and Oxidases

Some oxidases include flavin adenine dinucleotide (FAD), which can activate O₂ to produce diamagnetic products without radical chain reaction steps (Minaev 2002; Stryer 1995). For example, glucose oxidase (GO) is a well-known enzyme that can bind dioxygen directly from air and produce hydrogen peroxide in the metabolic

cycle (Prabhakar et al. 2003). An X-ray structure of GO from *Aspergillus niger* is available (Klinman 2001) and used for simulation of the active site in theoretical modeling (Prabhakar et al. 2002). The histidine residue has been assigned an important role in the substrate oxidation by the GO active site.

In the heme biosynthesis pathway a deficient activity of the enzymes often leads to excessive excretion of porphyrins (Silva and Ramos 2008). These toxic products can induce porphyria diseases due to decreased activity of coproporphyrinogen III oxidase (CPO), which catalyzes the oxidative decarboxylation of propionic acid chains. This oxidase is similar to GO in that they both activate dioxygen without using metals and reducing agents. However, mechanisms by which these two oxidases produce diamagnetic peroxides from paramagnetic dioxygen are not properly understood on experimental grounds (Klinman 2001; Prabhakar et al. 2002; Silva and Ramos 2008), while some theoretical predictions seem to shed light on the specific quantum nature of such chemical transformations (Minaev 2007; Minaev et al. 2007; Prabhakar et al. 2003; Silva and Ramos 2008). Density functional theory (DFT) calculations (Prabhakar et al. 2002, 2003; Silva and Ramos 2008) indicate the common feature of spin transitions in these oxidases' mechanisms.

Hemoglobin and myoglobin are important globular proteins that reversibly bind the O₂ molecule. Both proteins contain ferrous iron of a heme group, which is usually simulated by Fe(II) porphyrin, where the iron ion is tetra-coordinated to the nitrogen atoms of the tetra-pyrrole rings (Franzen 2002; Jensen and Ryde 2004; Shikama 2006; Sigfridsson and Ryde 2002). The proximal histidine residue from the protein side chain is also bound to the Fe(II) ion leaving one empty position in the octahedral coordination sphere around the ferrous iron. Hemoglobin and myoglobin bind several small gaseous molecules (CO and NO) besides dioxygen (Sigfridsson and Ryde 2002). The binding of these diatomic ligands to heme has been studied in biochemistry for over a 100 years (Shikama 2006). Dissociation of CO, NO, and O₂ from the iron of heme occurs by both photolysis and thermolysis. In the former case, the fate of the diatomic ligands depends on the competition between geminate recombination and protein relaxation. Binding to heme occurs by the reverse of the thermal process (Franzen 2002).

It is well established that the structure of surrounding protein affects the ligand binding ability of the heme group; carbon monoxide binds to free heme in solution 10⁵ times more effectively than dioxygen, but in myoglobin this ratio is strongly reduced (Franzen 2002). Thus, myoglobin seems to favor O₂ binding compared to CO (Franzen 2002; Jensen and Ryde 2004; Shikama 2006; Sigfridsson and Ryde 2002). Such discrimination is of vital importance: otherwise we would suffocate from CO produced in our body during metabolism. The reason for the discrimination has been connected with different geometric parameters of O₂ and CO binding to the heme iron: the Fe–C–O bond is linear, whereas the Fe–O–O bond is bent (Franzen 2002). Both hemoglobin and myoglobin have another, distal, histidine ligand (His-64) positioned above the Fe(II) ion, but too far away to coordinate directly to the iron. His-64 is in the right position to affect the Fe–CO group and to produce a tension. This idea was supported by early crystal structures,

showing an Fe–C–O angle of 120–130° (Shikama 2006), an idea that has penetrated into the textbooks (Shikama 2006; Stryer 1995). Some newer X-ray measurements, IR spectra, and DFT calculations indicate that the Fe–C–O angle is nearly linear in the heme models; thus one finds the idea of a strongly bent Fe–C–O unit as the reason for the discrimination but which nowadays is agreed to be incorrect (Sigfridsson and Ryde 2002). Recent DFT calculations show that the FeO₂ group is much more polar than the FeCO group in heme models (Franzen 2002; Jensen and Ryde 2004; Shikama 2006; Sigfridsson and Ryde 2002). The electronic structure of the FeP complex with dioxygen is close to that of a superoxide anion bound to a ferric ion, and this charge-transfer complex is in the singlet spin state (Franzen 2002; Jensen and Ryde 2004; Shikama 2006; Sigfridsson and Ryde 2002). Thus, electrostatic interaction between the distal histidine and the Fe(III)–O–O[−] group is stronger than that for the Fe–CO group; therefore, the protein discriminates between O₂ and CO by hydrogen bonding and electrostatic interaction in myoglobin (Franzen 2002; Jensen and Ryde 2004; Sigfridsson and Ryde 2002).

Another important factor in such discrimination is determined by electron spin (Franzen 2002; Jensen and Ryde 2004; Minaev 2010; Shikama 2006; Sigfridsson and Ryde 2002; Stryer 1995). The O₂ binding to heme is highly non-exponential at ambient temperature with a rapid phase of 3 ps, a longer phase of similar geminate recombination (20–200 ps), and a slow bimolecular process of 1 ms; all phases are spin-dependent processes (Franzen 2002; Jensen and Ryde 2004; Petrich et al. 1988; Shikama 2006). At this point we need to consider it in more detail.

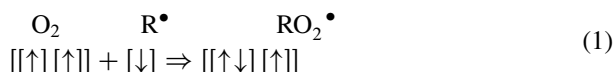
Definition of Spin and the Angular Momentum Conservation

Spin is derived from the Dirac equation and is a general intrinsic quantum property of electrons (Esposito 1999; Hagan et al. 2002; Hameroff and Penrose 1996; Hu and Wu 2004; Penrose 1960). Spin is an angular momentum with a length of the momentum vector $\sqrt{S^2} = \sqrt{S(S+1)}\hbar = (\sqrt{3}/2)\hbar$, where the spin quantum number for one electron is equal to $S = 1/2$ and two projection $M_S = \pm(1/2)\hbar$ described by α and β wave functions exist (the doublet state) (Esposito 1999). The role of electron spin in chemistry is rationalized in terms of spin-valence concept: a covalent chemical bond is formed when atomic orbitals of two electrons with opposite spins overlap each other (Stryer 1995). This spin wave function ($\alpha\beta - \beta\alpha$) is antisymmetric with respect to permutation of two electrons. The total spin of such pair is zero ($S = 0$) and the state is singlet (only one state; no spin, no intrinsic magnetic moment, the molecule is diamagnetic). When spins are parallel the total spin has a quantum number $S = 1$ in the equation for the length of the total spin angular momentum $\sqrt{S^2} = \sqrt{S(S+1)}\hbar = \sqrt{2}\hbar$ and there are three spin states ($\alpha\alpha$, $\beta\beta$, and $\alpha\beta + \beta\alpha$) with different projections on z-axis $M_S = \pm\hbar, 0$ (triplet state) (Penrose 1960). For a simple covalent σ -bond the triplet state is unstable and such simple molecule can exist only in the singlet state with two opposite spins. That is why almost all stable organic molecules contain an even number of electrons,

which can be divided into α and β pairs. They are diamagnetic, have paired spins and are in the singlet ground state.

Such diamagnetic substances usually do not change zero spin during chemical reactions (Blomberg et al. 2000a, 2005; Friedman and Campbell 1987; Jensen et al. 2005; Klinman 2001; Lane 2002; Minaev 2002, 2007, 2010; Minaev et al. 2007; Petrich et al. 1988; Prabhakar et al. 2002, 2004; Sawyer 1991; Silva and Ramos 2008; Strickland and Harwey 2007). The absence of the total electronic spin in stable molecules leads to the illusion that spin is not important in organic chemistry and biochemistry. In fact, the total electronic spin is the main regulating factor in many metabolic processes catalyzed by metal-organic enzymes, such as cytochromes (Blomberg et al. 2005; Minaev 2007; Minaev et al. 2007; Shaik et al. 2002; Stryer 1995), copper-aminoxidase (Prabhakar et al. 2003; 2004), and even in metal-free glucose oxidase (Minaev 2002; Prabhakar et al. 2002, 2003). Spin inversion is especially important in many dioxygen reactions, binding to heme (Jensen and Ryde 2004), combustion, and respiration (Minaev 2007). One can propose that the importance of spin inversion is also reflected in the Perutz model of the hemoglobin cooperativity (Franzen 2002; Minaev et al. 2007).

Dioxygen molecule is a well-known exception to the general rule: unlike many chemically stable organic compounds, the O_2 molecule has the triplet ground state (Sawyer 1991). According to Hund's rule, two unpaired electrons in two degenerate $\pi_{g,x}$ - and $\pi_{g,y}$ -orbitals have a lower repulsive energy in the triplet state compared to the singlet one. Because of this, oxygen is paramagnetic, it has intrinsic magnetic moment due to spins of two unpaired electrons and its addition to organic compounds is spin forbidden: starting reactants have the total spin $S = 1$ (from the O_2), whereas the oxidation products are diamagnetic ($S = 0$). This is the reason why organic matter may exist in the oxygen-rich atmosphere (Sawyer 1991). Because of the spin prohibition, combustion of organic fuels requires activation in the form of high-temperature ignition stage (Prabhakar et al. 2002; Sawyer 1991), i.e., generation of primary radicals. Reaction of radical R (the triplet, $S = 1/2$, doublet state) with O_2 molecule is spin-allowed, since the starting reactants ($O_2 + R^\bullet$) and product (RO_2^\bullet) both have the doublet states (reactants also possess quartet state, $S = 3/2$, being non-reactive),



which provide the radical chain character of the combustion reactions (Minaev 2007).

The quantum cell $[\uparrow]$ denotes (Eq. 1) a molecular orbital (MO) with α spin; double brackets embrace one molecule. The radical RO_2^\bullet can decompose into radical RO^\bullet and biradical $\bullet O$ thus providing a branching chain reaction. In radical chain combustion the energy is released in the form of heat and light without any specific control (until fuel exhaustion). Clearly, such mechanism of oxidation by molecular oxygen cannot be realized in living cells. Cells meet their energy needs in the course

of metabolic processes using strictly controlled energy of oxidation of organic compounds in their reactions with dioxygen, overcoming spin prohibition without high-temperature ignition step of a radical chain (Minaev 2007). An aerobic life evolved due to quantum kinetic prohibitions to reactions of paramagnetic oxygen with diamagnetic organic substances. The main reason for sluggish O₂ reactivity at the ambient conditions is the spin prohibition, namely, the starting reagents have two unpaired spins (from O₂ molecule), while in diamagnetic oxidation products (CO₂, H₂O, N₂) all spins are always paired. Overcoming this prohibition by generating radicals to interact with dioxygen (like in combustion) is not possible in living matter. Since the cells cannot resist large temperature gradients, they have to transform the energy released through oxidation to some kind of chemical energy prior to dissipation in the form of heat. This occurs by combining oxidation with ATP synthesis. All versatile energy-supplying metabolic processes and reactions occur under subtle enzymatic regulation, which is often spin-dependent (Blomberg et al. 2000a; Jensen et al. 2005; Minaev 2002, 2007, 2010; Petrich et al. 1988; Prabhakar et al. 2002, 2004).

Molecular Oxygen Structure and Spectra

The O₂ molecule in the triplet ground state has the following electronic configuration (Sawyer 1991) (1σ_g)²(1σ_u)²(2σ_g)²(2σ_u)²(3σ_g)²(1π_u)⁴(1π_g)² (Fig. 1). The two outer electrons in two degenerate 1π_g-MO's provide the lowest triplet state of the type [↑][↑], where the quantum cells [↑] denote the degenerate π_g-orbitals. These two unpaired electrons in antibonding π_g-MOs (Fig. 1) are responsible for the specific character of the dioxygen interaction with radicals (combustion) and chemically stable diamagnetic compounds (slow oxidation), as outlined above. Two antibonding π_g-vacancies make it possible to transform dioxygen into O₂⁻ and O₂²⁻ anions, the formation of the latter being strongly dependent on the presence of electron donors (enzymes) and magnetic perturbations that affect the spin prohibition.

There are four possible quantum states (Eq. 2) for the electronic configuration mentioned above. For the imaginary form of two degenerate 1π_g-MO's: 1π_g[±] = ψ(r, θ) e^{±iφ}, where θ is an angle between the radius vector (r) and the molecular axis (z), φ is the rotation angle of the radius vector about the z axis, the quantum states are represented by the scheme (Minaev 2007):

$$\begin{array}{cccc} [\uparrow][\uparrow] & [\uparrow\downarrow] & \square[\uparrow\downarrow] & [\uparrow][\downarrow] \\ \left| {}^3\Sigma_g^- \right\rangle & \left| {}^1\Delta_g \right\rangle & \left| {}^1\Delta_g \right\rangle & \left| {}^1\Sigma_g^+ \right\rangle \end{array} \quad (2)$$

Exchange interaction stabilizes the triplet state $\left| {}^3\Sigma_g^- \right\rangle$, while the degenerate singlet state $\left| {}^1\Delta_g \right\rangle$ is higher in energy by 22 kcal/mol, and the $\left| {}^1\Sigma_g^+ \right\rangle$ state is the

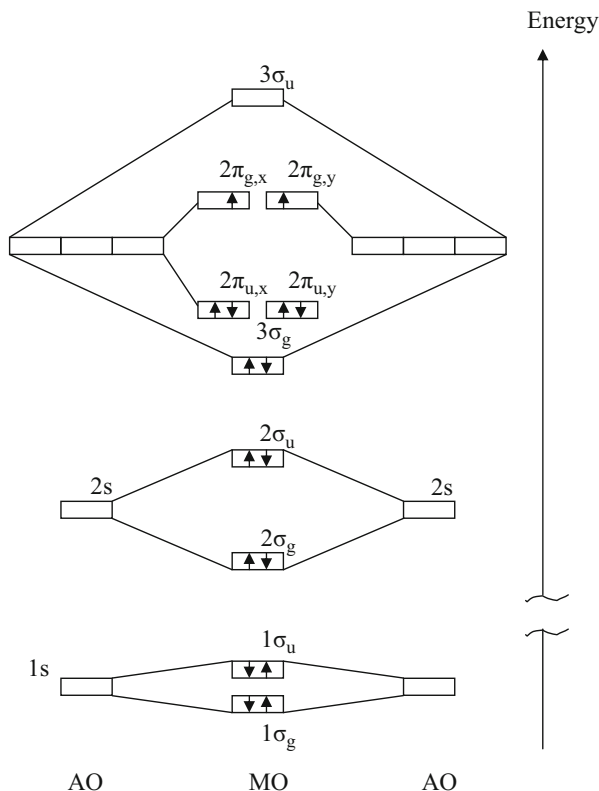


Fig. 1 Molecular orbital configuration of dioxygen

uppermost one (37 kcal/mol above the ground triplet $|^3\Sigma_g^- \rangle$ state). Interelectronic repulsion between the two electrons in the two closely space-distributed $1\pi_g$ -MO's, expressed through exchange integrals, is relatively high. This is an additional reason for a barrier in the $O_2(^3\Sigma_g^-)$ reactions with the closed shell molecules.

The $|^3\Sigma_g^- \rangle$ and $|^1\Sigma_g^+ \rangle$ states are mixed by weak internal magnetic perturbation, the so-called spin-orbit coupling (SOC) (Minaev 2007; Minaev and Lunell 1993). Some small admixture of one state (about 1.3%) is present in the other state, when SOC is additionally accounted for in the nonrelativistic Schrödinger equation. Such a small correction is very important because it makes the forbidden S-T transitions become allowed and can explain large difference in intensities of the atmospheric oxygen bands $^3\Sigma_g^- - ^1\Delta_g$ (1270 nm) and $^3\Sigma_g^- - ^1\Sigma_g^+$ (754 nm) (Minaev 1980, 1989, 2007; Minaev and Ågren 1995, 1996; Minaev and Ågren 1997; Minaev and Minaeva 2001; Minaev et al. 1993; 1996; 1997; 2008; Ogilby 1999; Paterson et al. 2006; Schweitzer and Schmidt 2003). In the following analysis

we want to understand the spin-dependent mechanism of the O₂ binding step once dioxygen has moved from the solvent through the protein chains and reached either the myoglobin distal cavity near the ferrous-heme cofactor or the flavin moiety in the active site of the GO enzyme. In both cases a similar SOC effect of the S-T states mixing is responsible for the biochemically important oxygen activation (Minaev 2002, 2007; Minaev and Ågren 1995; Prabhakar et al. 2002).

The heme-O₂ binding step has been extensively studied after flash-photolysis (Franzen 2002; Jensen and Ryde 2004; Shikama 2006; Sigfridsson and Ryde 2002), which shows interesting kinetic features, like non-homogeneous decay, recombination barriers, etc., and indicates complicated spin-dependence (especially in comparison with NO and CO binding to myoglobin). Kinetics of dioxygen reaction with GO has been investigated by studies of deuterium isotope effect during oxidation half reaction (Klinman 2001; Prabhakar et al. 2002). Explanation of such spin-dependence is the main purpose of the present review. In the beginning we want to consider some general features of spin-effect manifestation in chemistry and in dioxygen activation.

Spin-Prohibition of Dioxygen Reactions

Spins can undergo “depairing” when exposed to light; in (Eq. 3) an electron goes from a doubly occupied orbital [$\uparrow\downarrow$] of the ground singlet state to a vacant MO [] with the simultaneous spin flip to produce the triplet excited state (Minaev 2007):



According to the Pauli principle, both spins can be parallel (total spin $S = 1$) in such an excited state when two electrons occupy two different MOs. This triplet state has three possible orientations of the total spin vector, thus the singlet \rightarrow triplet excitation includes three possible transitions to three spin sublevels. All of them are spin-forbidden. This is a very strict prohibition, since it can be removed only by influence of magnetic interactions that are much weaker in general than the electric interactions. The latter determine the energetics of chemical bonding, electronic “depairing” excitation and the pathways of chemical reactions. The spin influences the energy through exchange interaction (Minaev 2007). A weak SOC slightly mixes the singlet and triplet states of molecules, which provides a non-zero rate for the $T \rightarrow S$ transitions. The late are observed in the form of phosphorescence and are well known as important quenching processes in photochemistry (Minaev 2004, 2007; Minaev and Minaev 2005). The nonradiative T-S transitions also play an important role in dark reactions, in particular, in catalysis (Metz and Solomon 2001; Minaev 2002). Weak SOC acts as a “key” needed to open a “heavy door”: that is, the system chooses a pathway of chemical reaction with low activation barrier in the triplet state instead of overcoming a high activation barrier in the singlet state. We remind that the exchange integral appears with different signs in the energy expression for the S and T states, namely, two radicals form a chemical bond in

the S state and repel each other in the T-state (Friedman and Campbell 1987). The different behavior of the S and T states is important not only for radical reactions, but also for many chemical transformations that include spin “depairing” during bond scission or proceed through biradical intermediates. This often occurs in catalysis by transition metal compounds (Minaev et al. 2007; Prabhakar et al. 2004; Sheldon 1993), especially in hemoproteins (Franzen 2002; Jensen and Ryde 2004; Shikama 2006; Sigfridsson and Ryde 2002; Stryer 1995).

The complete occupation of the degenerate-shell, Eq. 2, requires adding two additional electrons in order to become diamagnetic species, like usual chemical substance. Therefore, the oxygen molecule has a very strong tendency to take electrons from other substances and to make the complete electron-pairing in its unoccupied orbitals. This leads to the sequential formation of highly reactive or toxic oxygen species such as the superoxide anion O_2^- , peroxide anion O_2^{2-} , as by-products of many normal cellular metabolisms. Consequently, the development of enzymes to protect cells against such “oxidant stress” was of great urgency to aerobic organisms. This resulted in the ubiquitous occurrence of superoxide dismutase, catalase, peroxidase, and so on (Shikama 2006).

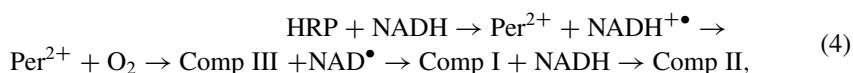
Cytochrome Oxidases and Related Heme-Containing Enzymes

Many heme-containing enzymes share the same active-species of the iron-ion types, but at the same time they exhibit significant differences in biochemical properties and reactivity with complicated spin-dependent kinetics (Burnold and Solomon 1999; de Winter and Boxer 2003; Kumar et al. 2005; Metz and Solomon 2001; Orlova et al. 2003; Shaik et al. 1998, 2002, 2005). Some of these heme-containing enzymes, cytochrome P450, cytochrome c oxidase, and horseradish peroxidase (HRP), are utilized as the active site of a high-valent oxo-iron heme species, called compound I (Shaik et al. 2002). Shaik et al. (Kumar et al. 2005; Shaik et al. 2002, 2005) proposed that the diverse reactions catalyzed by cytochrome P450 could be explained in terms of the difference in reactivity between the high-spin and low-spin states of the compound I. Their proposed two-state reactivity concept (Kumar et al. 2005; Shaik et al. 2002, 2005) has been widely used for the cytochrome P450 family and based on spin conversion transition between these states during reactions.

HRP, being similar to cytochrome P450 by crystal structure of the active site, differs in axial (proximal) ligand bound to the iron; P450 has a thiolate ligand of a cysteinate side chain, whereas HRP binds an imidazole group of a histidine residue side chain (Kumar et al. 2005). Although HRP functions predominantly as an electron sink (Kumar et al. 2005), it is also known to perform oxygen-transfer reactions (Afanasyeva et al. 2006; Chalkias et al. 2008; Shaik et al. 2002). Generally, however, the reactivity of HRP enzymes is sluggish compared with the P450 enzymes family; sulfoxidation is the most efficient while C-H hydroxylation is rather sluggish. This originates mainly in the much smaller substrate-binding pocket in HRP (Afanasyeva et al. 2006; Shaik et al. 2002). The recent DFT study of sulfoxidation by HRP and by cytochrome P450 reveals a spin-state selection

dependency on the proximal ligand: for thiolate ligand the reaction prefers the high-spin quartet state path while for the imidazole ligand the reaction mechanism involves a two-state reactivity (Kumar et al. 2005). This spin pattern also obeys orbital-selection rules, which are derived from the MO involvement of the ligands into the key oxo-iron porphyrin orbitals (Kumar et al. 2005; Shaik et al. 2005). In the case of the two-state reactivity the SOC-induced spin-transitions between the high-spin and low-spin reaction pathways could be magnetic-field dependent, if the viscosity of the media is sufficiently high (Afanasyeva et al. 2006; Buchachenko and Kouznetsov 2008; Buchachenko et al. 2005; Chalkias et al. 2008; Grissom 1995; Minaev 1983; Minaev and Lunell 1993; Serebrennikov and Minaev 1987). This is because SOC in the active sites is highly anisotropic and provides different rate constants for different spin-sublevels in zero field (Minaev 2002; Minaev and Lunell 1993), which become modified in the external field making the reaction to be magnetic-field dependent (Buchachenko and Kouznetsov 2008; Serebrennikov and Minaev 1987).

Such a dependence is found for the HRP-catalyzed oxidation of NADH, which can be presented by the schemes (Afanasyeva et al. 2006):

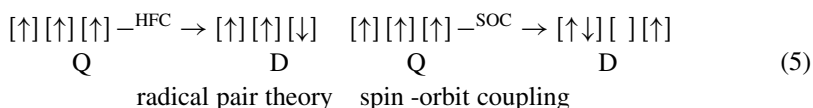


where compound II represents the Fe(IV) = O group connected with the porphyrin ring, compound I represents Fe(IV) = O connected with the radical cation of porphyrin, and compound III represents the porphyrin-containing Fe(IV)–OOH group. Here porphyrin denotes the protoporphyrin IX linked with protein chain by the histidine residue. The scheme (Eq. 4) is compatible with the magnetic-field dependence of the rate constants measured by stop-flow spectroscopy (Afanasyeva et al. 2006). The catalytic cycle starts with electron transfer from NADH to native HRP (Per³⁺) to produce NADH⁺• radical and the ferropoxidase intermediate (Per²⁺); (Per³⁺) actually means here the ferric porphyrin-Fe(III).

Interconversion of NADH to form NAD⁺ occurs by hydride transfer, in which the H⁺ ion and two electrons are transferred between the C(4) carbon atom of nicotinamide ring and substrate (Minaev et al. 1999). Catalytic cycle of NADH oxidation in the presence of hydrogen peroxide begins with a two-electron oxidation of the HRP enzyme to form Fe^{IV} and the porphyrin cation radical; this compound I is a highly reactive species that can accept one electron from NADH to form a NADH⁺• radical cation, which can undergo deprotonation to yield NAD[•] and compound II (Afanasyeva et al. 2006; Grissom 1995). The process is competitive with reverse electron transfer to regenerate NADH and Comp I, as shown by chemically induced dynamic nuclear polarization (CIDNP) of one of the C(4) protons in the NMR spectrum of NADH in presence of HRP (Afanasyeva et al. 2006). An alternative explanation of CIDNP is also possible (Minaev 2002; Serebrennikov and Minaev 1987), based on account of anisotropic SOC in compound I (Minaev 2010; Minaev and Lunell 1993; Minaev et al. 2007). The magnetic-field dependence of

the enzymatic reaction helps to reveal electron spin correlation among the catalytic states. Paramagnetic intermediates in the HRP catalytic cycle are difficult to detect, since direct observation by the EPR method depends on trapping unstable kinetically relevant radicals that may be transient or present at insignificant concentrations. Compound I is illusive in the sense that it does not accumulate in the cycle. This transient radical was inferred by cryogenic EPR/ENDOR technique (Davydov et al. 2008). The cryoreduction of the EPR-silent compound II produces Fe(III) species retaining the structure of precursor $[\text{Fe(IV)}=\text{O}]^{2+}$ or $[\text{Fe(IV)}=\text{OH}]^{3+}$; thus the EPR spectra of cryoreduced HRP II provide evidence of the low-spin hydroxy-Fe(III) heme species (Davydov et al. 2008; Kumar et al. 2005).

The first stage (Eq. 4) is thought to be magnitosensitive (Afanasyeva et al. 2006). Electron transfer from NADH to native HRP produces the quartet (Q) radical pair of Per^{2+} and $\text{NADH}^{+\bullet}$, which is unreactive toward recombination. The quartet – doublet (D) transition produces a radical pair that can recombine and quench the enzymatic process. It is claimed (Afanasyeva et al. 2006) that the Q-D transition is governed by isotropic hyperfine interaction and radical g-factors in terms of radical pair theory (Kumar et al. 2005). This Q-D transition is presented in the left part of the following scheme:



Here the set of three orbitals represents the triplet state of the Per^{2+} active site and one unpaired spin in the $\text{NADH}^{+\bullet}$ radical. The ferro-porphyrin intermediate (Per^{2+}) is known now to have a quintet ground state (Davydov et al. 2008) as follows from DFT calculations (Shaik et al. 2005). Thus, the main idea of the magnetic field mechanism (Afanasyeva et al. 2006) gives rise to some doubt. The most important objection is connected with the possible account of SOC, illustrated on the right side of the scheme (Eq. 5). It is possible to construct a number of configurations that could be definitely admixed to the open shell ground states of the Per^{2+} active site, which provides a strong SOC between Q-D states of the radical pair (Eq. 4). Such an SOC-induced mechanism could be applied to a number of magnetic field effects in metal-containing enzymes (Davydov et al. 2008; Engstrom et al. 1998, 2000; Gegear et al. 2008; Hoff 1986; Johnsen and Lohmann 2005; Maeda et al. 2008).

In the context of dioxygen activation we have to mention enzymes with binuclear metal centers, like cytochrome oxidase (Blomberg et al. 2000a, 2005), hemerythrin, and hemocyanin (Burnold and Solomon 1999; Metz and Solomon 2001). Generally speaking, the enzymes containing iron and copper active sites (Metz and Solomon 2001; Prabhakar et al. 2004) play key roles in dioxygen activation by generation of a peroxo intermediate; either O_2 is reduced by two electrons, provided by a binuclear metal site, or one electron is provided by metal and a second electron by a cofactor (Prabhakar et al. 2004). Cytochrome oxidase catalyzes the four-electron reduction of oxygen molecule to water (Stryer 1995). No intermediates were detected in the

reaction $O_2 + 4e^- + 4H^+ = 2H_2O$. However, many experimental measurements (Stryer 1995) have proved the formation of O_2^{2-} . The reaction center of cytochrome oxidase includes one heme ferrous ion and one copper ion (Blomberg et al. 2000a). The oxygen molecule binds to the heme Fe^{2+} cation and to the Cu^+ ion that donates one electron each to form an O_2^{2-} anion. This provides a way to overcome the major obstacle to oxygen activation, that is, spin inversion (T-S transition). The O_2^- and O_2^{2-} species have the following ground state configurations (in addition to Fig. 1):

$$\begin{array}{l} [\uparrow\downarrow][\uparrow] \quad [\uparrow][\uparrow\downarrow] \quad [\uparrow\downarrow][\downarrow\uparrow] \\ O_2^- |^2\Pi_g > \quad O_2^- |^2\Pi_g > \quad O_2^{2-} |^1\Sigma_g^+ > . \end{array} \quad (6)$$

Since the O_2^{2-} dianion has a filled electron shell, the ground state of this species is totally symmetric and characterized by the term $^1\Sigma_g^+$. Transfer of two electrons causes the ground-state term, $^3\Sigma_g^-$ of the O_2 molecule (Fig. 1), to transform smoothly to the term $^1\Sigma_g^+$ of the dianion. The spin-orbit coupling between the states $^3\Sigma_g^-$ and $^1\Sigma_g^+$ is symmetry-allowed (Minaev 1980, 2007); therefore, the reduction $O_2 \rightarrow O_2^{2-}$ is also symmetry allowed with inclusion of SOC. Transition of the active site of cytochrome oxidase to the singlet state removes spin prohibition for subsequent fast chemical reactions up to formation of stable diamagnetic products (Minaev 2010). The orbital doubly degenerate $|^2\Pi_g >$ ground state of the O_2^- ion is split by strong SOC (Minaev 2007), since it has a nonzero orbital angular momentum. This is an important key aspect of many enzymatic O_2 -activation reactions, considered in the following sections.

Dioxygen Reaction with Glucose Oxidase

It is often assumed (Sawyer 1991; Shikama 2006; Stryer 1995) that one can overcome the spin prohibition to oxidation of organic substrates with atmospheric oxygen by successive addition of single electrons and protons in the successive reduction of O_2 . Further reactions of diamagnetic hydrogen peroxide, produced in such reduction, are spin allowed. It is assumed (Sawyer 1991; Sheldon 1993; Stryer 1995) that removal of spin prohibition in such reactions proceeds as in the case of radical-chain oxidation, where the spin prohibition can be removed upon formation of primary radicals. It is important to stress a fundamental difference between the enzymatic reactions involving radicals and the radical reactions in chain oxidation processes. In the latter case radicals go to the bulk of the gaseous plasma flame (or in the solution bulk) and no longer retain the “spin memory” about precursors. All participants of biochemical oxidation reactions, i.e., dioxygen and electron transfer agents, are confined within the same active site of enzyme. If an electron is transferred to the oxygen molecule from a diamagnetic enzyme M , i.e., $O_2 + M \rightarrow O_2^- + M^+$, it produces a triplet radical ion-pair (triplet precursor), all spins remain correlated, the “spin memory” is retained, and the spin prohibition

to subsequent reactions of the radical ion-pair thus generated is not removed and cannot lead to a singlet product.

For example, reaction of O_2 with glucose oxidase (Klinman 2001; Minaev 2002; Prabhakar et al. 2002, 2003) involves flavin adenine dinucleotide (FAD) and includes two stages; namely, glucose oxidation to glucosolactone with reduction of FAD to $FADH_2$ and the reverse cycle $FADH_2 \rightarrow FAD$, with reduction of O_2 to H_2O_2 . From the standpoint of dioxygen activation it is interesting to consider only the second stage (Fig. 2). In this model the protonated histidine residue (His516) is included, which is in close proximity to FAD as follows from X-ray analysis (Klinman 2001). O_2 can occupy the cavity between $FADH_2$ and His516 (Fig. 2a). DFT calculations indicate that an electron is transferred immediately from the reduced cofactor to dioxygen (Fig. 2b); the electron transfer process is determined by low ionization potential of $FADH_2$, relatively high electron affinity of O_2 (0.45 eV), and attraction of the O_2^- to the protonated histidine. After formation of a triplet radical pair, $FADH_2 + O_2 \rightarrow FADH_2^{\bullet} + O_2^{\bullet-}$, the $T \rightarrow S$ transition has to occur in order to provide the final diamagnetic products $FAD + H_2O_2$, to which the singlet-spin stage (Fig. 2c) precedes.

After the $T \rightarrow S$ transition in the ion-radical pair (Fig. 2b), an ordinary chemical transformation on the singlet state PES occurs (Fig. 2c). It involves abstraction of hydrogen atom from the N1 atom of $FADH_2^{\bullet}$ and a proton abstraction from the nearest histidine residue in order to create H_2O_2 by reduction of the superoxide anion. This process, accompanied by the formation of hydrogen peroxide, can occur

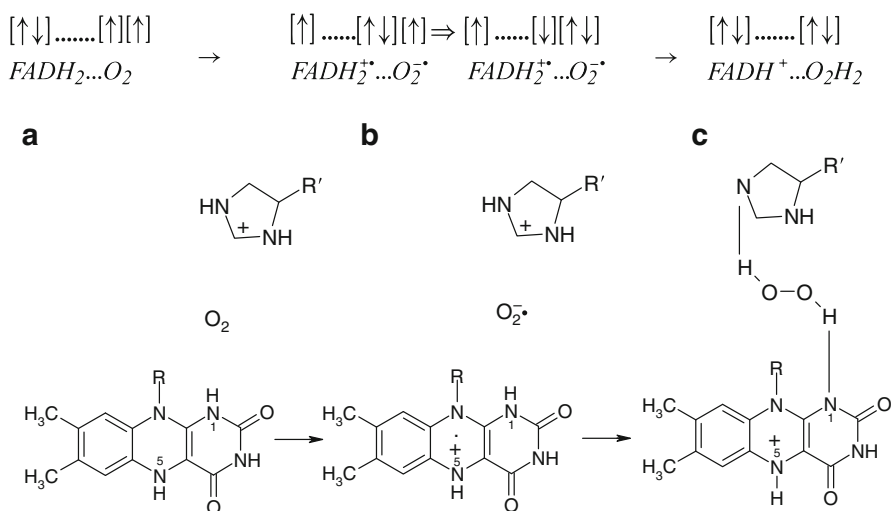


Fig. 2 An active center of the glucosidase model. The first phase of the catalytic cycle (not shown) includes the FAD to $FADH_2$ reduction. (a) starting model of the dioxygen entrance to the active site; (b) electron transfer stage and spin transition in the radical pair; (c) hydrogen peroxide production. The final stage (not shown) includes the proton transfer from the N5 atom to the histidine residue along the hydrogen-bond network in peptide chain of the enzyme

only in the singlet state. The final phase of the catalytic cycle (not shown in Fig. 2) consists of a subsequent proton transfer from FADH^+ ion back to histidine across the system of H-bonds in water-protein chain (Prabhakar et al. 2003). This proton transfer does not change the number of active electrons and the closed-shell spin state. The $T \rightarrow S$ transition has been explained (Prabhakar et al. 2003, 2004) by a relatively large SOC between the S and T states of the radical pairs (Fig. 2b), which have different orbital structures inside the superoxide ion. As one can see in the orbital-configuration scheme (Fig. 2), the $T \rightarrow S$ transition includes an electron jump from one $\pi_{g,x}$ molecular orbital of the dioxygen to another $\pi_{g,y}$ orbital (Spin on the FADH_2^{\bullet} moiety is non-active during the transition). Such transformation is equivalent to orbital rotation, or to a torque, which creates a transient magnetic field; finally this magnetic field induces a spin flip (Prabhakar et al. 2003, 2004). Thus, the reductive activation of dioxygen by GO is favored by an increased electron affinity of O_2 due to the proximity of protonated His516-residue and a fast electron transfer followed by the $T \rightarrow S$ transition induced by strong SOC in superoxide ion. This simple consideration is supported by direct quantum-mechanical calculations of the SOC integrals (Minaev 2010; Silva and Ramos 2008). Exactly the same mechanism of the $T \rightarrow S$ transition is applied for explanation of activity of coproporphyrinogen III oxidase (CPO) in catalysis of the oxidative decarboxylation during the heme biosynthesis (Silva and Ramos 2008). Silva and Ramos have supported the so-called Lash's model of CPO oxidation by dioxygen (Silva and Ramos 2008) and obtained the T-S crossing point that lies only 8 kcal/mol above the reactant state. They have calculated the SOC integral between T and S potential energy surfaces crossing to be equal to 77.45 cm^{-1} in an excellent agreement with the earlier prediction for GO reaction (Minaev 2002; Prabhakar et al. 2002).

A similar mechanism of the SOC during the $T \rightarrow S$ transition has been used to explain bioluminescence induced by the luciferase enzyme (Orlova et al. 2003). Attachment of triplet dioxygen to the closed shell luciferin to yield a singlet state intermediate (which finally provides bioluminescence) is a spin-forbidden process (Orlova et al. 2003). It is important to the use of previous GO model (Minaev 2002; Prabhakar et al. 2002) that a critically significant protonated histidine residue resides in the proximity of the acidic group of luciferin (Orlova et al. 2003), quite similar to the GO active site (Fig. 2). Photodestruction of this histidine inactivates luciferase completely, thus it is found plausible that the SOC-induced spin-flip in superoxide ion (Prabhakar et al. 2002) operates in the case of oxygenation by luciferase (Orlova et al. 2003). In the following we want to show that a similar mechanism of SOC enhancement by charge transfer can be applied for spin-dependent reaction of dioxygen binding to heme.

Dioxygen Binding to Heme

The O_2 binding with myoglobin model was studied recently by DFT methods (Blomberg et al. 2005; Franzen 2002; Jensen and Ryde 2004; Jensen et al. 2005; Minaev 2007; Minaev et al. 1993, 2007; Sigfridsson and Ryde 2002; Strickland and

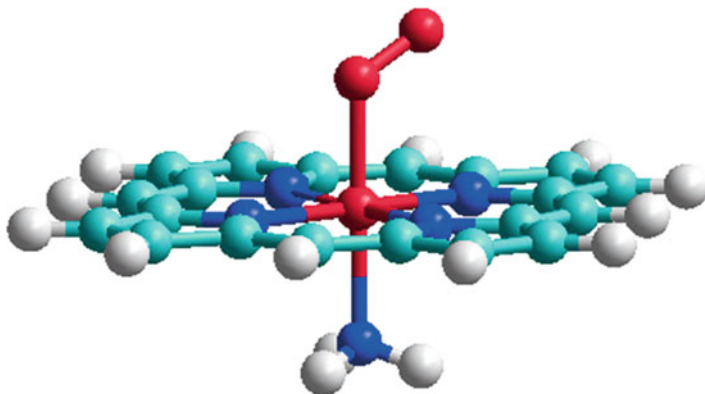


Fig. 3 A simple model of Fe(II)-porphine-NH₃-O₂ model of oxyhemoglobin (heme-cofactor coordinated with O₂. Simulates the heme (with NH₃ as a model of a proximal histidine residue)

Harvey 2007). Fully relaxed potential energy curves (PEC) were calculated for the seven lowest electronic states in Ref. Sigfridsson and Ryde (2002), Strickland and Harvey (2007), and Blomberg et al. (2005), while the PEC for spin states $S = 0, 1, 2, 3$ at fixed geometry as a functions of the Fe–O₂ distances were presented in Ref. Franzen (2002). In this work we have recalculated some points of the fixed Fe–O distance (1.8, 2.0 and 2.5 Å) with full geometry optimization of all other parameters for all possible multiplets and accounting for different symmetries (A' and A'') for the singlet and triplet states. The model of oxyheme is shown in Fig. 3, which includes Fe(II)-porphyrin coordinated with NH₃ as a model of proximal histidine residue. Its calculation gives the similar singlet ground state as for the simplified Fe(II)-Porphin-Imidazole-O₂-model, which poses C_s symmetry (Sigfridsson and Ryde 2002). The simpler Fe(II)-Porphine-NH₃-O₂ model (Fig. 3) has been used in this work for DFT calculations in the vicinity of the equilibrium. The B3LYP/6-31G* method (Frisch et al. 2003) has been employed and the results quite close to those presented in Ref. Sigfridsson and Ryde (2002) have been obtained. For the Fe(II)-Porphyrin-NH₃-O₂ model, shown in Fig. 3, all vibrational frequencies and their intensity in the infrared and Raman spectra have been calculated. Many normal modes are similar to those calculated in Refs. Minaev et al. (2007) and Minaev (2002). An additional Fe–O₂ stretching vibrational frequency is calculated at 539 cm⁻¹, which agrees qualitatively well with the resonance Raman band, observed at 567 cm⁻¹ for oxyhemoglobin by Soret excitation (Potter et al. 1987). This indicates reliability of the chosen model and of the DFT method used in this work.

In the entrance channel of the O₂ binding reaction to heme we have a number of different multiplets. At the infinite separation the deoxyheme has a quintet ground state with the triplet state being very close in energy. This is in agreement with the experimental data (Shikama 2006), showing that the isolated deoxyheme is a high-spin quintet (Friedman and Campbell 1987). The optimized structure of this Fe(II)P complex with NH₃ at the fifth coordination position agrees with the X-ray

analysis of the crystal structure of deoxymyoglobin. In this case the porphyrin Fe–N distances (2.08 Å) are larger than in the low-spin states of deoxymyoglobin (2.0 Å) and the iron ion is above the porphyrin ring plane by 0.3 Å in agreement with the X-ray data (0.36 Å) (Friedman and Campbell 1987; Sigfridsson and Ryde 2002). This illustrates the well-known fact that the high-spin iron ion Fe(II)(⁵D) is too large to fit into the porphyrin ring cavity (Jensen and Ryde 2004).

When this deoxyheme interacts with the triplet ground state dioxygen O₂ ($X^3\Sigma_g^-$), there are six unpaired electrons (two from O₂ plus 4 from heme). Their interaction can provide a variety of possible total spin states. The maximum spin corresponds to the septet ($^7A''$, $S = 5$) state, when both subsystems have parallel spins. If they are anti-parallel, the triplet state $^3A''$ occurs. At long Fe–O distances ($R > 2.5$ Å) these states are degenerate because of the absence of exchange interactions between O₂ and heme. The intermediate quintet $^5A''$ state is also degenerate together with the triplet and septet. The A'' symmetry is determined by the oxygen degenerate π_g orbitals, which have the a' and a'' symmetry, respectively, with respect to the plane, which contains the O₂ and Fe–N bond in ammonia molecule.

In general, all spin states that occur at each random collision of heme and O₂ should lead to oxygen binding. But the rate constants would be spin-dependent and different for the triplet, quintet, and septet states (even for different spin sublevels of one multiplet). More detailed information about O₂ binding has been obtained in flash-photolysis studies of O₂ dissociation from heme, when the fate of dioxygen depends on the competition between intrinsic recombination rate constant and protein relaxation, as well as the O₂ escape from the protein (Shikama 2006). The greatly different recombination dynamics of O₂, CO, and NO molecules with heme have been attributed to spin states of each ligand and to their possible combinations with the iron spin (Franzen 2002; Jensen and Ryde 2004; Shikama 2006; Sigfridsson and Ryde 2002). The observed recombination kinetics can be influenced by protein dynamics, if the intrinsic recombination rate constant is slower than this dynamics. By studies of viscosity and temperature dependence, the general averaged time scale for the recombination rate constant can be estimated (Friedman and Campbell 1987; Jensen and Ryde 2004; Shikama 2006).

It is interesting to compare O₂, CO, and NO molecules in this respect (Blomberg et al. 2005; Friedman and Campbell 1987; Strickland and Harwey 2007). The CO recombination with heme is a single exponential process characterized by a slow rate constant ($k \approx 10^6$ s⁻¹) at the ambient temperature and a low solvent viscosity (below viscosity of globin). This intrinsic (geminate) recombination rate is slower than both protein relaxation and CO escape, thus the recombination yield is very small (0.04) (Shikama 2006). The rebinding of the NO stable radical (doublet ground state, $S = 1/2$) is characterized by two-exponential kinetics with the rapid ($k_1 \approx 10^8$ s⁻¹) and slow ($k_2 \approx 5 \cdot 10^6$ s⁻¹) rate constants under ambient conditions (Friedman and Campbell 1987; Shikama 2006). Since the ground state heme has the quintet spin state ($S = 2$), there are two starting states: ($S = 2\frac{1}{2}$ and $S = 1\frac{1}{2}$) depending on mutual spin orientation of two species, while the recombination NO-heme product has the doublet state ($S = 1/2$) (Friedman and Campbell

1987; Strickland and Harwey 2007). Both types of geminate recombination require spin change. The slow process includes two-step spin-flip transformations ($S = 2\frac{1}{2} \rightarrow (S = 1\frac{1}{2}) \rightarrow (S = \frac{1}{2})$), since spin-orbit coupling can mix states and induces spin transition with the selection rule $\Delta S = 1$ (Minaev 2010). The rapid recombination occurs in one step ($S = 1\frac{1}{2} \rightarrow (S = \frac{1}{2})$).

The CO molecule is diamagnetic; all spins are paired, the total spin is zero. The heme-CO adduct is also diamagnetic ($S = 0$). Since the ground state heme has the quintet spin state ($S = 2$), the geminate recombination reaction is doubly spin forbidden. First it should be a quintet-triplet transition, which needs to overcome an additional activation barrier (besides the spin flip, induced by SOC) and then a final triplet-singlet transition. This is the reason why the CO recombination with heme is so slow, in spite of the high binding energy (Franzen 2002; Jensen and Ryde 2004; Shikama 2006; Sigfridsson and Ryde 2002; Strickland and Harwey 2007).

It is interesting to make the comparison: inorganic ferric ion (Fe^{3+}) has some catalytic activity for the decomposition of H_2O_2 into water and oxygen. When the ion is incorporated into the porphyrin molecule to form heme, the molecule is about a thousand times more effective than Fe^{3+} alone. If the protein component of the enzyme catalase then adds to the heme, the catalytic efficiency increases by a further factor of 10^9 times (Shikama 2006). This means that paramagnetic spin catalysis in combination with the SOC-induced spin-catalysis provide some kind of synergetic effect.

Spin-Orbit Coupling in O_2 -Heme Interaction

In this review we shall consider dioxygen binding in more detail, since the O_2 molecule is the main subject of spin-dependent biochemical phenomena in a general context. The ground state of the oxyheme product is an open-shell singlet in agreement with EPR experiment and Mossbauer spectra (Friedman and Campbell 1987; Sigfridsson and Ryde 2002) (the closed-shell singlet has been obtained in a number of calculations (Jensen et al. 2005; Shikama 2006), however, this result has since been revised (Blomberg et al. 2005; Strickland and Harwey 2007)). Thus, the reaction of O_2 binding to heme is spin forbidden. At least the T-S transition has to occur (Franzen 2002). Such spin flip can be induced by spin-orbit coupling (SOC) between the T and S states. One has to calculate the matrix element of the SOC operator, which can be presented in the effective single-electron approximation as

$$H_{\text{SO}} = \sum_A \zeta_A \sum_i \vec{l}_{i,A} \cdot \vec{s}_i = \sum_i \vec{B}_i \cdot \vec{s}_i = \sum_i (B_{i,x} s_{i,x} + B_{i,y} s_{i,y} + B_{i,z} s_{i,z}). \quad (7)$$

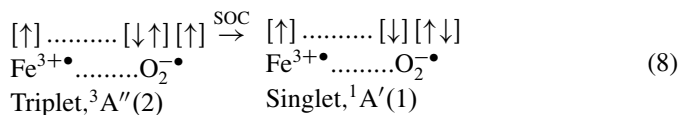
In (Eq. 7) ζ_A is a SOC constant for atom A ($\zeta_0 = 153 \text{ cm}^{-1}$), $\vec{l}_{i,A}$, \vec{s}_i – are the orbital and spin angular momentum operators for the i -th electron, respectively. This is the effective single-electron SOC approximation, which proved to be useful

in many spectroscopic and chemical applications including oxygen spin-forbidden atmospheric bands (Minaev 1989; Minaev and Ågren 1997; Minaev et al. 1996, 1997, 2008; Ogilby 1999; Paterson et al. 2006; Schweitzer and Schmidt 2003) and spin-forbidden enzyme reactions (Minaev 2002, 2007, 2010; Minaev et al. 2007; Prabhakar et al. 2003, 2004).

Deoxyheme has a quintet ground state (four spins are unpaired, $S = 2$) (Minaev et al. 2007; Sigfridsson and Ryde 2002) and the adduct with the triplet dioxygen (two unpaired spins) would be expected to have either six ($4 + 2 = 6$) unpaired spins or two ($4 - 2 = 2$) unpaired spins depending on ferromagnetic or antiferromagnetic relative orientation of the two magnetic moments of the deoxyheme and O_2 . An intermediate quintet spin state is also possible for the ground state species coupling. The triplet state of deoxyheme, being very close in energy, produces adducts with the triplet O_2 , which could be either quintet ($S = 2$), triplet ($S = 1$), or singlet ($S = 0$) depending on if it is ferromagnetic or antiferromagnetic coupling of two species. Thus, only triplet deoxyheme could provide the ground singlet state product in the process of the antiferromagnetic coupling with the triplet O_2 in a spin-allowed oxyheme formation without spin flip. Spin transition from the ground quintet to the close-lying triplet deoxyheme can be induced by SOC in the third shell of the iron ion. In this case the primary electronic reorganization takes place in the ferrous ion at the equilibrium between the quintet and triplet states already before dioxygen approaches the deoxyheme (Franzen 2002; Jensen and Ryde 2004). All other recombination processes include spin flip induced during heme- O_2 interaction; they seem to be more important for dioxygen binding (Blomberg et al. 2005; Franzen 2002; Jensen and Ryde 2004; Minaev 2007; Minaev et al. 2007; Shikama 2006; Strickland and Harwey 2007). Such natural heme - O_2 reactions could start with the $^3A''$ (2) state, which is repulsive at shorter distance ($R < 2.5$ Å) or with the septet $^7A''$ (1) state (Jensen and Ryde 2004) (both are the ground state of the entrance channel heme + O_2 and go in parallel with some other multiplets until the short distances limit (2.5–3 Å)). The energy gap is about 0.1 eV at these limits in agreement with Ref. Jensen and Ryde (2004). The optimized singlet ground state $^1A'$ (1) in the reaction product is lower in energy than other multiplets at least by 0.4 eV; this oxyheme is an open-shell singlet of a complicated orbital and spin structure (Franzen 2002; Jensen and Ryde 2004; Sigfridsson and Ryde 2002). It has a short Fe–O distance (1.81 Å) (Jensen and Ryde 2004) (reproduced in our DFT calculations, 1.84 Å) in contrast to the high-spin states (2–2.7 Å). Our result is close to the $Fe^{3+} - O_2^-$ radical-pair structure in agreement with other DFT calculations and with Weiss' model (Franzen 2002; Jensen and Ryde 2004; Shikama 2006; Sigfridsson and Ryde 2002). The spin densities in oxyheme are equal to 0.94, –0.31, and –0.72 for the Fe–O–O chain, respectively, at the optimized bond angle of 118° . The O–O bond distance (1.36 Å) and vibration frequency (1110 cm^{-1}) correspond better to superoxide ion (Jensen and Ryde 2004; Minaev 2007).

Potential energy surfaces for O_2 binding with the heme model (deoxy-Fe-porphyrin bound with imidazole) have been calculated by Jensen et al. (Jensen and Ryde 2004), who optimized by DFT method the C_s symmetry restricted reaction with fixed Fe–O bond length, which was systematically increased point by point.

Jensen et al. (Jensen and Ryde 2004) found that the spin-change occurs easily due to a broad crossing region of five electronic states. Similar PES crossing are obtained in other studies (Franzen 2002; Minaev et al. 1993; Strickland and Harwey 2007). Accounts of our data and the results of Refs. Sigfridsson and Ryde (2002), Franzen (2002), Jensen and Ryde (2004), and Strickland and Harwey (2007) allow us to consider the following scenario of the O₂ binding to myoglobin. At the intermediate distances 2.5–3 Å the starting ³A'' (2) state from the entrance channel transfers to the triplet Fe³⁺ – O₂⁻ radical-pair. In this region there are few crossing points between S and T states, including the ³A'' (2) – ¹A' (1) states crossing, where spin change could occur (Franzen 2002; Jensen and Ryde 2004; Sigfridsson and Ryde 2002). A simplified electronic structure of the ³A'' (2) and ¹A' (1) states near the crossing of the potential energy surfaces is presented in scheme (7). Outer electrons of the ground triplet state dioxygen in two degenerate π_g-MO's provide a scheme [↑][↑]; electron transfer from Fe²⁺ to O₂ in order to produce the radical pair Fe³⁺ – O₂⁻ (Eq. 8) can be accomplished by the occupation of either π_{g,x}- or π_{g,y}-orbitals. Both radical pairs could be in T and S states; all four states are almost degenerate at the intermediate distances. The most interesting spin states are those presented in scheme (Eq. 8), since they correspond to the desired T → S transition and to the final product of the O₂ binding by heme. The scheme (Eq. 8) is equivalent to the scheme (Eq. 5) and the same explanation for the high SOC matrix element (Minaev 2002; Prabhakar et al. 2002) is relevant.



In this model SOC perturbation occurs entirely in the oxygen moiety and the Fe(III) ion is magnetically silent. The starting triplet radical pair corresponds to a charge-transfer (CT) state described by ³A'' (2) wave function ${}^3\Psi_{CT_x} = \Re \left| (3d\alpha) (\pi_{g,x}\beta) (\pi_{g,x}\alpha) (\pi_{g,y}\alpha) \right|$, that is, transfer of an electron to the π_{g,x}-orbital of O₂ molecule, whereas the singlet radical pair corresponds to a CT state described by ¹A' (1) wave function ${}^1\Psi_{CT_y} = \Re \left| (3d\alpha) (\pi_{g,x}\beta) (\pi_{g,y}\alpha) (\pi_{g,y}\beta) \right|$ and a transfer of an electron to another degenerate orbital of oxygen, π_{g,y}. Here \Re means a proper anti-symmetrization of the wave function. Spin-orbit coupling between these CT states is the maximum possible for a system comprised of such light atoms like oxygen (Minaev 2002, 2004, 2007). The considered matrix element of the SOC operator (Eq. 7) is easily estimated to be equal:

$$\begin{aligned}
 \langle {}^3A''(2) | H_{SO}^Z | {}^1A'(1) \rangle &= \langle {}^3\Psi_{CT_x} | H_{SO}^Z | {}^1\Psi_{CT_y} \rangle = \frac{1}{2} \langle \pi_{g,x} | B_z | \pi_{g,y} \rangle \\
 &= \frac{i}{2} \xi_o = 76.5 \text{ icm}^{-1}
 \end{aligned} \quad (9)$$

The value is identical to those calculated for similar charge-transfer states including the GO (Minaev 2002; Prabhakar et al. 2002) and coproporphyrinogen III oxidase

(Silva and Ramos 2008). This SOC matrix element (Eq. 9) is very close to a value of about 80 cm^{-1} , postulated in estimation of the Landau-Zener rate constant for T-S transitions in spin-dependent O_2 and NO binding to heme (Franzen 2002; Friedman and Campbell 1987; Petrich et al. 1988). With such a proposal for the generally unknown SOC integral (Shaik et al. 2002) a quite reasonable estimation for the spin-dependent rate constants of the CO, O_2 , and NO recombination in heme proteins are obtained (Franzen 2002; Petrich et al. 1988). For CO binding to heme a quintet-triplet-singlet step-wise transition is necessary, which explains the million times slower recombination rate in this case in comparison with the O_2 and NO recombination in heme proteins (Franzen 2002; Shaik et al. 2002; Sigfridsson and Ryde 2002). The gradient difference at the location of crossing points that enters the denominator of the Landau-Zener expression of the rate constant for spin transition is quite small ($0.1\text{--}0.2 \text{ eV/\AA}$) for O_2 and NO binding to heme (Franzen 2002; Jensen and Ryde 2004), thus the topology of the binding curves supports a rapid recombination of both ligands to hemoglobin and myoglobin. The rapid NO rebinding to heme ($k_1 \approx 10^8 \text{ s}^{-1}$) includes one-step quartet-doublet transition; since the NO radical has one outer electron at the degenerate π_x and π_y orbitals, a quite similar theory of SOC in quasi-degenerate charge transfer states, like that, presented in Eqs. 8 and 9, can be applied. The only difference is that the SOC integral (9) now includes π_x and π_y orbitals of the NO molecule and thus is slightly smaller (about 60 cm^{-1}). The rate constant of spin transition in the Landau-Zener model is determined by the square of the SOC integral. This explains that the rapid NO rebinding rate constant is about three times slower than the rapid rate constant of the O_2 recombination in heme proteins (Franzen 2002; Jensen and Ryde 2004; Shikama 2006).

All previous analyses of SOC effects in hemoproteins were based on the assumption that the SOC integral in dioxygen binding to heme is determined by the iron ion and no attempt at direct calculation has been done (Blomberg et al. 2005; Franzen 2002; Jensen and Ryde 2004; Sigfridsson and Ryde 2002; Strickland and Harwey 2007). As follows from our simple analysis, the SOC integral (Eq. 9) is determined entirely by SOC in oxygen molecule and is connected with the degeneracy of the two π_{g,x^-} and π_{g,y^-} -orbitals in the open-shell of dioxygen. This enhancement of SOC effect by inclusion of charge-transfer to O_2 and involvement of superoxide-ion structure seems to be quite general in biochemistry (Minaev 2007, 2010); it is applied also to those enzymes that have no transition atoms, like glucose oxidase, or coproporphyrinogen III oxidase (Minaev 2007; Prabhakar et al. 2002; Silva and Ramos 2008). It is important to stress that the rapid T-S transition in O_2 binding to heme is not only forbidden by spin, but also by orbital symmetry (it includes the $A'' - A'$ symmetry change). Such double prohibition is necessary in order to make the spin change in chemical reaction to be effectively allowed (Minaev and Lunell 1993).

The iron ion can also contribute to the SOC integral in the O_2 binding to heme. In general, contributions from different atoms can cancel each other in SOC calculations; it happens when orbital rotation for different atoms occurs in opposite directions along the same axis during the T-S transition (Minaev 2007; Minaev and

Lunell 1993). This type of SOC suppression is known in the form of the destructive external heavy atom effect (Minaev 2010). However, such an effect cannot be applied for SOC analysis in heme-O₂ interaction. The main axis of the third-orbitals rotation in ferrous ion during the T-S transition is perpendicular to the porphyrin tetrapyrrole plane, while the SOC contribution from dioxygen is determined by the $\pi_{g,x}$ - and $\pi_{g,y}$ -orbitals' rotation around the O–O axis and the two axes do not coincide. The angle of 118° between these axes excludes possibility of cancellation: only the sum of the squares of the x,y,z projections, of the type given by Eq. 9, contributes to the final SOC value, which determines the rate constant of the T-S transition in terms of the Landau-Zener approximation (Harvey 2004; Kondo and Yoshizawa 2003; Minaev 2002; Minaev and Ågren 1995; Poli and Harvey 2003).

The Role of Spin in Biochemistry

We can see that spin-dependent quantum effects and transformations are important in a number of biochemical processes (Afanasyeva et al. 2006; Blomberg et al. 2000a, 2005; Buchachenko 1977; Buchachenko and Kouznetsov 2008; Buchachenko et al. 2005; Burnold and Solomon 1999; Chalkias et al. 2008; Davydov et al. 2008; de Winter and Boxer 2003; Franzen 2002; Grissom 1995; Jensen and Ryde 2004; Jensen et al. 2005; Kumar et al. 2005; Lane 2002; Metz and Solomon 2001; Minaev 1983, 2002, 2007, 2010; Minaev et al. 1999, 2007; Orlova et al. 2003; Petrich et al. 1988; Prabhakar et al. 2002, 2003, 2004; Serebrennikov and Minaev 1987; Shaik et al. 1998, 2002, 2005; Sigfridsson and Ryde 2002; Silva and Ramos 2008; Strickland and Harwey 2007). Spin is a very fundamental quantum phenomenon associated with the structure of space-time (Esposito 1999; Hameroff and Penrose 1996; Hu and Wu 2004; Penrose 1960). Various models of elementary particles in modern physics and even space-time itself are built with spinors (Esposito 1999; Penrose 1960). Spin of the Dirac electron is qualitatively shown to be responsible for all known quantum effects and the quantum potential is a pure consequence of “internal motion” evidencing that the quantum behavior is a direct consequence of the fundamental existence of spin (Penrose 1960). These results have been expanded recently by deriving a spin-dependent gauge transformation between the Hamilton-Jacobi equation of classical mechanics and the time-dependent Schrödinger equation of quantum mechanics that is a function of the quantum potential in Bohm mechanics (Esposito 1999). It is quite natural that the most mysterious peculiarities of living matter, the neuron network machinery of brain, hemoglobin, cytochromes, and oxydases, cannot avoid utilizing such a fundamental property of electrons like spin and quantum behavior. The occurrence of non-zero electron spin in any bio-system indicates that we are at a level of the life science where classical concepts are insufficient for its proper description and understanding. Electron spin is much more important than nuclear spin since it determines the exchange interaction, the most significant part of chemical forces, which are finally responsible for the functions of cells and for metabolism processes.

External Magnetic Field Effects in Biochemistry

Regardless of the comments above, we need to speak about nuclear spins, since they are everywhere (at least in the form of protons) and because they also can be connected with quantum effects in bio-systems. Recent experiments (Buchachenko and Kouznetsov 2008; Buchachenko et al. 2005) demonstrate that intramitochondrial nucleotide phosphorylation is a nuclear spin controlled process because the magnetic magnesium isotope $^{25}\text{Mg}(\text{II})$ increases the rate of mitochondrial ATP synthesis in comparison with the spinless nonmagnetic $^{24}\text{Mg}(\text{II})$, $^{26}\text{Mg}(\text{II})$ ions. Such nuclear spin isotope effect is usually interpreted in terms of radical pair theory for separated spins in solvent cage (Buchachenko 1977), but an alternative explanation based on electronic spin-transition in the active center (Minaev 1983; Minaev and Ågren 1995; Serebrennikov and Minaev 1987) is also possible. New developments in magnetic resonance imaging of the brain demonstrate that induced quantum coherences of proton spins separated by long distances ranging from thousands to millions of nanometers (from 1 μm to 1 mm) are sustained for milliseconds and longer (Hu and Wu 2004). While these quantum couplings are not the type of quantum processes that are likely to prove useful in brain function (Hagan et al. 2002), they nonetheless show that biology can take advantage of mesoscopic quantum coherence in clever ways. These quantum modes are not the stable entangled superpositions required for quantum computations in brain, but the proton spins' coherence can indeed survive in DNA, peptides, and in the brain's milieu (Hameroff and Penrose 1996; Hu and Wu 2004). Some new theory of consciousness is based on an idea that spin is the "mind-pixel" (Hagan et al. 2002; Hameroff and Penrose 1996; Hu and Wu 2004); it postulates that consciousness is intrinsically connected to quantum spin, since the latter is the origin of quantum effects in Bohm formalism and a fundamental quantum process associated with the structure of space-time (Esposito 1999; Hagan et al. 2002; Hameroff and Penrose 1996; Hu and Wu 2004; Penrose 1960). Spin involvement in the theory of consciousness can be connected also with memory storage and magnetic perturbations.

External magnetic field effects (MFE) in biochemistry are well known (Afanasyeva et al. 2006; Buchachenko and Kouznetsov 2008; Buchachenko et al. 2005; Chalkias et al. 2008; Davydov et al. 2008; Engstrom et al. 1998, 2000; Grissom 1995; Hoff 1986; Minaev et al. 1999; Serebrennikov and Minaev 1987) and could be analyzed in terms of competition with the internal magnetic interactions including SOC as the most important one (Minaev 1983; Minaev and Ågren 1995; Minaev and Lunell 1993) and hyperfine coupling (Burnold and Solomon 1999; Metz and Solomon 2001; Minaev 2004); the spin-dependent exchange interactions are also involved in the MFE theories as well as the whole spin-catalysis concept (Minaev and Ågren 1996; Orlova et al. 2003). MFE on enzymatic reactions have been first interpreted (Grissom 1995) entirely in the context of RPT (Buchachenko 1977). It was clear since the 1980a that RPT cannot explain all variety of MFE in chemistry and biology and SOC effects could be taken into account (Minaev 1980, 1983). But only recently, a number of publications (Afanasyeva et al. 2006;

Chalkias et al. 2008; Minaev 2002; Minaev and Ågren 1995) examined the external MFE in terms of SOC effects on the kinetics of enzymatic reactions. Activity increase of horseradish peroxidase (HRP) in the presence of magnetic particles (Fe_3O_4) (Chalkias et al. 2008) and external MFE in enzymatic oxidation of NADH by HRP (Afanasyeva et al. 2006) are the most interesting examples. Interconversion of NADH to form $\text{NADH}^{+\bullet}$ and NAD^+ occurs by hydride transfer, in which H^+ ion and two electrons are transferred between the C(4) carbon atom of nicotineamide ring and substrate (Afanasyeva et al. 2006). Catalytic cycle of NADH oxidation in the presence of hydrogen peroxide begins with a two-electron oxidation of the HRP enzyme to form Fe^{IV} and the porphyrin cation radical; this compound I is a highly reactive species that can accept one electron from NADH to form $\text{NADH}^{+\bullet}$ radical cation, which can undergo deprotonation to yield NAD^\bullet and compound II (Afanasyeva et al. 2006; Grissom 1995).

MFE on the photosynthetic light-harvesting reaction center in bacterium *Rhodospseudomonas sphaeroides* was the first successful observation and interpretation of spin-dependent intermediates in biology (Grissom 1995; Hoff 1986). The quantum yield of the triplet chromophore decreases by 50 % at 0.05 T. As the applied magnetic field increases, the Zeeman interaction between $M_S = \pm 1$ spin sublevels grows and causes a decrease of hyperfine coupling between triplet $T(\pm 1)$ and singlet S states, thus only T(0)-S states can interconvert. At high magnetic fields greater than 0.5 T the quantum yield of the triplet chromophore begins to increase slowly and attains parity with the zero-field ($B = 0$) value at $B = 3$ Tesla. At higher fields a net increase is detected, which is explained by T-S transitions via Δg mechanism (Afanasyeva et al. 2006; Buchachenko 1977; Engstrom et al. 2000). Such biphasic MFE dependence on B is typical for many photochemical (Hoff 1986; Serebrennikov and Minaev 1987) and enzymatic (Afanasyeva et al. 2006; Grissom 1995) reactions that occur through the radical pair mechanism. No total MFE on the photosynthetic reaction center would be expected since the total quantum yield is composed of a series of coupled vectorial processes that create an irreversible free energy cascade and drive proton pumping (Grissom 1995; Hoff 1986). For similar reasons it is impossible to expect other MFE with biological relevance. For example, MFE is possible, in principle, for the spin inversion step in GO, if the rotation of reactants in the step shown on Fig. 2b would be frozen. The T-S transitions for different $M_S = \pm 1, 0$ spin sublevels are characterized by different rate constants even in zero field in the fixed molecular frame. But rotation at room temperature quenches all anisotropy of zero-field splitting and SOC-induced rate constants (Serebrennikov and Minaev 1987). That is why no MFE has been observed for glucose oxidase in vivo (Chalkias et al. 2008).

Many mammals, reptiles, birds, insects, and fish are known to use the Earth's magnetic field for orientation and navigation (Gegear et al. 2008; Johnsen and Lohmann 2005; Maeda et al. 2008). Birds have been studied very intensively, but biophysics and neurobiology of their avian magnetoreception are still poorly understood (Grissom 1995; Johnsen and Lohmann 2005). Maeda et al. (2008) have recently proposed a carotenoid-porphyrin model system to demonstrate that the lifetime of a photochemically induced radical pair is changed by MFE with a low

magnetic field of about 50 μT , which is comparable with the Earth's magnetic field. Radiofrequency MFE can disrupt the ability of birds to orient and also has profound influence on radical pair reactions *in vitro*, which supports the RPT concept (Maeda et al. 2008). The recent DFT study of the electronic mechanisms of oxidative cleavage of the O–O bond by apocarotenoid oxygenase (ACO), which occurs in dark condition and includes a number of biradical and radical stages, is important additional support to the carotenoid-porphyrin model of Maeda et al. (2008).

In this context it is interesting to pay attention to molecular oxygen. Each O_2 molecule contains two unpaired electron spins and is a paramagnetic species capable of producing a fluctuating magnetic field along its diffusing pathway. Thus, O_2 serves as a natural contrast agent in magnetic resonance imaging (Hu and Wu 2004). The existence of unpaired electrons in stable molecules is very rare indeed. Dioxygen is the only paramagnetic species found in large quantities in the brain (besides the nitric oxide) and in other aerobic tissues. The singlet excited $^1\Delta_g$ oxygen (Minaev 2007) is a metastable species with a relatively short lifetime (few microseconds in water) (Schweitzer and Schmidt 2003) and long diffusion path (Ogilby 1999); it could serve as a light signal transmitter or transfer other types of information along definite distance in tissue. Generation and quenching of the singlet $^1\Delta_g$ oxygen are governed by SOC and other quantum perturbations (Minaev 1980, 1989, 2007; Minaev and Ågren 1995, 1996; Minaev and Ågren 1997; Minaev and Minaev 2005; Minaev and Minaeva 2001; Minaev et al. 1993, 1996, 1997, 2008; Ogilby 1999; Paterson et al. 2006; Schweitzer and Schmidt 2003). But the ground state triplet $^3\Sigma_g$ oxygen bound and transported by hemoglobin is a source of many other quantum effects besides those described above.

Spin Neuroscience

Conventional neuroscience has been unable to provide a complete understanding of cognitive processes (Esposito 1999; Hagan et al. 2002; Hameroff and Penrose 1996; Hu and Wu 2004; Penrose 1960). It has been accepted that the brain can be simulated as a neural network organized by the principles of classical physics (Hameroff and Penrose 1996; Shaik et al. 2002). Such an approach has delivered successful implementation of learning and memory and promoted optimism that a sufficiently complex artificial neural network would (in principle, at least) reproduce the full extent of brain processes involved in cognition and consciousness (Hagan et al. 2002). But the functioning of the nervous system lie outside the realm of classical mechanics; one finds ample support for this in analysis of sensory organs of vision, the operation of which is quantized at levels varying from the reception of individual photons by the retina to thousands of quanta in the auditory system (Hagan et al. 2002; Hameroff and Penrose 1996). It is also proposed that synaptic signal transmission has a quantum character (Hagan et al. 2002). The Penrose-Hameroff orchestrated objective reduction (OOR) model (Hameroff and Penrose 1996) of quantum computation in microtubules within neurons of the brain is compatible with modern physiology and can generate testable predictions. Coherent

superpositions of tubulin proteins are unstable in the OOR model and subject to self-collapse under quantum criterion. The phase of quantum superposition/computation is a preconscious process and each self-collapse event corresponds to “moment of conscious experience.” An apparent shortcoming of the Penrose-Hameroff OOR model and other models of quantum processes relevant to consciousness is the question of environmental decoherence (Hagan et al. 2002).

The dioxygen utilization in living matter provides many interesting manifestations of spin-dependent quantum effects that are under a control by very weak relativistic perturbations (Minaev 2007). This illustrates a long-lived quantum coherence in O₂ binding to hemoglobin (Franzen 2002; Friedman and Campbell 1987; Shikama 2006). Spin coherence can occur in photosynthetic systems, in oxidases, and in cytochromes. It depends on hyperfine interactions and spin–spin coupling in the ensembles of paramagnetic complexes. The decoherence can be induced by spin-flip and this review illustrates how to calculate and interpret such phenomena. It seems promising that the SOC-induced spin effects connected with dioxygen activation in bio-systems could be useful in other areas of life science.

Conclusions

In this review we consider spin effects in a number of biochemical processes with special attention to O₂ activation by heme, oxidases, and cytochromes. The universal role of the spin-orbit coupling (SOC) effect in such activations is stressed. Oxidation of organic materials by dioxygen, to give CO₂ and water, is very favorable (exothermic); fortunately, the unfavorable kinetics preclude this spontaneous combustion of living matter into “a puff of smoke” (Sheldon 1993). The reason for the sluggish O₂ reactivity is a spin barrier: the direct reaction of triplet dioxygen with singlet organic molecules to give stable diamagnetic products is a spin-forbidden process with a very low rate. The common way of circumventing this kinetic barrier via a free radical pathway (combustion) is impossible for living cells. The reaction of singlet (diamagnetic) molecules with ³O₂ forming two radicals is a spin-allowed process. Usually it is highly endothermic; Sheldon provides as an example of such a reaction observed at moderate temperature the activation of ³O₂ by flavin-dependent oxygenases (Sheldon 1993). He argues that it is possible only for very reactive substrates (flavins) that form resonance-stabilized free radicals, e.g., reduced flavins, like FADH^{•+}, and HO₂[•] as a counterpart. We have stressed that, even in this case, the complete oxidation in the enzyme (e.g., glucose oxidase) is spin-forbidden if the HO₂[•] radical is not the final product that leaves the active center of enzyme and goes into the bulk of the cell, but still participates in further oxidative transformations (Minaev 2002). We have shown that flavin-dependent oxygenases provide a very efficient way to overcome the spin-prohibition by spin-orbit coupling perturbation at the stage of electron transfer (Eq. 4).

A second way to overcome the obstacle of spin conservation is for ³O₂ to combine with a paramagnetic transition metal ion (Burnold and Solomon 1999; Metz and Solomon 2001; Prabhakar et al. 2003, 2004). Such reactions are spin-

allowed and governed by exchange interactions (exchange-induced spin-catalysis) (Minaev 2007; Prabhakar et al. 2004). Peroxo-, dioxo-, and superoxo-complexes with metals in different oxidation degree are considered during the search of “dream reactions” for selective industrial catalysts (Sheldon 1993). The expectation that the resulting metal-dioxygen complexes may react selectively with organic substrates at moderate temperature forms a background for the extensive studies of metal-catalyzed oxidation during the last four decades (Prabhakar et al. 2002, 2004). In this respect the reactions of $^3\text{O}_2$ with hemoglobin and myoglobin at moderate temperature are very peculiar, since they are still spin-forbidden and quite efficient (Franzen 2002; Jensen and Ryde 2004; Sigfridsson and Ryde 2002).

We have recalculated some potential energy surface cross-sections for different multiplets along the heme- O_2 binding reaction coordinate and obtained results in agreement with Refs. Sigfridsson and Ryde (2002), Franzen (2002), and Jensen and Ryde (2004). The Fe(II)porphine molecule coordinated with ammonia molecules (Fig. 3, NH_3 as a model of the proximal histidine) are used as in other similar studies (Blomberg et al. 2005; Franzen 2002; Jensen et al. 2005; Sigfridsson and Ryde 2002; Strickland and Harwey 2007). The more realistic protoporphyrin IX model provides similar result for the ground state of the heme active site. Results of previous works (Blomberg et al. 2005; Franzen 2002; Jensen and Ryde 2004; Minaev 2007; Minaev et al. 2007; Sigfridsson and Ryde 2002; Strickland and Harwey 2007) indicate that the main reason for the facilitated binding of O_2 to heme is a broad crossing region of the relevant spin states, which provides significant spin transition probabilities. They have shown that porphyrin is an ideal Fe(II) ligand for the spin-flip problem, because it tunes the spin states to be close in energy, giving parallel binding potential energy surfaces, small activation energies and large transition probabilities in terms of the Landau-Zener approach (Prabhakar et al. 2003). But none of these studies (Blomberg et al. 2005; Franzen 2002; Jensen and Ryde 2004; Jensen et al. 2005; Sigfridsson and Ryde 2002; Strickland and Harwey 2007) have considered the reason for relatively large spin-orbit coupling, which induces the necessary spin flip in the heme- O_2 binding reaction; a general assumption that the SOC integral at Fe(II) ion of about 80 cm^{-1} , postulated in Ref. Petrich et al. (1988), has been put forward instead. We have shown that such SOC integral (Eq. 9) is determined entirely by SOC in the oxygen moiety and is connected with the two $\pi_{g,x}$ - and $\pi_{g,y}$ -orbitals in the open-shell of dioxygen. This is also connected with charge transfer (CT) and with the $\text{Fe}^{3+} - \text{O}_2^-$ radical-pair structure of the ground state $^1\text{A}'$ (1) and the close-lying $^3\text{A}''$ (2) state near the crossing of the potential energy surfaces, see scheme (Eq. 8). This scheme indicates that the triplet and singlet states, $^3\text{A}''$ (2) and $^1\text{A}'$ (1), differ by a single electron jump inside O_2 from the $\pi_{g,x}$ MO to the $\pi_{g,y}$ orbital. Such transformation is equivalent to the electronic orbital rotation, i.e., a torque, which creates transient magnetic field during the T-S transition and this magnetic field is responsible for the spin flip. In this model the magnetic perturbation occurs entirely in the oxygen moiety and the iron ion is silent. Account of SOC in the third-shell of the metal has to increase the total SOC integral.

From a broader perspective, the non-zero electron spin and the concepts of quantum mechanics play a fundamental role in our understanding of the mystery of

life. Spin of dioxygen is a property of the ground state of the molecule and protects O₂ from its involvement in the realm of ordinary chemistry, where the brute force of activation energy just governs all spin-allowed biochemical transformations of diamagnetic species. Dioxygen is protected from brute force via the $X^3\Sigma_g^- - a^1\Delta_g$ energy gap (22 kcal/mol) (Hagan et al. 2002) and via spin-prohibition; at the same time O₂ is extremely fragile and easily activated by the presence of small amount of radicals or by magnetic perturbations (Minaev 2007). Such quantum spin protection is very important for biological systems, which operate at room temperature and are extraordinary complex, diverse, noisy, and “wet.” The quantum spin protection of dioxygen reactivity and its dependence on weak internal and external magnetic perturbations is important in connection with a fundamental fact, namely that biological systems are open driven systems. Self-organization in such biosystems should involve additional and very common mechanisms of self control, like S-T transitions in the O₂ activating enzymes, cytochromes, heme, photosynthetic centers, and probably in neural networks.

A fundamental conclusion of this review is an assertion of the fact that the functioning of various enzymes is controlled by spin conversion determined by internal and external magnetic forces. The most important magnetic interaction is spin-orbit coupling and we indicate a simple physical origin of this perturbation as responsible for the functionality of enzymes like glucose oxidase, coproporphyrinogen oxidase, luciferase, and myoglobin, hemoglobin, and other hemoproteins.

Acknowledgments This work is supported by the State Foundation of Fundamental Investigations (DFFD) of Ukraine, the project F25.5/008, and by Visby project No = 01403/2007.

Bibliography

- Afanasyeva, M. A., Taraban, M. B., Purtov, P. A., Leshina, T. V., & Grissom, C. B. (2006). Magnetic field effects in enzymatic reactions: Radical oxidation of NADH by horseradish peroxidase. *Journal of the American Chemical Society*, *128*, 8651.
- Blomberg, M. R. A., Siegbahn, P. E. M., Babcock, G. T., & Wikstrom, M. (2000a). Modeling cytochrome oxidase: A quantum chemical study of the O–O bond cleavage mechanism. *Journal of the American Chemical Society*, *122*, 12848.
- Blomberg, M. R. A., Siegbahn, P. E. M., Babcock, G. T., & Wikstrom, M. (2000b). O–O bond splitting mechanism in cytochrome oxidase. *Journal of Inorganic Biochemistry*, *80*, 261.
- Blomberg, L. M., Blomberg, M. R. A., & Siegbahn, P. E. M. (2005). A theoretical study of the binding of O₂, NO and CO to heme proteins. *Journal of Inorganic Biochemistry*, *99*, 949.
- Buchachenko, A. L. (1977). Enrichment of magnetic isotopes—new method of investigation of chemical reaction mechanisms. *Russian Journal of Physical Chemistry*, *51*, 2461.
- Buchachenko, A. L., & Kouznetsov, D. A. (2008). Magnetic field affects enzymatic ATP synthesis. *Journal of the American Chemical Society*, *130*, 12868.
- Buchachenko, A. L., Kouznetsov, D. A., Orlova, M. A., & Markarian, A. A. (2005). Spin biochemistry: Intramitochondrial nucleotide phosphorylation is a magnesium nuclear spin controlled process. *Mitochondrion*, *5*, 67.
- Burnold, T. C., & Solomon, E. I. (1999). Reversible dioxygen binding to hemerythrin. *Journal of the American Chemical Society*, *121*, 8288.

- Chalkias, N. G., Kahawong, P., & Giannelis, E. P. (2008). Activity increase of horseradish peroxidase in the presence of magnetic particles. *Journal of the American Chemical Society*, *130*, 2910.
- Davydov, R., Osborne, R. L., Kim, S. H., Dawson, J. H., & Hoffman, B. M. (2008). EPR and ENDOR studies of cryoreduced compound I. *Biochemistry*, *47*, 5147.
- de Winter, A., & Boxer, S. G. (2003). Energetics of primary charge separation in bacterial photosynthetic reaction center mutants: Triplet decay in large magnetic field. *Journal of Physical Chemistry A*, *107*, 3341.
- Engstrom, M., Minaev, B. F., Vahtras, O., & Agren, H. (1998). MCSCF linear response calculations of electronic g-factor and spin-rotational coupling constants for diatomics. *Chemical Physics Letters*, *237*, 149.
- Engstrom, M., Himo, F., Graslund, A., Minaev, B. F., Vahtras, O., & Ågren, H. (2000). H-bonding to the tyrosyl radical analyzed by ab initio g-tensor calculations. *Journal of Physical Chemistry A*, *104*, 5149.
- Esposito, S. (1999). On the role of spin in quantum mechanics. *Foundations of Physics Letters*, *12*, 165.
- Franzen, S. (2002). Spin-dependent mechanism for diatomic ligand binding to heme. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 16754.
- Friedman, J., & Campbell, B. (1987). *Structural dynamics and reactivity in hemoglobin*. New York: Springer.
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., et al. (2003). *Gaussian 03, Revision B. 03*. Pittsburg: Gaussian Inc.
- Gegeer, R. J., Conelman, A., & Waddell, S. (2008). Cryptochrome mediates light-dependent magnetosensitivity in *Drosophila*. *Nature*, *454*, 1014.
- Grissom, C. B. (1995). Magnetic field effects in biology: A survey of possible mechanisms with emphasis on radical-pair recombination. *Chemical Reviews*, *95*, 3.
- Hagan, S., Hameroff, S. R., & Tuszynski, J. A. (2002). Quantum computation in brain microtubules. *Physical Review E*, *65*, 061901.
- Hameroff, S., & Penrose, R. (1996). Conscious events as orchestrated space-time selections. *Journal of Consciousness Studies*, *3*, 36.
- Harvey, J. N. (2004). Spin-forbidden CO ligand recombination in myoglobin. *Faraday Discussions*, *127*, 165.
- Hoff, A. J. (1986). Magnetic interactions between photosynthetic reactants. *Photochemistry and Photobiology*, *43*, 727.
- Hu, H. P., & Wu, M. X. (2004). Spin-mediated consciousness theory. *Medical Hypotheses*, *63*, 633.
- Jensen, K. J., & Ryde, U. (2004). How O₂ binds to heme. *Journal of Biological Chemistry*, *279*, 14561.
- Jensen, K. J., Ross, B. O., & Ryde, U. (2005). The CAS PT2 study of oxymyoglobin model. *Journal of Inorganic Biochemistry*, *99*, 45.
- Johnsen, S., & Lohmann, K. J. (2005). The physics and neurobiology of magnetoreception. *Nature Reviews Neuroscience*, *6*, 703.
- Klinman, J. P. (2001). Life as aerobes: Are there simple rules for activation of dioxygen by enzymes? *Journal of Biological Inorganic Chemistry*, *6*, 1.
- Kondo, M., & Yoshizawa, K. (2003). A theoretical study of spin-orbit coupling in an Fe(II) spin-crossover complex. Mechanism of the LIESST effect. *Chemical Physics Letters*, *372*, 519.
- Kumar, D., de Viser, S. P., Sherma, P. K., Hirao, H., & Shaik, S. (2005). Sulfoxidation mechanisms catalyzed by cytochrome P450 and HRP models: Spin selection induced by the ligand. *Biochemistry*, *44*, 8148.
- Lane, N. (2002). *Oxygen: The molecule that made the world*. Oxford: Oxford University Press.
- Maeda, K., Henbest, K. B., Cintolesi, F., Kuprov, I., Rodgers, C. T., Liddell, P. A., Gust, D., Timmel, C. R., & Hore, P. J. (2008). Chemical compass model of avian magnetoreception. *Nature*, *453*, 387.
- Metz, M., & Solomon, E. I. (2001). Dioxygen binding to deoxyhemocyanin. *Journal of the American Chemical Society*, *123*, 4938.

- Minaev, B. F. (1980). Intensities of spin-forbidden transitions in molecular oxygen. *International Journal of Quantum Chemistry*, 89, 367.
- Minaev, B. F. (1983). Theoretical analysis and prognostication of spin-orbit coupling effects in molecular spectroscopy and chemical kinetics. DrSc Dissertation, Institute of Chemical Physics, Moscow.
- Minaev, B. F. (1989). Solvent-induced emission of molecular a $^1\Delta_g$ oxygen. *Journal of Molecular Structure (Theochem)*, 183, 207.
- Minaev, B. F. (2002). Spin effects in reductive activation of O₂ by oxidase enzymes. *RIKEN Review*, 44, 147.
- Minaev, B. F. (2004). Ab initio study of the ground state properties of molecular oxygen. *Spectrochimica Acta Part A-Molecular and Biomolecular Spectroscopy*, 60, 1027.
- Minaev, B. F. (2007). Electronic mechanisms of molecular oxygen activation. *Russian Chemical Review*, 76, 1039.
- Minaev, B. F. (2010). Environment friendly spin-catalysis for dioxygen activation. *Chemistry & Chemical Technology*, 4, 1–16.
- Minaev, B. F., & Ågren, H. (1995). Spin-orbit coupling induced chemical reactivity and spin-catalysis phenomena. *Collection of Czechoslovak Chemical Communications*, 60, 339.
- Minaev, B. F., & Ågren, H. (1996). Spin-catalysis phenomena. *International Journal of Quantum Chemistry*, 57, 510.
- Minaev, B. F., & Ågren, H. (1997). Collision-induced b $^1\Sigma_g^-$ -a $^1\Delta_g$, X $^3\Sigma_g^-$ -b $^1\Sigma_g^-$, and X $^3\Sigma_g^-$ -a $^1\Delta_g$ transition probabilities in molecular oxygen. *Journal of the Chemical Society-Faraday Transactions*, 93, 2231.
- Minaev, B. F., & Lunell, S. (1993). Classification of spin-orbit coupling effects in organic chemical reactions. *Zeitschrift Fur Physikalische Chemie-International Journal of Research in Physical Chemistry & Chemical Physics*, 182, 263.
- Minaev, B. F., & Minaev, A. B. (2005). Calculation of the phosphorescence of porphyrins by the density functional method. *Optics and Spectroscopy*, 98, 214.
- Minaev, B. F., & Minaeva, V. A. (2001). MCSCF response calculations of the excited states properties of the O₂ molecule and a part of its spectrum. *Physical Chemistry Chemical Physics*, 3, 720.
- Minaev, B. F., Lunell, S., & Kobzev, G. I. (1993). The influence of intermolecular interaction on the forbidden near-IR transitions in molecular oxygen. *Journal of Molecular Structure (Theochem)*, 284, 1.
- Minaev, B. F., Vahtras, O., & Ågren, H. (1996). Magnetic phosphorescence of molecular oxygen. *Chemical Physics*, 208, 299.
- Minaev, B. F., Mikkelsen, K. V., & Ågren, H. (1997). Collision-induced electronic transitions in complexes between benzene and molecular oxygen. *Chemical Physics*, 220, 79.
- Minaev, B. F., Lyzhenkova, I. I., Minaeva, V. A., & Boiko, V. I. (1999). A quantum chemical approach to the mechanism of biochemical action of nicotinamide. *Theoretical and Experimental Chemistry*, 35, 258.
- Minaev, B. F., Minaeva, V. A., & Vasenko, O. M. (2007). Spin states of the Fe(II)-pporphin molecule: Quantum-chemical study by DFT method. *Ukrainica Bioorganica Acta*, 1, 24.
- Minaev, B. F., Minaeva, V. A., & Evtuhov, Y. V. (2008). Quantum-chemical study of the singlet oxygen emission. *International Journal of Quantum Chemistry*, 108, 500.
- Ogilby, P. R. (1999). Singlet oxygen. *Accounts of Chemical Research*, 32, 512.
- Orlova, G., Goddard, J. D., & Brovko, L. Y. (2003). Theoretical study of the amazing firefly bioluminescence. *Journal of the American Chemical Society*, 125, 6962.
- Paterson, M. J., Christiansen, O., Jensen, F., & Ogilby, P. R. (2006). Singlet oxygen. *Photochemistry and Photobiology*, 82, 1136.
- Penrose, R. (1960). A spinor approach to general relativity. *Annals of Physics*, 10, 171.
- Penrose, R. (1994). *Shadows of the mind*. Oxford: Oxford University Press.
- Petrich, J. W., Poyart, C., & Martin, J. L. (1988). Spin-forbidden binding of O₂ to hemoglobin. *Biochemistry*, 27, 4049.

- Poli, R., & Harvey, J. N. (2003). Spin-forbidden chemical reactions. *Chemical Society Reviews*, 32, 1.
- Potter, W. T., Ticker, M. P., & Caughey, W. S. (1987). Resonance Raman spectra of myoglobin. *Biochemistry*, 26, 4699.
- Prabhakar, R., Siegbahn, P. E. M., Minaev, B. F., & Ågren, H. (2002). Activation of triplet dioxygen by glucose oxidase: Spin-orbit coupling in the superoxide ion. *Journal of Physical Chemistry B*, 106, 3742.
- Prabhakar, R., Siegbahn, P. E. M., & Minaev, B. F. (2003). A theoretical study of the dioxygen activation by glucose oxidase and by copper amine oxidase. *Biochimica et Biophysica Acta*, 1647, 173.
- Prabhakar, R., Siegbahn, P. E. M., Minaev, B. F., & Agren, H. (2004). Spin transition during H₂O₂ formation in the oxidative half-reaction of copper amine oxidase. *Journal of Physical Chemistry B*, 108, 13882.
- Sawyer, D. T. (1991). *Oxygen chemistry*. New York: Oxford University Press.
- Schweitzer, C., & Schmidt, R. (2003). Physical mechanisms of generation and deactivation of singlet oxygen. *Chemical Reviews*, 103, 1685.
- Serebrennikov, Y. A., & Minaev, B. F. (1987). Magnetic field effects due to spin-orbit coupling in transient intermediates. *Chemical Physics*, 114, 359.
- Shaik, S., Filatov, M., Schroder, D., & Schwarz, H. (1998). Electronic structure makes a difference: Cytochrome P-450 mediated hydroxylation of hydrocarbons as a two-state reactivity paradigm. *Chemistry—A European Journal*, 4, 193.
- Shaik, S., Ogliaro, F., de Visser, S. P., Schwarz, H., & Schroeder, D. (2002). Two state reactivity mechanism of hydroxylation and epoxidation by cytochrome P450 revealed by theory. *Current Opinion in Chemical Biology*, 6, 556.
- Shaik, S., Kumar, D., de Viser, S. P., Altun, A., & Tiel, W. (2005). Theoretical perspective on the structure of cytochrome P450 enzymes. *Chemical Reviews*, 105, 2279.
- Sheldon, R. A. (1993). A history of oxygen activation. In D. Barton et al. (Eds.), *The Activation of dioxygen and homogeneous catalytic oxidation*. New York: Plenum.
- Shikama, K. (2006). Nature of the FeO₂ bonding in myoglobin and hemoglobin. A new molecular paradigm. *Progress in Biophysics and Molecular Biology*, 91, 83.
- Sigfridsson, E., & Ryde, U. (2002). Theoretical study of discrimination between O₂ and CO by myoglobin. *Journal of Inorganic Chemistry*, 91, 101.
- Silva, P. J., & Ramos, M. J. (2008). A comparative DFT study of the reaction mechanism of the O₂-dependent coproporphyrinogen III oxidase. *Bioorganic Medical Chemistry*, 16, 2726.
- Strickland, N., & Harvey, J. N. (2007). Spin-forbidden ligand binding to the ferrous-heme group. *Journal of Physical Chemistry B*, 111, 841.
- Stryer, L. (1995). *Biochemistry* (4th ed., p. 152). New York: Freeman.

G. Náray-Szabó, A. Perczel, A. Láng, and D. K. Menyhárd

Contents

Introduction	1590
Structure Determination	1590
Experiment	1591
Computer Modeling	1596
Structure Representation	1600
Molecular Graphics	1600
Electrostatics	1602
Solvent-Accessible Surface	1604
Dynamics	1605
Time Scales	1605
Dynamical Structures	1608
Dynamics of Interactions	1608
Interactions	1609
Protein Hydration	1609
Ligand Binding	1610
Protein-Protein Interactions	1613
Enzyme Mechanisms	1617
Conclusion and Outlook	1621
Bibliography	1621

Abstract

Proteins play a crucial role in biological processes; therefore, understanding their structure and function is very important. In this chapter, we give an overview on computer models of proteins. First, we treat both major experimental

G. Náray-Szabó (✉) • A. Perczel • A. Láng • D. Menyhárd

Laboratory of Structural Chemistry and Biology, Institute of Chemistry, Eötvös Loránd

University and MTA-ELTE Protein Modelling Research Group, Budapest, Hungary

e-mail: naraysza@chem.elte.hu; perczel@chem.elte.hu; langax@chem.elte.hu;

dmenyhard@chem.elte.hu

structure determination methods, X-ray diffraction and NMR spectroscopy. In subsequent sections, computer modeling techniques as well as their application to the construction of explicit models are discussed. An overview on molecular mechanics and structure prediction is followed by an overview of molecular graphics methods of structure representation. Protein electrostatics and the concept of the solvent-accessible surface are treated in detail. We devote a special section to dynamics, where time scales of molecular motions, structures, and interactions are discussed. Protein in relation to its surroundings is especially important, so protein hydration, ligand binding, and protein-protein interactions receive special attention. The case study of podocin provides an example for the successful application of molecular dynamics to a complex issue. At last, computer modeling of enzyme mechanisms is discussed. It is demonstrated that protein representation by computers arrived to a very high degree of sophistication and reliability; therefore, even lots of experimental studies make use of such models. A list with a large number of up-to-date bibliographic references helps the reader to get informed on further details.

Introduction

In order to obtain versatile models, we attribute a definite spatial structure and shape to molecules, which may, however, change in time. In case of proteins, the structure is too complicated and complex; thus, the use of computer graphics is unavoidable. In this chapter, we give an overview on the experimental and computational methods, which provide information on protein structure and allow deriving its relation to function. Neither of the techniques currently available is able to yield full information. X-ray diffraction works only for crystalline samples, NMR techniques refer to restraints, which determine a manifold of related but nonidentical structures, while modeling still lacks full reliability. Accordingly, in order to obtain thorough information on protein structure, various techniques should be combined, although in several cases even a single technique provides a model, which is adequate for some considerations. The present chapter may contribute to an overview on up-to-date techniques with reference to more detailed reviews on specific subjects, like the book published recently (Náray-Szabó [2014](#)).

Structure Determination

Both experimental and modeling techniques are available for the determination of the three-dimensional structure of proteins. They vary in performance and neither of them offers a unique tool to be applied for any type of proteins and their complexes. In most cases, publications refer to a single method only, structures are determined either by X-ray diffraction, nuclear magnetic resonance (NMR) spectroscopy or, eventually, by computer modeling. In order to get a more realistic, broad and rather complete information on structural and dynamic properties of proteins, a comparison of data obtained by a variety of methods is required.

Experiment

The first protein structures were determined by X-ray diffraction techniques in the late 1950s, and due to intensive work 95,375 structures are available now in the Protein Data Bank (as of 2015). Structure determination by NMR spectroscopy is also rapidly expanding; the number of deposited structures is 10,789 (Protein Data Bank 2015).

Protein Crystallography

Protein structures can be visualized at the atomic level by X-ray crystallography. An appropriate crystal has to be irradiated and the obtained diffraction pattern is recorded and analysed by mathematical tools. The result is an electron density map with relative maxima at the position of atomic nuclei. Using classical chemical information, the density map can be evaluated and the three-dimensional molecular model of the protein can be constructed. In order to determine the structure of a protein, the following steps have to be performed (Brandén and Tooze 1999; Drenth 2007).

Sample preparation and crystallization. Its goal is to produce a well-ordered crystal that is large enough to provide an appropriate diffraction pattern when irradiated with X-rays. It is inherently difficult because of the fragile nature of protein crystals. Proteins have irregularly shaped surfaces, which form large channels within the crystal. The noncovalent bonds that hold together the lattice must often be formed by flexible amino-acid residues of the protein surface. The success of protein crystallization depends on several factors, like purity, pH, concentration, temperature, precipitants, or additives. In order to ensure sufficient homogeneity, the protein should be pure to at least 97 %. It is necessary to screen a reasonable number of conditions, for that at least 200 μ l of material in a concentration of about 10 mg/ml is needed. The most commonly used techniques for crystallization are the hanging- and sitting-drop methods (Horsefield and Neutze 2006), although alternative approaches also have emerged, e.g., the microbatch method (Luft et al. 2003; Merlino et al. 2011). Efficient high-throughput methods have been developed to speed up the process and to determine optimal crystallization conditions (Chayen 2007).

Data collection. Having an appropriate crystal, its symmetry, the unit cell parameters and the resolution limit have to be determined. Then, it will be irradiated by an X-ray beam and continuously rotated by a small angle. Thus, the X-ray diffraction pattern containing relevant information on the position of scattering centers can be recorded. For crystals of lower quality, application of synchrotron radiation is recommended, where the beam intensity is greater; therefore, data collection times are shorter, sometimes less than 10 min. Synchrotron radiation is also the basis of time-resolved crystallography (Bourgeois and Royant 2005; Petsko and Ringe 2015). The relatively slow reaction rates allow following changes in atomic positions during the enzymatic process.

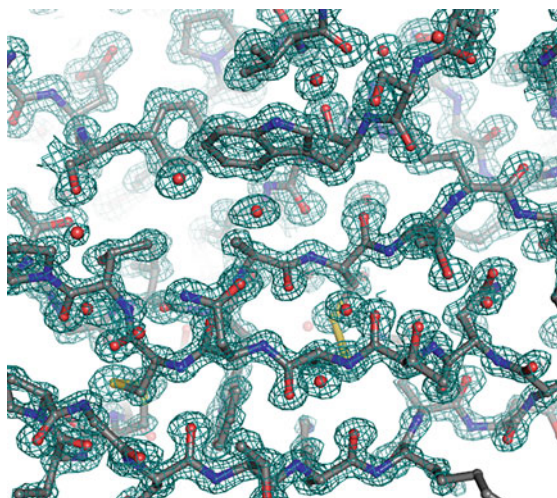
Structure determination. The relationship between the scattering angle and the distance between planes passing through the atoms in the crystal is given by

the Bragg's law. The electron density is related to the diffraction pattern by a mathematical function termed inverse Fourier transform. In order to compute this, both the amplitude and the phase of the diffracted waves should be known. Since experimental measurement of the relative phase angles is impossible, we face the so-called phase problem, often the most serious bottleneck in determining a new structure. A possible technique to solve this issue is the isomorphous replacement, i.e., to provide a crystal, which is nearly identical to the one studied, except that a few atoms have been replaced or added (e.g., using selenomethionine instead of methionine). If these atoms have a large atomic number, they will strongly perturb the diffraction pattern and allow deducing possible values for the phase angles. A similar technique applies the multiple-wavelength anomalous dispersion (MAD). By changing the wavelength of the X-rays, the degree to which the anomalously scattering atoms perturb the diffraction pattern also changes. This gives the same kind of information as isomorphous replacement.

Model building. If the studied crystals were perfectly ordered, all the atoms would scatter in phase and the electron density map would have peaks at each of the atomic positions. However, in electron density maps of most proteins, atoms are not resolved from each another; we need a model to be fitted to the electron density. Fitting models to density requires the use of computer graphics programs such as COOT (Emsley and Cowtan 2004) and as a result the electron density map is interpreted in terms of a set of atoms. In most cases, the protein backbone is fitted first then the side-chain atoms are positioned if the resolution permits. The result is dependent on the resolution and the quality of the phases (Fig. 1). Often regions of high flexibility (e.g., surface lysine side chains) are not visible since the corresponding electron density is smeared.

Refinement and validation. An atomic model can be improved by refinement, in which it is adjusted to the measured diffraction data. Refinement improves the

Fig. 1 Atomic (1.2 Å) resolution electron density map of a section of the complex between crayfish trypsin and an inhibitor from the desert locust *Schistocerca gregaria* (SGTI) (V. Harmat, unpublished)



phases leading to clearer maps and better models. The quality of an atomic model can be judged by the so-called *R*-factor, which is the average fractional error in the calculated amplitude compared to the observed one. Since, in general, there are insufficient diffraction data, these may be completed by restraints on geometry, which keep the bond lengths, angles, and close contacts within a reasonable range. Main-chain torsion angles are difficult to restrain but the distribution of these angles is strongly restricted as shown by the Ramachandran plot (Hovmöller et al. 2002), which is thus an additional indicator of the quality of a structure.

Neutron diffraction provides special information on proteins, a special method requiring high thermal-neutron fluxes, which can be obtained only from nuclear reactors (Fitter et al. 2006; Dambrot 2012). Hydrogen atoms can be precisely located, which is almost impossible by X-ray diffraction. A diffraction experiment can be performed on a crystal; the results can be evaluated similarly as done for the X-ray technique.

Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a versatile tool of modern structural biology, which allows determining protein structures even in some cases, where X-ray diffraction fails. It has become an independent method of routine structure determination for proteins with up to about 300 residues. Structure determination by NMR consists of multiple consecutive steps.

Sample preparation and stability testing. For successful NMR experiments, a protein solution of high purity (>95 %), stability (over a week), and appropriate concentration (0.1–1 mM) is needed; the total sample volume and mass should be 350–550 μ l and 3–30 mg, respectively. 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) rather than tetramethylsilane (TMS) is used as a reference compound. Although both natural and synthetic polypeptides and proteins can be used, most samples are expressed via a suitable prokaryotic or eukaryotic cellular or cell-free recombinant technique. The latter method eliminates most toxic effects attributed to the overproduction of recombinant proteins. Biotechnological approaches have the common advantage of large-scale protein production, specific or nonspecific stable isotope labeling, and easy sequence variation.

Data acquisition (Schorn and Taylor 2004). Some magnetically active nuclei (e.g., ^1H , ^{13}C , or ^{15}N) experiencing a large magnetic field (10–20 T) possess two energy states. The energy difference between these states is large enough to be detectable if the resonance criterion is met during a radiofrequency irradiation (at or very near to the Larmor-frequency of the nucleus in question) as long as relaxation mechanisms are long enough. For example, the resonance frequencies, mostly between 500 and 950 MHz for ^1H nuclei, referring to the energy of the absorption are proportional to the strength of the static magnetic field. Depending on the local chemical environment, nuclei of the same type in a molecule resonate at slightly different frequencies. The relative shift from a reference (DSS for proteins in aqueous solution) resonance frequency is called the chemical shift. Peptides and proteins from natural sources are typically restricted to utilizing resonances based solely on protons; however, to get more information and/or to decrease spectral

crowding/overlap and hence to decrease assignment ambiguity, further NMR active nuclei have to be introduced or enriched in the molecule (mostly ^{13}C and ^{15}N).

In ideal signal dispersion, each magnetically active atom has its unique chemical shift by which it can be recognized and assigned. Assignment of each measured chemical shift to a single atom (or a group of atoms, e.g., a methyl group) is a mandatory step toward structure determination. Resonance frequency assignment is a very complicated and laborious procedure and often needs to introduce further “dimensions” into measurements but provides useful information on protein structure, dynamics, interaction, and ligand binding. Beside the most common direct dimension, typically associated with ^1H , experiments exploiting other nuclei (vide supra) are run to generate additional dimensions to minimize spectral overlap. For structure determination, two major types of NMR experiments are in everyday use: (i) the one based on coherence transfer between chemically connected atoms via indirect (through bond or J) coupling (e.g., COSY, TOCSY, HSQC, HNCQ, HNCA, ...) and (ii) those relying on magnetization transfer between spatially close atoms (through space or direct coupling); see below. The first set is for the identification of connected chemical shifts, while the second one is to localize spatial proximity. In both cases, it is the electron that mediates these interactions.

Resonance frequency assignment. The basis of the analysis of NMR-spectra is the distinction of real and artificial (background) signals; in other words, a large signal-to-noise ratio is needed, which is met if using concentrated samples. Then, one has to find out which chemical shift corresponds to which nuclei within the macromolecule. Several different protocols were invented to achieve this tedious goal, primarily depending on the type of isotope labeling introduced for proteins. Homonuclear techniques are used for unlabeled typically smaller molecules by taking advantage of a series of specific experiments. Resonance assignment for proteins of medium size ($5 < \text{MW} < 15$ kDa) relies on ^{15}N -labeled samples by using suitable parameterized methods. Larger proteins require double (^{15}N -, ^{13}C) or triple (^2H -, ^{15}N -, ^{13}C) labeling to achieve resonance assignment for both backbone and side-chain atoms. In principle, by filtering peaks, grouping of resonances, identifying spin-systems, and by grouping stretches of sequences into segments, the full assignment procedure can be automated. A sophisticated software, like DYNASSIGN (Schmucki et al. 2009), ensures high completeness of the assignment procedure.

Restraint collection. Structure determination by NMR-spectroscopy depends on various structural restraints. Distance restraints are typically obtained from NOESY experiments, where cross peaks appear due to successful magnetization transfer events between spatially close nuclei. Once assigned, each of the NOESY cross peaks can be converted into the approximate internuclear distance. The success of a structure determination depends on several factors, such as the total number of Nuclear Overhauser Effect (NOE) defined hydrogen-hydrogen distance restraints, on their distribution within the molecule, accurate assignment, and on their proper calibration. Because of various limitations, these distances cannot be determined precisely; rather they are sorted as short (1.8–2.7 Å), medium (1.8–3.5 Å), or long (1.8–6.0 Å) ones. Restraints on the backbone (φ and ψ) and side-chain (χ_1 , χ_2 , etc.) dihedral torsion angles can be obtained by measuring conformation sensitive

indirect coupling constants. Various 2D and 3D methods have been developed for the accurate measurement of coupling constants by NMR both in solution and in solid state. By suitable measurements followed by careful calibrations, these data can be converted into dihedral angle restraints. Orientation NMR restraints are commonly obtained for macromolecules. Dipolar couplings are averaged in solution because of fast molecular tumbling. However, due to slight overpopulation of one state, residual dipolar coupling (RDC) can be measured even in solution by using suitable oriented media to restrict overall molecular tumbling. Dipolar coupling is routinely used in solid state NMR and provides useful information about the relative orientation of bond vectors. Hydrogen bonds as structural restraints are typically determined by hydrogen/deuterium exchange experiment. H/D exchange can be quantitatively monitored by the isotope effect of deuterium on the chemical bond strength. The slower the NH exchanges to ND, the more the amide group is buried within the protein. Temperature dependence of NH chemical shifts (termed as temperature coefficients) can also be detected and used to estimate the involvement of the NH bond in an intramolecular H-bond, especially for peptides and shorter proteins. Various paramagnetic restraints can also be used for solution structure determination. Finally, order parameters introduced as restraints reflect an advanced view on the dynamic structure of proteins. Raw data typically involving longitudinal R_1 , transversal R_2 relaxation rate constants, and $^{15}\text{N}\{-^1\text{H}\}$ -NOE are required to determine general order parameters (S^2) for each and every backbone NH beside the overall- (τ_c) and local-correlation times (τ_l) and chemical exchange rates (R_{ex}). Recently, chemical shifts have been used to obtain preliminary models of protein structures (Cavalli et al. 2007).

Structure calculation and refinement. Experimentally determined structural restraints are used as input for these structure elucidations. Computer programs, like XPLOR (Schwieters et al. 2006; Güntert 2011) combine as many restraints as possible in conjunction with general properties of proteins. These approaches convert the different structural restraints into target functions with general protein properties describing energy terms, all to be minimized at once. A bundle of structures will be obtained reflecting a selected molecular fold. Structure determination by NMR is yet a time-consuming iterative process requiring well-trained and devoted specialists. The most time-consuming processes are the chemical shift and the NOE assignments. Automating and black-boxing the assignment to increase the throughput and to eliminate subjective decisions during structure elucidation is an important goal, yet partially achieved. Efforts are made to fully automate resonance assignment in conjunction with structure calculations in an integrated manner (Volk et al. 2008).

Structure validation is strongly recommended by using special software (e.g., CheckShift (Ginzinger et al. 2007)) to check both the quality of the NMR experimental data and structural information before publication. In general, these programs provide a means of validating both the geometry and restraint violations of an NMR-driven protein structure bundle. The output provides both statistical summaries and visualizations on restraint violations, agreement with already deposited model parameters, quality of geometrical properties, etc.

Although NMR is a versatile tool for protein structure determination, it may have severe limitations such as spectral crowdedness, too fast relaxation, poor automation, etc. Traditionally, NMR has been used for small- or medium-size proteins. Drawbacks and technical problems originate from the limited spectral resolution and signal overlap. Introducing multidimensional NMR spectroscopy, in conjunction with the different isotope labeling schemes, as well as the increase of the applied static magnetic field equipped with cryoprobe helped a lot in enhancing the power of NMR spectroscopy. An additional limitation is connected to spin relaxation; coherence (T_2) and magnetization (T_1) vanishes fast in time for larger macromolecules. This means that there is less and less time to acquire NMR signals (uncertainty of chemical shifts), causing signal-broadening (half-width at half-height of Lorentzian) and weakening. Two techniques have been introduced lately to overcome this problem, a chemical one by the introduction of sample deuteration and a spectroscopic one, namely the transverse relaxation optimized spectroscopy (TROSY). Today, these technical novelties allow determining even structures of larger water soluble and membrane proteins (Wüthrich 2002).

Computer Modeling

Modern computer hardware and software technology allows to model proteins at the atomic level quite precisely. Several computer methods are available for experiment-based structure determination, interpretation of structure-function relationships, as well as construction of explicit models of proteins from atomic or molecular fragments.

Molecular Mechanics

Since proteins are very large, containing several thousand atoms, quantum mechanical calculations are not tractable for doing calculations on the whole system. Instead, the much simpler molecular mechanics (MM) or force-field methods are applied, where a molecular system is treated as an ensemble of atoms connected by strings (bonds). A price for simplicity is that electrons are not treated explicitly and correctly. Forces acting between atoms are described by simple mathematical expressions, adapted from classical physics. A certain force field is composed by interaction terms describing different types of strain present in a molecule. The total energy of the system is typically calculated as the sum of the following terms:

$$E_{\text{MM}} = V_{\text{stretch}} + V_{\text{bend}} + V_{\text{torsion}} + V_{\text{outofplane}} + V_{\text{nonbonded}} \quad (1)$$

where stretching, bending, torsion, out-of-plane bending, van der Waals, and electrostatic terms, respectively, are defined as follows

$$V_{\text{stretch}} = \sum_{\text{bonds}} K_{\text{stretch}} (r_{ij} - r_{ij}^0)^2 \quad (2)$$

$$V_{\text{bend}} = \sum_{\text{angles}} K_{\text{bend}} \left(\theta_{\text{ijk}} - \theta_{\text{ijk}}^0 \right)^2 \quad (3)$$

$$V_{\text{torsion}} = \sum_{\text{dihedrals}} K_{\text{torsion}} [1 + \cos(n\varphi - \delta)] \quad (4)$$

$$V_{\text{outofplane}} = \sum_{\text{impropers}} K_{\text{outofplane}} c_{\text{ijk;l}}^2 \quad (5)$$

$$V_{\text{nonbonded}} = \sum_{\text{non-bonded pairs}} [A_{\text{nonbonded}} (C_{12} r_{\text{ij}}^{-12} - C_6 r_{\text{ij}}^{-6}) + q_i q_j / D r_{\text{ij}}] \quad (6)$$

where r_{ij} , r_{ij}^0 , θ_{ijk} , and θ_{ijk}^0 are the actual and equilibrium bond lengths (interatomic distances) and bond angles, respectively, φ is the dihedral angle and δ is its phase shift, $\chi_{\text{ijk;l}}$ is the angle between the bond jl and the plane ijk , where j is the central atom. K_{stretch} , K_{bend} , K_{torsion} , and $K_{\text{outofplane}}$ refer to respective force constants; n is the multiplicity of the torsional function. $A_{\text{nonbonded}}$, C_6 and C_{12} are adjustable parameters. q_i is the net charge on atom i , D is the dielectric parameter. For review and comparison of various force fields see (Wang et al. 2001; MacKerell 2004; Hu and Jiang 2009; Khoruzhii et al. 2014).

Van der Waals and electrostatic parameters in Eq. 6 are usually not calculated for atoms that are relatively close (at a distance of one or two bonds). In order to reduce computational efforts (e.g., methyl) groups of atoms may be considered as a single interacting unit. On this basis a variety of so-called coarse-grained models has been developed (Giorgetti and Carloni 2014). Force constants and other parameters in the energy expression are fitted to experimental or calculated data. Different parameters are used, e.g., for sp^3 -, sp^2 -, and sp -type atoms and additional ones for atoms in carbonyl or peptide groups or in aromatic systems. Accordingly, the total number of adjustable parameters may be quite large. As the number of atom pairs in $V_{\text{nonbonded}}$ increases quadratically, in order to reduce computation time most programs use a cut-off value, beyond which the interactions are set to zero over an additional short distance. It has to be noticed that a certain set of parameters works only for a given force field, and a single parameter has no meaning in itself, it cannot be transferred to another force field.

Basically two ways of parameterization can be followed. Class I force fields like AMBER (Case et al. 2005) or GROMOS (2011) work with a simpler energy expression and their parameterization is based on experimental data. They find a wide application to proteins, nucleic acids, and carbohydrates, as well as their complexes. Class II force fields, e.g., the Merck Molecular Force Field (Halgren 1996), include higher order and cross terms as well; they are calibrated to quantum mechanically calculated energies and gradients, which increases their transferability and reliability.

Various, freely or commercially available protein modeling software packages make use of different force fields. AMBER (Case et al. 2005) is a package of molecular simulation programs including both source code and demos. GROMACS (Van der Spoel et al. 2005) is a highly optimized and fast program, which supports

several force fields and has an open source code. INSIGHT II from ACCELRYNS (<http://accelrys.com/>), SYBYL from TRIPOS (<http://www.tripos.com/>), CHARMM (<http://www.charmm.org>) and SCHRODINGER (<http://www.schrodinger.com/>) are licensed packages with many components, including force fields and programs for energy calculations.

A basic limitation to classical force fields is that they are valid only near local minima of the potential energy surface (PES) of a molecular system. Thus, they cannot be applied to, e.g., enzyme reactions, which pass through other regions of the PES, where the electronic distribution is strongly perturbed. In order to overcome this limitation, quantum mechanics and molecular mechanics may be combined by partitioning large covalent systems in a quantum region, which is embedded in a classical one. The total energy in this approach (hybrid quantum mechanical/molecular mechanical, QM/MM, method) (Gao and Truhlar 2002; Hu and Yang 2008) is as follows:

$$E = E_{QM} + E_{MM} + E_{QM/MM} \quad (7)$$

where E_{QM} , E_{MM} , and $E_{QM/MM}$ refer to the energies of the quantum and classical regions and their interaction, respectively. $E_{QM} = \langle \Psi | H_{QM} | \Psi \rangle$ is the expectation value of the Hamiltonian for the atoms in the quantum region, which is composed of kinetic, electron-electron, and electron-nucleus interaction energy terms.

$$H_{QM} = \sum \Delta_i + \sum 1/r_{ij} + \sum Z_a/r_{ia} \quad (8)$$

Here, subscripts “i,” “j,” and “a” refer to electrons and nuclei of the quantum region. E_{MM} is calculated for atoms of the classical region as in Eq. 1; the third term in Eq. 7 refers to the interaction between atoms of the quantum and the classical region, respectively. $E_{QM/MM} = \langle \Psi | H_{QM/MM} | \Psi \rangle$ with

$$H_{QM/MM} = \sum q_M/r_{iM} + \sum [Z_a q_M/r_{aM} + A_{\text{nonbonded}} (C_{12} r_{aM}^{-12} - C_6 r_{aM}^{-6})] \quad (9)$$

where M denotes an atom in the classical region.

For most QM/MM studies on proteins definition of the quantum region, (mostly the active site) needs cutting covalent bonds, which requires special care in treating the boundary. This is very important to avoid artifacts and to ensure a relatively smooth variation of the energy and its gradient when passing from the quantum to the classical region, prerequisite for a successful molecular dynamics calculation. Two main approaches are available for the appropriate treatment of atoms at the boundary: the link atom and the localized orbital methods. In the first method an artificial link (dummy or hydrogen) atom is used to saturate the dangling bonds, the quantum calculation should be done for the saturated system. In the second method, the bonds connecting the quantum and classical regions are cut and represented by

a set of strictly localized molecular orbitals, which are assumed to be transferable (Ferenczy and Náray-Szabó 2014).

Structure Prediction and Homology Modeling

Protein structures are very complicated; therefore, their determination by energy minimization, like in case of small molecules, would need astronomical computational work. However, several methods are available for the prediction of such structures (Bujnicki 2009; Krieger et al. 2003). Below we discuss some of them.

Homology modeling. Proteins, derived from a common ancestor, are called homologous and their primary sequence overlaps to a certain extent. It has been observed that overlap in the sequences results in the partial conservation of three-dimensional structures. This allows constructing a model of an unknown protein on the basis of other, homologous ones, possessing some sequence similarity. Conservation of folds among homologues guides modeling, which follows four steps: fold recognition and template selection, target-template alignment, model building, and model assessment.

The first two steps are often performed together, as the most common methods of identifying templates rely on the production of sequence alignments. Selection of the correct template is of basic importance to the quality of the final model. Subsequent steps can be performed iteratively to improve the quality of the result. After the template was chosen on the basis of percentage of identities and similarities with the target protein it has to be corrected with respect to missing residues and improper bond lengths and mutated to the corresponding residues of the target protein. The gaps in the target sequence have to be excised; loops should be built and inserted by selecting anchoring residues (cf. Fig. 2).

Loops satisfying the distance condition for connecting secondary structures can be located by a search in the Protein Data Bank. At last, using an appropriate force field the generated complete model has to be optimized.

Assessment of homology models can be performed with two main methods: statistical potentials or energy calculations. Both methods produce an estimate of the energy (or an analogous quantity) for the model being assessed. Unfortunately, neither of the two methods correlates very well with structural accuracy. Statistical potentials are computationally quite efficient and can be constructed empirically on the basis of observed residue-residue contact frequencies among proteins of known structure. Energy calculations are performed using a molecular mechanics force field (cf. section “[Molecular Mechanics](#)”). The use of these methods is based on the hypothesis stating that a native structure of a protein refers to a minimum on the energy hypersurface; thus, the lower the energy the better the structure. Some methods employ implicit solvation, representing the solvent around a single protein

```
ARIFYDKGFQGH CYKSEC-NLQP  
ARIFYDKG---RCYKSECPNLQP
```

Fig. 2 Sequence alignment of template (*top*) and target (*bottom*) in homology modeling

molecule as a continuous dielectric medium. Quality of homology models decreases with decreasing sequence identity. At 70 % and 25 % identity, the RMS deviation between the matched C^α atoms ranges between 1–2 and 2–4 Å, respectively. Several computer programs, like WHAT IF (Krieger et al. 2003), MODELLER (Fiser and Sali 2003), and SWISS-MODEL (2011), are available for homology modeling.

Protein threading (fold recognition) is based on the observation that the number of different folds in nature is limited to approximately 1000. Furthermore, 90 % of the new structures submitted to the Protein Data Bank in the past few years have similar folds to the ones already stored. The method uses sequences which have the same fold as some known structures but are not homologous with any of these. Prediction uses statistical relationships between structure and sequence. The prediction is made by aligning each amino acid in the target sequence to a given position in the template structure, and evaluating how well the target fits the template. After that the best-fit template is selected, the structural model of the sequence is built on the basis of the alignment with the chosen template.

Ab initio methods. When structural analogues are not available, prediction has to be based on other information. Various types of potentials as well as evolutionary information may be used in generating spatial restraints or identifying local structural building blocks. A popular ab initio method is Rosetta@home (2011), where the conformations, specified by the nonhydrogen backbone and C^β atoms, observed for each short sequence segment in known protein structures are used as an approximation of the set of local conformers, which the same sequence segment would occupy in the modeled protein. The combination of local conformers, which has the lowest overall energy, is considered as the best model. Before its structure has been determined experimentally, ROSETTA provided a model for the protein T0283 having 112 residues. Comparison of this model to the experimental structure indicates an RMS deviation of 1.8 Å over 92 residues. Despite significant success, the enormous computer cost of the procedure is still prohibitive for routine use.

Structure Representation

Protein structures contain typically thousands of atoms; therefore, their graphical representation, replacing old mechanical models, is essential for understanding structure, function, and their relationship. Modern computer technology provides an irreplaceable tool and allows insight into structural and functional details.

Molecular Graphics

Three-dimensional structure of a molecule may be displayed on the computer screen in various representations (Lesk et al. 2008). The simplest is where the molecule is displayed as a chemical graph, in which vertices stand for atoms and edges for bonds. Usually this representation is not very useful for molecules as large as proteins, as the figure gets too crowded. However, for structures derived

from X-ray data it is often used when mobility of various atoms is represented by spheres or ellipsoids centered at edges. In most cases, evaluation of data from NMR spectroscopy provides a bundle of structures, a superposition of those 10 or 20 ones, which fit best to restraints extracted from the spectrum. In order to simplify the model often a single structure, adopting the “average” conformation of the bundle, is shown.

More sophisticated is the ball-and-stick model, displaying atoms as balls and bonds as sticks connecting them. In space-filling, or CPK, representation atoms, represented by spheres with van der Waals radii, within which most of the electron density can be located, are drawn to suggest the amount of space they occupy. For proteins, ribbon diagrams offer a simple yet powerful means to visualize special features of protein structure. They indicate whether the polypeptide chain is twisted, folded, or unfolded. They are generated by interpolating a smooth curve through the polypeptide backbone. α -helices are shown as coiled ribbons, β -strands as arrows, random coils as lines or thin tubes. As an example for the performance of ribbon diagrams, we display a β -barrel, a bunch of large β -sheets that twist and coil to form a closed structure, on Fig. 3. The ribbon representation allows easy recognition of backbone structural motifs, while the space-filling model refers to other steric aspects of the macromolecule, e.g., location and shape of crevices.

A special feature of molecular graphics is that it can generate various surfaces around the molecule, on which a variety of its properties can be displayed. For proteins, the electrostatic potential (see section “[Electrostatics](#)”), water accessibility, and hydrophobicity (see section “[Solvent-Accessible Surface](#)”) patterns are most important. Commercial (see section “[Molecular Mechanics](#)”) and open source code software, like PyMOL (<http://www.pymol.org/>) and Jmol (<http://jmol.sourceforge.net/>), are available for molecular graphics.

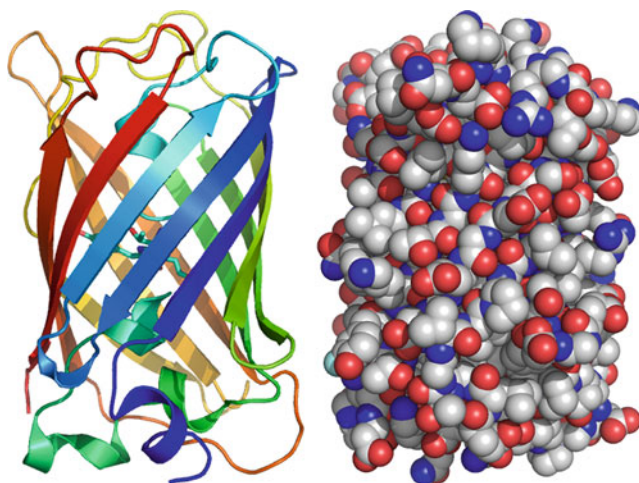


Fig. 3 Ribbon (*left*) and space-filling (*right*) representation of green fluorescent protein (PDB code: 1RRX)

Electrostatics

Since proteins have several polar or charged side chains, especially on their surface, electrostatics is a simple, yet quite reliable tool to study various processes involving proteins (Baker and McCammon 2009; Rocchia and Spagnuolo 2015). Electrostatic interactions may play a role in protonation, ligand binding, enzymatic catalysis, redox processes, and electron-proton coupling (e.g., in photosynthesis).

The electrostatic potential generated by the electrons and the nuclei of a protein molecule in a given spatial point, r_i , is calculated as follows

$$V(r_i) = \int \rho(r) |r - r_i| dr + \sum Z_a |r_i - R_{ia}| \quad (10)$$

where ρ is the electron density of the protein and Z_a is the nuclear charge of an atom located at the point R_{ia} . Exact calculation of the electron density involves the quantum mechanically determined total wave function; for proteins, this would need astronomical computer time. In order to reduce computational work, the density can be approximated by a sum of quantum mechanically defined molecular fragments, classical atomic monopoles, or multipoles.

Equation 10 refers to a protein molecule in vacuo; however, a realistic model should include explicit treatment of the surrounding biophase. This is done in the Poisson-Boltzmann approach, where the mobile ions around the protein are represented by a mean field approximation. A linearized equation can be derived, which combines the electron density of the protein and the charge represented by the mobile ions.

$$\nabla [\varepsilon(r)\nabla V(r)] + 4\pi [-2IV(r)/kT + \rho(r)] = 0 \quad (11)$$

Here $V(r)$, $\varepsilon(r)$, and $\rho(r)$ are the electrostatic potential, the dielectric constant and the protein electron density in point r , while I stands for the ionic strength. DelPhi, a versatile software package providing numerical solutions to the Poisson-Boltzmann equation, is widely used for the calculation and visualization of the protein electrostatic potential (DelPhi 2014).

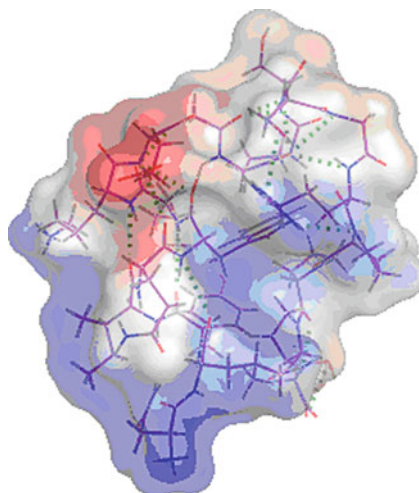
An application of the Poisson-Boltzmann equation is the estimation of the effective pK_a of a given side-chain, which is related to the electrostatic potential generated by the whole protein near the protonated atom. pK shifts caused by, e.g., site-directed mutagenesis can be precisely calculated using electrostatics (Gilson and Honig 1987). In subtilisin mutants the calculated, ionic strength-dependent pK shifts lie very close to the experimental values. The trend is correctly reflected for all but the lowest ionic strengths (cf. Table 1).

The electrostatic approach is quite useful in describing the binding of small ligands to proteins. Figure 4 shows a representation of the protein electrostatic potential on the surface, providing a simple means to roughly estimate position and binding strength of charged ligands. This point will be discussed in more detail in section “Protein Hydration.” Molecular recognition by proteins, as well as enzyme

Table 1 Comparison of calculated and experimental (in parentheses) pK_a shifts in subtilisin mutants as a function of the ionic strength

I	D99S mutant	E156S mutant
0.500	0.10 (0.09)	0.19 (—)
0.100	0.18 (0.23–0.29)	0.27 (0.25)
0.025	0.25 (0.36)	0.34 (0.41)
0.010	0.29 (0.42)	0.37 (0.42)
0.005	0.31 (0.38)	0.39 (—)
0.001	0.34 (—)	0.42 (0.39)

Fig. 4 Electrostatic potential of the miniprotein Tc5b_D9E mapped on the molecular surface (*red*: negative, *white*: neutral, *blue*: positive)



catalysis is also partly determined by electrostatics (cf. sections “[Ligand Binding](#)” and “[Enzyme Mechanisms](#),” respectively).

Redox processes involve a change in the charge distribution of the participating protein fragments; therefore, electrostatics may play here again a significant role. If a negative charge is located near the redox center, addition of an electron becomes more difficult, which lowers the redox potential. This effect is formally similar to the pK shift upon point mutation of a neutral side chain to a negative one. A specific advantage of the calculations is that contributions of various protein groups, the biophase, and ionized side chains can be treated separately. This information is very useful in the design of experiments, because probable effect of point mutations, ionic strength, and pH variation, respectively, can be predicted with a fair accuracy.

As an example, values of the one- and two-electron reduction potential of pheophytin-*a* were calculated in *N-N*-dimethylformamide by a combination of quantum mechanical, statistical mechanical, and polarizable continuum methods (Mehta and Datta 2008). Two different computational methods gave -0.92 and -1.03 V for the one-electron, while -1.34 and -1.30 V for the two-electron reduction potential values, respectively. The observed one- and two-electron potentials are -0.90 V and -1.25 V, respectively. Solution of the Poisson-Boltzmann equation

gave for the reduction potential of pheophytin-*a* within the thylakoid membrane -0.58 V, which is in good agreement with the reported value of -0.61 V.

Protonation-mediated modulation of electron transfer has been supposed to explain the location of the radical site in Compound I, an intermediate of the enzymatic reaction, of two related heme peroxidases (Menyhárd and Náray-Szabó 1999). Both molecular orbital and electrostatic potential calculations suggest that the spin distribution depends on the protonation state of the proximal His-Asp-Trp triad. If the transferable proton is shifted from the Trp side chain to Asp, the radical localizes on the indole group, while if it remains on the Trp the unpaired electron transfers to the heme group. Therefore, in cytochrome C peroxidase, Trp is deprotonated in Compound I, while it is protonated and neutral in cytosolic ascorbate peroxidase. Protonation state of the proximal residues is influenced by the electrostatic effect of the protein environment that differs in these enzymes, especially in the immediate vicinity of the Asp side chain.

Solvent-Accessible Surface

Protein surfaces are quite irregular; their analysis has application in structure-based drug design, since drugs bind in protein crevices. Since the hydrogen atom positions are generally not known, heavy atoms binding hydrogens are represented by a single sphere, whose van der Waals radius is increased by 0.1 – 0.3 Å. Several surfaces can be defined around proteins, maybe most important is the solvent- (typically water-) accessible surface (Connolly 1996). This is constructed using a probe sphere, representing the solvent molecule, which is rolled around the protein surface. A typical value of its radius is 1.4 Å, which corresponds to a water molecule. The contact surface is that part of the van der Waals surface which can be touched by the probe sphere; the solvent-accessible surface is derived as the location of the center of the probe as it rolls over the van der Waals surface (see Fig. 5). The solvent-excluded

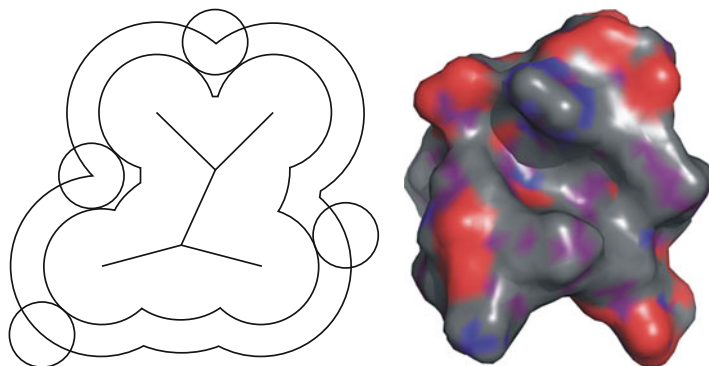


Fig. 5 *Left*: construction scheme of the solvent-accessible surface from the van der Waals surface (*inside*, obtained as a union of van der Waals spheres representing protein atoms). *Right*: solvent-accessible surface of the miniprotein Tc5b_D9E

volume of a protein is that volume of space where the probe sphere is in contact but does not collide with the atoms of the molecule.

The solvent-accessible surface area (SASA) is often used as a descriptor in quantitative structure-activity relationships (Connolly 1996). For a wide variety of molecules there is an approximate linear relation between solvation free energies and SASA. However, theoretical considerations indicate that the SASA model is incapable of accurately describing nonpolar solvation phenomena at length-scales comparable to the size of a water molecule. It is more useful at large length-scales when more extended hydrophobic surfaces are in contact with the solvent.

Dynamics

Finding the optimum structure representation of a protein is often not easy, as for larger molecules even the simplest ball-and-stick models can be too cumbersome. Thus, any simplification (neglecting H-atoms, fold representation by ribbons, etc.) of the graphics (see above) especially on a paper plot is important. It can also be misleading to use rigid representations of a flexible macromolecule. Graphical representations are based on set(s) of congruent experimental or computational restraints typically not shown in a figure. Clearly, it is inconvenient to display the electron density alone when reporting an X-ray structure or plotting all collected NOE restraints of structural models determined by NMR spectroscopy. However, the necessity to provide simple and easy-to-understand 3D structures of proteins drives scientists to neglect inherent dynamical properties of these molecules. For the very same reason, NMR spectroscopists tend to overrefine protein structures. Striving to obtain low RMSD values and to display an “X-ray-like” static structure, the reported model can be biased. The static picture often provided fails to satisfactorily explain the results of associated biological assays and is not fully compatible with ligand docking data. With the recent appearance of intrinsically unstructured proteins (IUPs) on stage, the importance of inherent dynamics is unquestionable. In other words, the absurdity of reporting only a single mean conformer is becoming evident. The present chapter gives an account of the dynamics occurring on different time scales and their interpretations and includes them in protein representation; it also gives some hints on the relevance of dynamics to biology.

Time Scales

Motions and events – especially when occurring periodically – are commonly characterized by their repetition rate, frequency, or by their time requirement. For example, our heart beats about once every second, so the heartbeat cycle, incorporating several steps such as the relaxation of the heart, opening of valves, etc., is occurring on the time scale of seconds. There are many slower events requiring more time than a second, and thus introduction of minutes, hours, days, seasons, years, etc., is easily justified. In our everyday life, people are experiencing

motions happening simultaneously but often on very different time scales. While going to work, which could take about an hour, our heartbeats about 4000 times and this daily routine is repeated nearly 300 times in a year. Proteins also have their “life,” which starts with folding after its translation, which event typically occurs on the millisecond time scale. As proteins are continually synthesized and degraded in all cells, the half-life of a protein is about an hour, varying from a few minutes to several weeks. However, protein dynamics typically comprises much faster events such as secondary structure formation (e.g., winding and unwinding of an α -helix) on the 10 ns–1 ms, or opening and closing of loops and hairpins on the 10 μ s–1 ms time scale.

The rotational diffusion correlation time, τ_c , is characteristic of any molecule as τ_c^{-1} is the average frequency of molecular tumbling in solution. The *correlation time* for short of a protein stands for the time that it takes for the molecule to reorient by 1 rad via tumbling in solution. Among others, being dependent on molecular size, τ_c is in the range of 1–10 ns for proteins in aqueous solution at normal conditions. However, there are additional interesting types of motions on an even faster time scale, such as protein backbone (1 ps–10 ns) and protein side-chain (0.1–10 ps) dynamics. Although even faster events may take place in proteins (e.g., bond vibrations studied by Fourier transform infrared or vibrational circular dichroism spectroscopy), these are less characteristic to the whole and more complicated to interpret.

NMR has long been used to decipher dynamics of biopolymers as the relaxation times and the line shapes of selected resonances are rather sensitive to chain motions. For proteins in solution, relaxation of the magnetically excited states occurs because of the above described very rapid and slower segmental motions. The somewhat localized motions are effective in causing relaxation in solution. As more than a single relaxation process often spanning a large time scale of motions (ps \leftrightarrow ms) characterizes the correlation function $G(\tau)$, it is advised to measure a number of different relaxation parameters at various magnetic fields and temperatures. During typical relaxation measurements, nuclear spins are excited by suitable radio frequency (RF) pulses, and then relaxation to equilibrium is monitored by heteronuclear single quantum coherence (HSQC) type 2D experiments. Commonly measured relaxation parameters are longitudinal (R_1), and transverse (R_2) relaxation rate constants, constraints obtained on the basis of the heteronuclear Overhauser effect ($^{15}\text{N}\{-^1\text{H}\}$ NOE) and additional cross-correlated relaxation parameters (Jarymowycz and Stome 2006).

Today, when routinely working with ^{15}N - and ^{13}C -labeled proteins, it is common to estimate the degree of main-chain mobility by using ^{15}N NMR relaxation. The associated experiments are straightforward to complete and the interpretation is relatively easy as it probes the reorientation of N–H bond vectors of protein amide bonds. Coupled primarily to fast (1 ps–10 ns) motion of the protein, heteronuclear NMR relaxation of backbone ^{15}N atoms is usually interpreted in the “model-free” framework, where it is supposed that global (τ_c) and local (τ_l) motions are uncorrelated. R_1 , R_2 , and $^{15}\text{N}\{-^1\text{H}\}$ NOE relaxation data of each ^{15}N atom are measured and subsequently fitted to the model to reconstruct spectral density

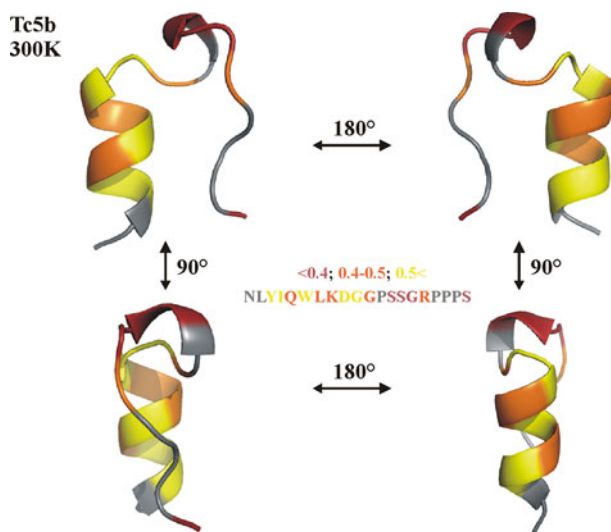


Fig. 6 Residue-specific square of the generalized order parameters, S^2 , mapped on the Tc5b miniprotein represented by its ribbon structure. Mobility is color-coded: *brown* ($S^2 < 0.4$) stands for high while *yellow* ($S^2 > 0.5$) for lower backbone NH mobility occurring on the 1 ps \leftrightarrow 10 ns time scale of motion. *Orange* represents 0.4–0.5 S^2 values (Pohl et al. unpublished)

components, $J(\omega_j)$. Results are analyzed in terms of the lowest number of sufficient secondary parameters. Local N–H motions are quantified by the square of the generalized order parameter, $0 \leq S^2 \leq 1$, reflecting the amplitude of N–H fluctuations with characteristic time constants (e.g., effective internal correlation time) (Cavanagh et al. 2007). A possible interpretation and graphical representation of the backbone dynamics is to map S^2 values on the graphical representation of the structure (cf. Fig. 6). In case of the present example, conformation of the central Trp residue, forming the hydrophobic cluster of the protein, is most restricted.

Molecular dynamics techniques allow consideration of S^2 values and NOE restraints as an ensemble property (Lindorff-Larsen et al. 2005; Richter et al. 2007). The obtained ensembles represent a more realistic view of these flexible molecules in solution than those calculated with conventional NMR structure refinement methods. The dynamically restrained ensembles occupy a considerably larger conformational space than the conventionally calculated ones and reproduce independent NMR parameters (e.g., chemical shifts) much better. The reconstruction of spectral density components, $J(\omega_j)$, from R_1 , R_2 , and $^{15}\text{N}\{-^1\text{H}\}$ NOE measurements is possible but more experiments are required (e.g., relaxation measurements at several B_0 values). Nevertheless, IUPs can also be analysed in terms of raw data as done recently, e.g., for the unstructured protease inhibitor calpastatin (Kiss et al. 2008).

Until now, we have discussed molecular motions 10 or 100 times faster than the rotational diffusion correlation time, τ_c , of the protein. As mentioned above, slower motions occurring on the microsecond-millisecond time scale are also of

significance (loop fluctuations, secondary structure reorientations, etc.). Motions of this regime are accessible to spin relaxation arising from modulation of the isotropic chemical shift of ^1H , ^{13}C , and ^{15}N spins coupled to chemical exchange. NMR can also provide information on this much slower motional regime by recording ^{13}C and ^{15}N heteronuclear ZZ-exchange, NOESY-spectroscopy, or various relaxation data (Cavanagh et al. 2007). More recently, residual dipolar couplings (RDCs) are also used for assessing dynamics of supra- τ_c time scales (Lange et al. 2008). These techniques are sensitive to protein motions or chemical kinetic processes happening on microsecond time scales (Palmer et al. 2001). Beside conformational events such as protein folding, other biochemical processes, e.g., enzyme catalysis, also occur on this time scale. It seems that information on the slower dynamics could be of great use in the future when interconversion rates and populations of states, cooperative conformational transition(s), functional aspects of enzymes, enzyme dynamics, etc., are modeled.

Dynamical Structures

Partially folded and molten globule states are of considerable interest and can be characterized by NMR and additional spectroscopic techniques. The molten globule state of a protein is folded to a given extent; it has native-like secondary structural elements and comprises a native-like 3D structure. The main difference from the native state is the absence of important side-chain/side-chain interactions. This intermediate structure can be stabilized under selected conditions making possible to obtain structural information. Proteins studied in their molten globule form are characterized as ensembles of a diverse set of interchanging conformers (Gspöner et al. 2006). The study of molten globule states by X-ray diffraction is often impossible and can be difficult using NMR techniques. However, from simple line shape and H/D exchange analysis obtained by NMR spectroscopy, measurement of the hydrodynamic radius, and deciphering stable (or residual) secondary structural subunits (by, e.g., circular dichroism or fluorescence spectroscopy), molten globule states can be appropriately characterized (Redfield 2004). Thus, protein modeling can be possible even in the molten globule state.

Dynamics of Interactions

Protein-protein and protein-ligand interactions have two apparently contradicting aspects. As these molecules are interacting 3D objects getting in contact with each other, partners should have complementary shapes of considerable stiffness over time (cf. section “[Ligand Binding](#)”). However, several phenomena cannot be understood without considering the inherent flexibility and plasticity of proteins. Thus, the shape of a protein should be kept rigid to enhance selectivity, but it is also dynamically changing upon binding of a partner molecule. Thus, proteins should both have a well-defined shape and the ability to change it.

Although the rigid lock-and-key model has been revisited in the past decades, one of the last standing cases is that of the so-called canonical protease inhibitors. Proteases have the task to cut other proteins into pieces for various reasons such as: digestion, activation, degradation, etc. Thus, they should be under severe control both in space, time, and concentration, often manifested by fine-tuned canonical inhibitors. Upon binding, they are not expected to influence any conformational changes around the binding site of the protease, as they have to bind in a substrate-like manner to inhibit activity. Recently, fast (ps-ns) backbone dynamics of some canonical inhibitors was measured and found that in contrast to the expectations they are quite mobile. In fact, these 35–40-residues-long inhibitor proteins stabilized by three disulfide bonds are as mobile as their protease binding region. Thus, even for canonical inhibitors, the classical lock-and-key model cannot remain valid. It was shown that the interaction cannot be properly described without taking into account the internal mobility of the partner proteins. Interestingly enough, the presence of flexible rather than rigid keys removes the necessity of an entirely complementary shape of the interacting partners. In fact, “keys” can gently adapt to the locks, which are slightly different from case to case (Fodor et al. 2005). Moreover, conformational changes or flexibility of partners was found to be synchronized and also encoded (Szenthe et al. 2004). This might even mean that the lock-and-key concept, proven utterly useful in the early days of biochemistry, should be interpreted dynamically and replaced by the more general hand-and-a-glove analogy (cf. section “[Ligand Binding](#)”) (Goh et al. 2004).

Interactions

Interaction of proteins with water, drugs, and other biological macromolecules plays an essential role in physiological processes. Owing to the very high number of atoms in a protein, in general, such interactions cannot be handled at the atomic level explicitly; simplified, still reliable models are needed.

Protein Hydration

Hydration of proteins is especially important in determining their properties (Makarov et al. 2002; Chaplin 2015). For example, protein activity is practically zero in the absence of hydrating water. The first hydration shell around proteins is ordered; surface water forms hydrogen-bonded patterns and positively charged surface side chains also take part in stabilization. The first hydration shell is denser by about 10–20 % than bulk water and this is probably responsible for keeping the molecules sufficiently separated so that they remain in solution.

Proteins are flexible in solution, which involves a variety of hydration states, absent in the crystal or in nonaqueous environments. Hydration also affects redox potentials. The biological activity of proteins appears to depend on the formation of a hydrogen-bonded network connecting the protein surface and water clusters.

Water molecules can form a bridge between the carbonyl oxygen atoms and amide protons of different peptide residues to stabilize protein-ligand and protein-protein complexes. Enzyme-substrate contact may be controlled by water molecules located in the protein environment, and one or more water molecules may be part of the catalytic machinery of enzymes. These water molecules act as proton donors or acceptors (or both) during reaction.

Computer modeling of protein hydration follows essentially two ways. One is to treat all water molecules in the hydration sphere explicitly and apply an appropriate force field (cf. section “[Molecular Mechanics](#)”) in molecular mechanics or molecular dynamics calculations. As an example, we present results of molecular dynamics calculations on aquaporin, a membrane protein facilitating the transport of water and other small solutes across biological membranes (Hub et al. 2009). Simulation was carried out with full electrostatics in a periodic simulation box containing the aquaglyceroporin tetramer, embedded within a lipid bilayer surrounded by water. The whole model consists of about 100,000 explicit atoms. The mechanism of water permeation through aquaporin was described, and it was found that protons are strictly excluded from the channel by a large electrostatic barrier. Both protein electrostatic field and desolvation contributes to this effect. Permeation of apolar molecules like carbon dioxide is hindered by a high energy barrier; therefore, this process can only be expected to take place in membranes with low intrinsic gas permeability. Analysis of MD simulations indicated that a narrowing region, containing aromatic and arginine side chains, filters uncharged solutes. Here, selectivity is controlled by the hydrophobic effect and steric restraints. As we see, simulation provides a specific tool to understand membrane permeation in atomic detail and to recognize its basic factors. No experimental means is available at present to provide such a detailed picture of the process.

Another way to consider protein hydration is to extend the concept of the solvent-accessible surface (cf. section “[Solvent-Accessible Surface](#)”) applying continuum dielectrics models. For example, combination of the electrostatic and an appropriate hydrophobic potential, accounting for hydrophilic and hydrophobic interactions, respectively, may be used as a solvation function, which considerably increases the reliability of protein structure predictions (Lin et al. 2007).

Ligand Binding

Binding of small-molecule ligands, such as drugs, vitamins, flavors, or pheromones, takes place in crevices of the protein surface. In a broader sense, ligand binding means the formation of a stable complex between a host and a guest (molecular recognition (Harmat and Náray-Szabó 2009)), and it can be best understood in terms of the *lock-and-key* analogy. Under lock we mean here the protein crevice, while the key stands for a small molecule or a fragment of a larger one, fully or partly embedded in the crevice. In general, we suppose that both lock and key are rigid; however, in many cases they are flexible and adopt their final shapes only upon binding. This process is best illustrated by the *hand-and-a-glove* analogy. If the macromolecular

crevice undergoes conformational changes in order to provide the optimal shape for embedding the relatively rigid guest molecule, we speak about induced fit. Alternatively, the guest molecules may adopt different conformations, some of which are appropriate for binding. This alternative is called conformational selection.

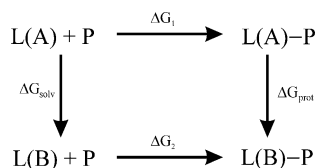
Steric fit means that interacting atoms may not interpenetrate beyond their van der Waals spheres and, simultaneously, the host crevice should be filled as perfectly as possible reducing the empty space between host and guest atoms to a minimum. Electrostatic fit refers to the maximization of ionic, hydrogen bonding and other types of polar interaction and is well accounted for by the molecular electrostatic potential provided by the host acting on a charge distribution representing the guest. The term hydrophobic fit refers to the association trend between apolar groups in the biophase. This may be explained either in terms of density differences between water and the host, or macromolecular crowding, an entropy effect reducing water-accessible surfaces of dissolved molecules upon association, in order to avoid unfavorable perturbation of water structure around the separated solutes.

The accurate prediction of the binding mode and affinity of a small molecule to its putative receptor is crucial for successful drug design and optimization; therefore, a variety of computer methods have been developed for that. *Docking* is an optimization method which predicts the preferred orientation of the ligand bound to the protein and forming a stable complex. It may be also used to predict the binding affinity using energy terms or an appropriately defined scoring function. Essentially three types of docking procedures are available (Mohan et al. 2005; Brás et al. 2014).

Rigid body docking assumes both the receptor and the small molecule to be rigid. *Flexible ligand docking* means that the receptor crevice is held fixed, but the ligand is treated as flexible (conformational selection, see above); this is the most commonly used method. In *flexible docking* both partners are considered to be able to change conformation easily. Following another classification, either a matching technique is used, which describes the protein crevice and the ligand as complementary surfaces, or the actual docking process is simulated and the ligand-protein interaction energies are calculated (Guvencs and MacKerrell 2009). Furthermore, modeling the role of the aqueous solvent in ligand-protein interactions is one of the key components. Three main computational techniques have been developed to model hydration: free energy methods, ligand-protein docking and scoring, and the explicit inclusion of tightly bound water molecules in modeling (Mancera 2007).

Shape complementarity methods describe the protein and ligand in terms of complementary molecular surfaces. The host and guest protein surfaces are defined in terms of the SASA (cf. section “[Solvent-Accessible Surface](#)”) and the matching surface description, respectively. The complementarity between these two surfaces is defined as the shape matching, which may help finding the appropriate orientation of the ligand in the crevice. These methods are typically fast and robust; however, in general, they cannot model the dynamic changes in the conformation of the protein-ligand complex. They can scan through several thousand guest molecules in some seconds and figure out which of them can bind at the crevice.

Fig. 7 The thermodynamic cycle used to compute solvation (*left*) and protein binding (*right*) free energies. *L*: ligand with substituents A or B, *P*: protein



Differences in binding energies of similar ligands, A and B, differing only by a few atoms, can be precisely calculated by the free-energy perturbation method (see Fig. 7). A thermodynamic cycle is constructed, where the horizontal free energies are determined experimentally, since they are difficult to calculate. The vertical free energies are calculated by simulation based on a force field or a continuum electrostatic model. ΔG_{solv} corresponds to the hydration energy change mutating A into B in water, while ΔG_{prot} is the binding energy change upon replacing A by B in the binding pocket. Since the thermodynamic cycle is closed and the free energy is a state function, we may write: $\Delta\Delta G_{\text{bind}}(A - B) = \Delta G_2 - \Delta G_1 = \Delta G_{\text{prot}} - \Delta G_{\text{solv}}$, the difference can be obtained from calculated quantities. Instead of simulating the horizontal physical binding processes, which is quite uncertain, the nonphysical transformations represented by the vertical reactions can be calculated relatively easily and precisely. An example for the successful application of the free-energy perturbation method is the calculation of the absolute binding free energy of benzamidine to trypsin, which is found to be $-6.78 \text{ kcal} \cdot \text{mol}^{-1}$ (Jiao et al. 2008), while the experimental binding free energy ranges from 6.3 to 7.3 $\text{kcal} \cdot \text{mol}^{-1}$. The calculated binding energy of a close analogue, diazamidine is weaker than that of benzamidine by $1.21 \text{ kcal} \cdot \text{mol}^{-1}$, agreeing well with the experimental value of $1.59 \text{ kcal} \cdot \text{mol}^{-1}$.

Docking by simulation is a much more complicated process. In this approach, the protein and the ligand are separated by a physical distance, and the guest finds its position in the crevice after a certain readjustment in the conformational space. This refers to translations, rotations, and conformational changes. Each of these induces an energy change of the system, which is calculated on the basis of a force field or a scoring function. Clearly, this method is much more sophisticated; it incorporates ligand flexibility, which is, in general, absent from complementarity techniques (see Zsoldos et al. 2007; Bras et al. 2014 and references therein). A disadvantage is that it takes much longer time to evaluate the optimal binding arrangement, since a rather large energy landscape has to be explored.

Docking is most commonly used in the field of drug design to quickly screen large databases of potential drugs *in silico* and to identify molecules that are likely to bind to a certain protein target (virtual screening). In lead optimization, docking may predict how a ligand binds to a protein crevice, which information may be used to design more potent and selective analogues. Several computer platforms are available for efficient and relatively fast docking, e.g., AUTODOCK (<http://autodock.scripps.edu/>).

Protein-Protein Interactions

Proteins often exert their function through interaction with other ones, forming complexes, e.g., in the ribosome or in cell membranes, during immune response or enzyme catalysis. Complex structures are essential in understanding action; therefore, beyond experimental techniques, like X-ray diffraction, electron microscopy, or mass spectrometry, *in silico* methods also are of great interest. However, due to a number of practical difficulties, it seems unlikely that solving the structures of protein complexes will become routine in the near future. Hence, computational docking techniques play an important role (Ritchie 2008).

The simplest protein-protein docking methods are based on the complementarity principle and focus on the properties of the contact surface. To do this, interaction patterns, observed in protein-protein complexes of known structure, have to be analysed. An example is based on shape and electrostatic complementarity (Fig. 8), where spatial matching of surfaces with complementary electrostatic potential patterns can be observed (Harmat and Náray-Szabó 2009). It can be supposed that much less improper orientations of the interacting partners would be obtained if one could first identify the interaction surfaces on them. For many (e.g., antibody-antigen or most enzyme-inhibitor) interactions, the binding site is known on one or more of the proteins to be docked. In other cases, a binding site may be strongly suggested by mutagenic or phylogenetic evidence.

Although the properties of protein-protein interfaces have been analysed in considerable detail, it remains a significant challenge to reliably predict the loca-

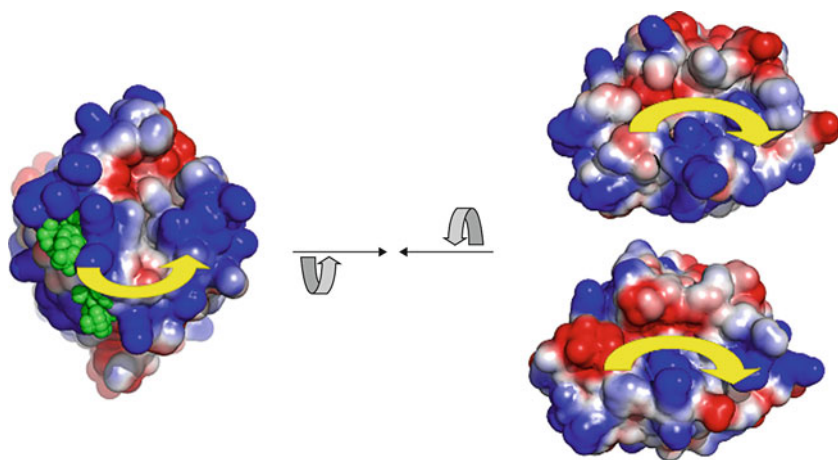


Fig. 8 Steric and electrostatic complementarity between C1 inhibitor (*left*, the docked heparin disaccharides are shown in *green*) and target proteases, factor XIa (*top right*) and C1s (*bottom right*). *Blue*, positive, *red*, negative electrostatic potential. *Yellow* arrows indicate the presumed position of the reactive center loop of C1 inhibitor in the complex. *Gray* arrows indicate that one of the interacting molecules must be rotated by 180° in order to get in the right position for overlap with the other at the contact surface (Figure by V. Harmat, Budapest)

tions of protein-protein interaction surfaces using computational techniques alone (Porollo and Meller 2007). Machine learning techniques are being used to develop automated protein-protein interface prediction software. These systems are typically trained using various combinations of, e.g., buried surface areas, desolvation and electrostatic interaction energies, hydrophobicity scores, and residue conservation scores.

Like in case of protein-ligand docking (cf. section “[Ligand Binding](#)”), for protein-protein docking we have essentially two alternatives (Ehrlich and Wade 2001; Vakser 2014). In the rigid-body approach interacting partners do not relax, their structures remain unchanged during complex formation. Quite often substantial conformational change occurs within the components as a result of their interaction. In this case a soft docking technique has to be applied, where conformational relaxation is allowed.

Docking involves two main steps: (i) generating a set of reliable configurations and (ii) distinguishing correct configurations from those not appropriate for complex formation. Docking algorithms may start with a rigid body representation of the protein, obtained by projecting interacting partners onto a three-dimensional grid and by distinguishing grid cells according to whether they are near or intersect the protein surface or are deeply buried within the core of the protein. A docking search is then performed by scoring the degree of overlap between pairs of grids in different relative orientations. After making exclusions based on experimental evidence or steric conflict, the remaining structures must be sampled by appropriately scoring each configuration. In the geometric hashing approach, each protein surface is first preprocessed to give a list of a few hundred critical points, which are then compared in a special algorithm to generate a relatively small number of trial docking orientations for grid scoring (Fischer et al. 1995).

An automated algorithm was developed to identify molecular surface complementarity (Katchalski-Katzir et al. 1992). It involves a digital representation of the molecules distinguishing between surface and interior, calculates a correlation function assessing the degree of surface overlap and penetration upon displacement using Fourier transform, and scans relative orientations in three dimensions. The algorithm estimates the degree of geometric match between the surfaces of the interacting molecules. The procedure is equivalent to a six-dimensional search, but it is much faster. The computation time depends only moderately on molecular size. Such methods can be used to evaluate very large numbers of configurations; however, this advantage is lost if conformational changes are also considered. Furthermore, it is possible to construct reasonable convolution-like scoring functions combining both steric and electrostatic fit.

Conformational changes can be considered in Monte Carlo methods, where an initial configuration is refined by taking random steps which are accepted or rejected based on their induced improvement in score, until a certain number of steps have been tried (Gray et al. 2003). It is assumed that convergence to the best structure should occur from a large class of initial configurations, only one of which needs to be accepted. Because of the difficulty of finding a scoring function which is highly discriminating and also converges to the correct configuration, various

refinements have been proposed. Monte Carlo methods are not guaranteed to search exhaustively, so that the best configuration may be missed even using a scoring function which would in theory identify it.

To find a scoring function, which forms an appropriate basis for selecting the best configuration, studies should be carried out on a standard benchmark of protein-protein interaction cases. For example, a benchmark of several dozens of known protein-protein complexes has been developed for testing docking methods (Mintseris et al. 2005). The set is chosen to cover a wide range of interaction (e.g., enzyme-inhibitor, antigen-antibody) types and to avoid repetition. Scoring functions are assessed on the rank they assign to the best structure and on the proportion of the benchmark cases for which they provide an acceptable result. Scores may be based on residue contacts, shape complementarity, interaction energies estimated by a force field, or on phylogenetic considerations. Often hybrid scores are created by combining one or more categories above.

Although solvation and desolvation effects are crucial in complex formation, most docking algorithms do not take into account the presence of water molecules near the protein-protein interface. An exception is the work by Van Dijk and Bonvin (2006).

In the following, we give an example for the successful application of molecular dynamics to a special problem, where impairing the flexibility of a small protein leads to severe consequences. A rare kidney disease, nephrotic syndrome type 2, is due to the pathogenicity of an allele encoding the R229Q polymorph variant of podocin, a small protein, which exists in the podocytes of the kidney, key members of its filtration system (Tory et al. 2014). Malfunction of the polymorph variant claiming the life of young patients is caused by the mutation it carries, but only under special circumstances. If both parents carry the polymorph – R229Q – variant, they will not necessarily have an unhealthy child. Only the polymorphism in association with a certain set of missense mutations will be detrimental. It was proposed that the R229Q mutation must impair the ability of podocin to dimerize: paired combination of missense mutations on one monomer of podocin and the polymorph mutation on the other must lead to nonfunctional dimers.

In a computational study, homology models of monomers of the wild-type and the polymorph variant of podocin were subjected to a 40 ns MD simulation. Podocin monomers contain a globular head domain and a long, helical tail. The mutated R229 residue is located in the head-domain. It turns inside and is stabilized by at least two H-bonds formed with negatively charged amino acids of the head-domain (E233, E237, and D244). In the R229Q polymorph monomer, where Q229 is too short to allow these contacts, E233 and E237 flip toward the tail domain and form H-bonds with positively charged amino acids of it (Fig. 9).

While the domains themselves are unchanged by the R229Q mutation, interactions between the head and tail domains restrict their hinge-like movement, and strongly affect the dynamics of the monomer (Fig. 10).

All the associated mutations that lead to unhealthy dimer formation when paired with the R229Q mutant are found in the helical tail region of the protein; thus, it was

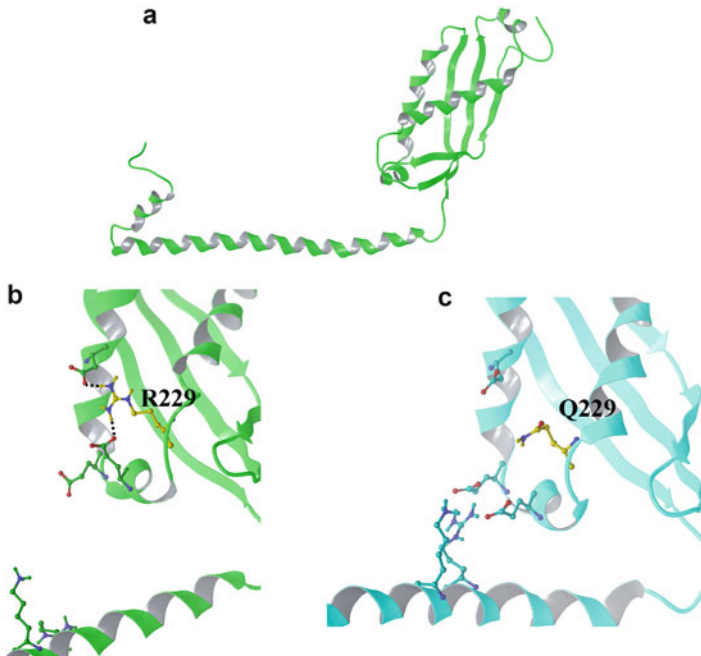
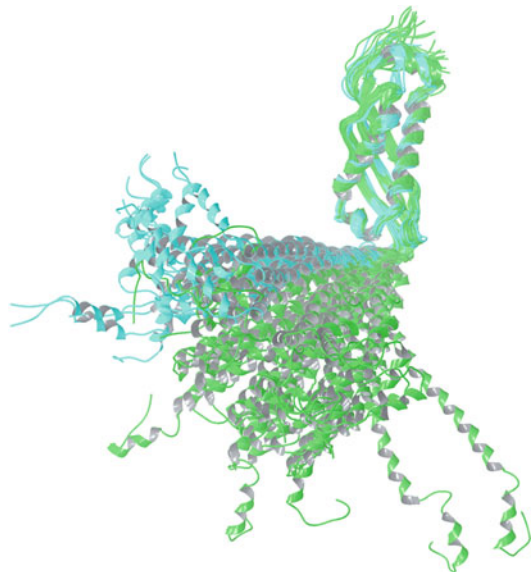


Fig. 9 (a). Structure of the native podocin monomer. (b). Interactions of the R229 residue of the wild-type protein with negatively charged groups of the head domain. (c). Interactions of the Q229 residue of the polymorph variant. Note that a new link is formed between the head and tail domains

Fig. 10 Difference in the flexibility of the wild-type protein (*green*) and the polymorph (*cyan*) monomers. All clusters of more than 0.1 % population of the last 10 ns of our simulations are shown, with their head domains superimposed



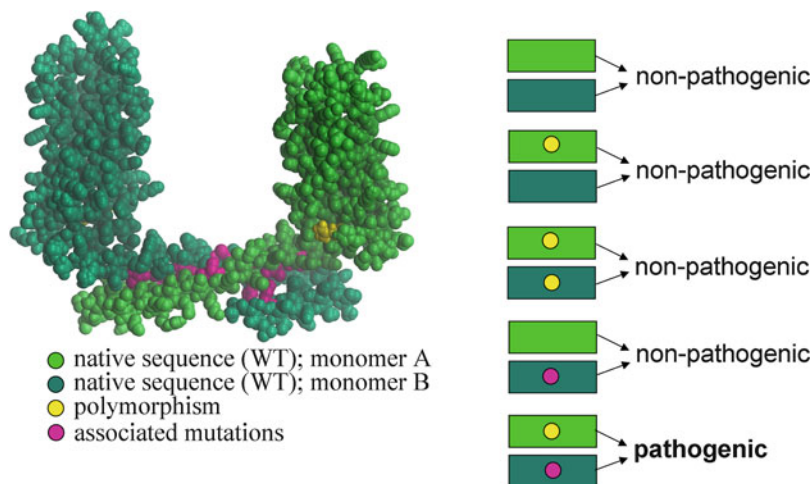


Fig. 11 Wild-type podocin dimer formation. Residue 229 is shown in *yellow*, mutation sites that lead to pathogenic dimer formation when faced with the polymorph monomer are shown in *magenta*. *WT* wild type

suggested that this must be the primary site of dimerization. A next set of 100 ns MD simulations involved dimer pairs, both nonpathogenic and pathogenic combinations (Fig. 11).

Interestingly, it was found that the loss of flexibility as a result of the R229Q mutation leads to formation of distorted dimers when pairing up with monomers carrying missense mutations on the helical tail region of podocin. On the other hand, it seems that the flexibility of the wild-type monomer allows for correction of small misfits in the dimerization process; thus, paired up with the same tail-mutated variants, the resulting dimers will not differ greatly from the wild-type dimer structure (Fig. 12). It was therefore shown that an inheritance pattern that defies the laws of Mendel could be explained by the impaired dimer-forming capacity of the R229Q polymorph, which is the result of the change of its overall flexibility, caused by a single mutation. Molecular dynamics simulations proved to be an ideal method to follow such changes.

Enzyme Mechanisms

Catalysis is a molecular transformation where the rate of the chemical reaction increases through the interaction of the reacting partners with the catalyst without changing the equilibrium constant. Accordingly, the catalyst modifies the potential surface of the noncatalyzed reaction and reduces the activation barrier. In case of

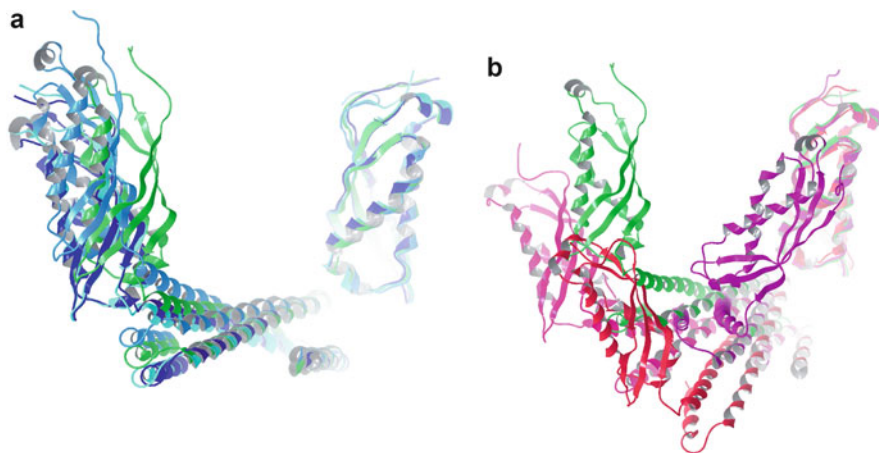


Fig. 12 Structure of nonpathogenic and pathogenic dimers of podocin. (*Left*) nonpathogenic dimers (colored by indifferent shades of *blue*: WT/R229Q, R229Q/R229Q, A284V/WT) superimposed on the WT/WT dimer (shown in *green*). (*Right*) pathogenic dimers (colored indifferent shades of *red*: A284V/A284V, A284V/R229Q, A297V/R229Q) superimposed on the WT/WT dimer (shown in *green*). *WT* wild type

enzyme catalysis, the reference is the activation energy in aqueous solution, where the reaction mechanism may considerably differ from that in the enzyme. The enzymatic process is especially effective, since rate increase may be as large as 10^{10} -fold in some cases. Protein modeling plays a crucial role in the understanding of the basis of enzymatic rate acceleration since experiments do not provide direct information on details, e.g., potential transition-state stabilization by the protein environment (Náray-Szabó et al. 1997, 2013).

Entropy effects may play an important role, though the actual advantage of their contribution to rate acceleration is still unresolved. This is partially due to the difficulties in obtaining estimates of entropic effects by modeling. Quite often, it is argued that the binding of reacting partners to an enzyme active site freezes their motion reducing the entropy contributing to the activation energy. However, it has been shown that this effect is much smaller than thought, because many of the movements that are free in the initial state of the reference solution reaction are also free in the transition state of the enzymatic process. The binding to the enzyme does not completely freeze the motion of the reacting fragments so that the entropy change during reaction is not zero. Furthermore, the binding entropy in the enzyme active site is not necessarily equal to the entropy change in the reference solution reaction (Warshel 2003).

Many enzymes make use of general acid or base catalysis, when a reaction step is modified and proton transfer is mediated by an acidic or basic side chain. This reduces the free energy of the proton transfer step and accelerates the reaction. For proton transfer, it may be supposed that rate acceleration is a consequence of the reduced distance between proton donor and proton acceptor, which is known to

determine the activation energy. However, it has been shown that both in enzymes and the reference reaction in water H-bond distances are similar. General acid/base catalysis plays a role in enzymes by reducing the free energy of the proton transfer step using a different reaction path with different reactants, which is, however, not a true catalytic effect. Simulation studies have indicated that special effects such as desolvation, low-barrier hydrogen bonds, steric strain, or entropy traps do not account for the major part of the catalytic power of enzymes (Warshel 2003).

Molecular modeling and experiments provide more and more evidence for the crucial role of protein electrostatics in enzymatic rate acceleration. Enzymes, acting as a “supersolvent,” can strongly stabilize polar structures, like ion pairs or the $(- + -)$ charge distribution located in their active sites. Stabilization in enzymes is stronger than in water, since preorganized and fixed protein dipoles are almost optimally oriented at the active site. The enzyme, therefore, provides a preoriented environment that stabilizes the transition state of the reaction. In contrary, water dipoles are distributed randomly and have to reorient in order to stabilize transition states. Reduction of the reorganization free energy in the enzyme is due to its folding into its final configuration, which precedes the catalytic process and takes place independently, during protein synthesis.

This contribution of the protein reorganization rather than the enzyme-substrate interaction is the most important factor in enzyme catalysis (Warshel 2003). In other words, enzymes have evolved to bind reacting groups in the correct position for reaction with a very high accuracy.

Because of the extreme importance of the protein environment in determining enzymatic reaction rates, the use of extended models and QM/MM computations methods is a must (cf. section “[Molecular Mechanics](#)”). Until now, a wide variety of QM/MM methods are available, which can be applied to the calculations on various enzyme reactions; it is therefore not easy to make a ranking. It seems, however, that basic features of most processes can be described by practically all types of methods, if a large enough number of protein atoms and surrounding water molecules is considered explicitly or if using a continuum approach combined with a relatively large model of the active site.

In the following we discuss a specific example, the catalytic mechanism of enzymatic phosphoryl transfer (Lasilla et al. 2011). This reaction plays a crucial role in most cellular functions, as DNA replication, signal transduction, metabolism, and transcription; it is involved in several enzymatic processes. Kinases catalyze the transfer of γ -phosphate from ATP to other molecules; mutases transfer phosphoryl groups between two atoms of the same molecule. Phosphatases accelerate the removal of phosphoryl groups, while polymerases catalyze the synthesis of nucleic acid polymers. Endonucleases are able to cleave the phosphodiester bond by transferring a phosphoryl group from a bridging oxygen atom to an activated water molecule.

It is generally claimed that phosphoryl transfer may follow basically two pathways (see Fig. 13). In the dissociative mechanism a trigonal metaphosphate intermediate is formed (Xu and Guo 2008), while the associative mechanism involves a relatively stable, trigonal bipyramidal intermediate (Lahiri et al. 2003).

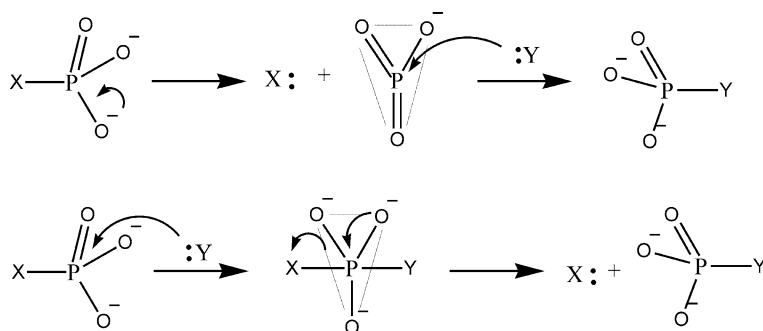


Fig. 13 Two pathways of the phosphoryl transfer reaction. *Top*: dissociative, *bottom*: associative. X and Y denote an enzyme or some functional group

Note that an intermediate refers to a local energy minimum on the reaction path. However, a third option has to be mentioned as well. This is the classical S_N2 mechanism with a trigonal bipyramidal transition state, referring to a maximum on the reaction path (Bernardi et al. 2002). The preferred pathway is determined by the nature of the phosphorus electrophile, the nucleophile, and the reaction medium (solvent or enzyme active site). Earlier computer simulations indicate that associative and dissociative mechanisms are similarly favored in the aqueous phase (Florian and Warshel 1998), and also calculations for different enzymes support either dissociative or associative pathways depending on a variety of factors (Klahn et al. 2006).

It is not easy to propose a mechanism on the basis of experimental evidence alone, since measured data provide in most cases indirect information on the energy and structure of the transition-state complex. Mildvan (1997) evaluated possible mechanisms by using the distances between the attacking and leaving axial oxygen atoms and phosphorus in the trigonal bipyramidal state to decide between associative or dissociative character. These distances may be considered as a measure of the strength of the P–O bond. The Mildvan criterion may be presented as an example for the unclear information often inherent in experiment. Based on X-ray diffraction experiments, Lahiri et al. (2003) reported on a pentacovalent phosphorane intermediate in the phosphoryl transfer reaction catalyzed by β -phosphoglucomutase, which refers to an associative mechanism. However, the associative character, as calculated from experimental bond distances in the intermediate, is quite weak (0.2), while for a truly associative reaction it should be near 1.0 (Berente et al. 2007). Another example of a trigonal bipyramidal intermediate, as observed by X-ray diffraction experiments, has been published by Barabas et al. (2004).

In general, the character of the transition state (TS) changes from associative to dissociative upon decrease in the pK_a of the leaving group, but the potential surface is very flat; therefore, careful calculations have to be performed in each special case in order to get reliable information on the associative character of the reaction

(Florián and Warshel 1998). Recently calculations have been performed on a variety of enzyme reactions where the associative mechanism seems to be supported; see, e.g., Elsasser et al. (2009).

Enzymatic phosphoryl transfer reaction provides an example where quantum mechanical calculations are irreplaceable in the precise description of the mechanism. The process is very complicated; however, up-to-date methods are available for sophisticated studies, and it can be anticipated that in the not too far future, mechanistic studies will be considered even by experimentalists as very useful in understanding subtle details.

Conclusion and Outlook

Molecules are not “seen” like macroscopic or microscopic objects, which provide a picture in our eyes by using visible light and human or artificial lens(es). As molecules are submicroscopic, they are too small to be visible even under the most powerful optical microscope. Although special optical microscopy allows seeing even cell cycles, the entire universe of atoms and molecules remains hidden for such a direct observation scheme. We may obtain “pictures” even on viruses having the size of 10–50 nm by transmission electron microscopy via electrons passing through an extensively prepared sample. Scanning tunneling microscopy and atomic force microscopy can improve resolution; however, neither of these techniques enable us to investigate molecules at an atomic level, especially not in solution. Accordingly, we must use indirect methods for visualization, by which we collect information on various physical properties related to the spatial position of atoms within a molecule. An ideal scenario is when a representation, or in other words a model, of a molecule on the screen is extracted and subsequently reconstructed from a suitable set of experimental data typically processed and evaluated by the computer.

Acknowledgments We are indebted to our colleagues Dr. Z. Gáspári, Dr. V. Harmat, and Ms. P. Rovó for important remarks on the manuscript and for providing most of the figures.

Bibliography

- Baker, N. A., & McCammon, J. A. (2009). Electrostatic interactions. In J. Gu & P. E. Bourne (Eds.), *Structural bioinformatics* (2nd ed., p. 575). Chichester: Wiley-Blackwell.
- Barabás, O., Pongrácz, V., Kovári, J., Wilmanns, M., & Vértesy, B. G. (2004). Structural insights into the catalytic mechanism of phosphate ester hydrolysis by dUTPase. *Journal of Biological Chemistry*, 279, 42907.
- Berente, I., Beke, T., & Náray-Szabó, G. (2007). Quantum mechanical studies on the existence of a trigonal bipyramidal phosphorane intermediate in enzymatic phosphate ester hydrolysis. *Theoretical Chemistry Accounts*, 118, 129.
- Bernardi, F., Bottoni, A., De Vivo, M., Garavelli, M., Keserú, G. M., & Náray-Szabó, G. (2002). A hypothetical mechanism for HIV-1 integrase catalytic action: DFT modelling of a bio-mimetic environment. *Chemical Physics Letters*, 362, 1.

- Bourgeois, D., & Royant, A. (2005). Advances in kinetic protein crystallography. *Current Opinion in Structural Biology*, *15*, 538.
- Brandén, C., & Tooze, J. (1999). *Introduction to protein structure*. New York: Garland.
- Brás, N. F., Cerqueira, N. M. F. S. A., Sousa, S. F., Fernandes, P. A., & Ramos, M. J. (2014). Protein ligand docking in drug discovery. In G. Náray-Szabó (Ed.), *Protein modelling* (p. 249). Cham-Heidelberg/New York/Dordrecht/London: Springer.
- Bujnicki, J. M. (Ed.). (2009). *Prediction of protein structures, functions, and interactions*. Chichester: Wiley-Blackwell.
- Case, D. A., Cheatham, T. E., III, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., & Woods, R. (2005). The Amber biomolecular simulation programs. *Journal of Combinatorial Chemistry*, *26*, 1668.
- Cavalli, A., Salvatella, X., Dombson, C. M., & Vendruscolo, M. (2007). Protein structure determination from chemical shifts. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 9615.
- Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, Rance, M., & Skelton, N. J. (2007). *Protein NMR spectroscopy, principles and practice*. Amsterdam: Elsevier.
- Chaplin, M. (2015). *Protein hydration*. London South Bank University, http://www1.lsbu.ac.uk/water/protein_hydration.html. Downloaded 18 Mar 2015.
- Chayen, N. E. (Ed.). (2007). *Protein crystallization strategies for structural genomics* (Biotechnology series). La Jolla: International University Line.
- Connolly, M. L. (1996). *Molecular surfaces: A review*. <http://www.netsci.org/Science/Compchem/feature14.html>. Retrieved on 7 Mar 2011.
- Dambrot, S. M. (2012). Small is beautiful: Viewing hydrogen atoms with neutron protein crystallography, <http://phys.org/news/2012-09-small-beautiful-viewing-hydrogen-atoms.html>. Downloaded: 17 Mar 2015.
- DelPhi. (2014). http://wiki.c2b2.columbia.edu/honiglab_public/index.php/Software:DelPhi. Retrieved on 14 Mar 2016.
- Drenth, J. (2007). *Principles of protein x-ray crystallography* (3rd ed.). New York: Springer.
- Ehrlich, L. P., & Wade, R. C. (2001). Protein-protein docking. *Reviews in Computational Chemistry*, *17*, 61.
- Elsasser, B., Valiev, M., & Weare, J. H. (2009). A dianionic phosphorane intermediate and transition states in an associative A(N) + D-N mechanism for the ribonucleaseA hydrolysis reaction. *Journal of the American Chemical Society*, *131*, 3869.
- Emsley, P., & Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. *Acta Crystallographica*, *D60*, 2126.
- Ferenczy, G. G., & Náray-Szabó, G. (2014). Strictly localised molecular orbitals in QM/MM methods. In G. Náray-Szabó (Ed.), *Protein modelling* (p. 71). Cham-Heidelberg/New York/Dordrecht/London: Springer.
- Fischer, D., Lin, S., Wolfson, H. L., & Nussinov, R. (1995). A geometry-based suite of molecular docking processes. *Journal of Molecular Biology*, *248*, 459.
- Fiser, A., & Sali, A. (2003). Modeller: Generation and refinement of homology-based protein structure models. *Methods in Enzymology*, *374*, 461.
- Fitter, J., Gutberlet, T., & Katsaras, J. (Eds.). (2006). *Neutron scattering in biology, techniques and applications*. Berlin: Springer.
- Florián, J., & Warshel, A. (1998). Phosphate ester hydrolysis in aqueous solution: Associative versus dissociative mechanisms. *Journal of Physical Chemistry B*, *102*, 719.
- Fodor, K., Harmat, V., Kardos, J., Antal, J., Hetényi, C., Perczel, A., Szenthe, B., Gáspári, Z., Katona, G., & Gráf, L. (2005). Conformational adaptation of a canonical protease inhibitor upon its binding to the target protease increases specificity. *FEBS Journal*, *272*, 167.
- Gao, J., & Truhlar, D. G. (2002). Quantum mechanical methods for enzyme kinetics. *Annual Review of Physical Chemistry*, *53*, 467–505.
- Gilson, M. K., & Honig, B. (1987). Calculation of electrostatic potentials in an enzyme active site. *Nature*, *330*, 84.

- Ginzinger, S. W., Gerick, F., Coles, M., & Heun, V. (2007). CheckShift: Automatic correction of inconsistent chemical shift referencing. *Journal of Biomolecular NMR*, *39*, 223.
- Giorgetti, A., & Carloni, P. (2014). Molecular mechanics/coarse-grained models. In G. Náray-Szabó (Ed.), *Protein modelling* (p. 165). Cham-Heidelberg/New York/Dordrecht/London: Springer.
- Goh, C. S., Milburn, D., & Gerstein, M. (2004). Conformational changes associated with protein-protein interactions. *Current Opinion in Structural Biology*, *14*, 104.
- Gray, J. J., Moughan, S. E., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. A., & Baker, D. (2003). Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *Journal of Molecular Biology*, *331*, 281.
- GROMOS. (2011). Dynamic modelling of molecular systems. <https://www1.ethz.ch/igc/GROMOS/>. Retrieved on 10 Mar 2011.
- Gsponer, J., Hoppearuoho, H., Whittaker, S. B. M., Spence, G. R., Moore, G. R., Paci, E., Radford, S. E., & Vendruscolo, M. (2006). Determination of an ensemble of structures representing the intermediate state of the bacterial immunity protein Im7. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 99.
- Guvenç, O., & MacKerrell, A. D., Jr. (2009). Computational evaluation of protein-small molecule binding. *Current Opinion in Structural Biology*, *19*, 56.
- Güntert, P. (2011). Automated protein structure determination from NMR data. In A. J. Dingley & S. M. Pascal (Eds.), *Advances in biomedical spectroscopy, vol. 3: Biomolecular NMR spectroscopy* (p. 341). Amsterdam: Ios Press. doi:10.3233/978-1-60750-695-9-338.
- Halgren, T. A. (1996). Merck molecular force field. I-V. *Journal of Combinatorial Chemistry*, *17*, 490.
- Harmat, V., & Náray-Szabó, G. (2009). Theoretical aspects of molecular recognition. *Croatica Chemica Acta*, *82*, 277.
- Horsefield, R., & Neutze, R. (2006). Crystallization of lysozyme by the hanging drop method, http://www.csb.gu.se/rob/PDFs/Crystallisation_Course_2006_Part-I.pdf. Retrieved on 10 Mar 2011.
- Hovmöller, S., Zhou, T., & Ohlson, T. (2002). Conformations of amino acids in proteins. *Acta Crystallographica*, *D58*, 768.
- Hu, Z., & Jiang, J. (2009). Assessment of biomolecular force fields for molecular dynamics simulations in a protein crystal. *Journal of Computational Chemistry*. doi:10.1002/jcc.21330.
- Hu, H., & Yang, W. (2008). Free energies of chemical reactions in solution and in enzymes with ab initio quantum mechanics/molecular mechanics methods. *Annual Review of Physical Chemistry*, *59*, 573.
- Hub, J. S., Grubmüller, H., & de Groot, B. L. (2009). Dynamics and energetics of permeation through aquaporins. What do we learn from molecular dynamics simulations? In E. Beitz (Ed.), *Handbook of experimental pharmacology, vol. 190, aquaporins* (p. 57). Berlin: Springer.
- Jarymowycz, V. A., & Stome, M. J. (2006). Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences. *Chemical Reviews*, *106*, 1624.
- Jiao, D., Golubkov, P. A., Darden, T. A., & Ren, P. (2008). Calculation of protein-ligand binding free energy by using a polarizable potential. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 6290.
- Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A. A., Aflalo, C., & Vakser, I. A. (1992). Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques. *Proceedings of the National Academy of Sciences of the United States of America*, *89*, 2195.
- Khoruzhii, O., Butin, O., Illarionov, A., Leontyev, I., Olevanov, M., Ozrin, V., Pereyaslavets, L., & Fain, B. (2014). Polarizable force fields for proteins. In G. Náray-Szabó (Ed.), *Protein modelling* (p. 91). Cham-Heidelberg/New York/Dordrecht/London: Springer.
- Kiss, R., Kovács, D., Tompa, P., & Perczel, A. (2008). Local structural preferences of calpastatin, the intrinsically unstructured protein inhibitor of calpain. *Biochemistry*, *47*, 6936.

- Klähn, M., Rosta, E., & Warshel, A. (2006). On the mechanism of hydrolysis of phosphate monoesters dianions in solutions and proteins. *Journal of the American Chemical Society*, *128*, 15310.
- Krieger, E., Nabuurs, S. B., & Vriend, G. (2003). Homology modeling. *Methods of Biochemical Analysis*, *44*, 509.
- Lahiri, S. D., Zhang, G., Dunaway-Mariano, D., & Allen, K. N. (2003). The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction. *Science*, *299*, 2067.
- Lange, O. F., Lakomek, N. A., Fares, C., Schroder, G. F., Walter, K. F. A., Becker, S., Meiler, J., Grubmuller, H., Griesinger, C., & de Groot, B. L. (2008). Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science*, *320*, 1471.
- Lasilla, J. K., Zalatan, J. G., & Herschlag, G. (2011). Biological phosphoryl transfer reactions: Understanding mechanism and catalysis. *Annual Review of Biochemistry*, *80*, 669–702. doi:10.1146/annurev-biochem-060409-092741.
- Lesk, A. M., Bernstein, H. J., & Bernstein, F. C. (2008). Molecular graphics in structural biology. In M. Peitsch & T. Schwede (Eds.), *Computational structural biology, methods and applications* (p. 729). Singapore: World Scientific Publishing.
- Lin, M. S., Fawzi, N. L., & Head-Gordon, T. (2007). Hydrophobic potential of mean force as a solvation function for protein structure prediction. *Structure*, *15*, 727.
- Lindorff-Larsen, K., Best, R. B., Depristo, M. A., Dobson, C. M., & Vendruscolo, M. (2005). Simultaneous determination of protein structure and dynamics. *Nature*, *433*, 128.
- Luft, J. R., Collins, R. J., Fehrman, N. A., Lauricella, A. M., Veatch, C. K., & DeTitta, G. T. (2003). A deliberate approach to screening for initial crystallization conditions of biological macromolecules. *Journal of Structural Biology*, *142*, 170.
- MacKerell, A. D., Jr. (2004). Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, *25*, 1584.
- Makarov, V., Pettitt, B. M., & Feig, M. (2002). Solvation and hydration of proteins and nucleic acids: A theoretical view of simulation and experiment. *Accounts of Chemical Research*, *35*, 376.
- Mancera, R. L. (2007). Molecular modeling of hydration in drug design. *Current Opinion in Drug Discovery & Development*, *10*, 275.
- Mehta, N., & Datta, S. N. (2008). Theoretical investigation of redox species in condensed phase. *Journal of Chemical Sciences*, *119*, 501.
- Menyhárd, D. K., & Náray-Szabó, G. (1999). Electrostatic effect on electron transfer at the active site of heme peroxidases: A comparative molecular orbital study on cytochrome C peroxidase and ascorbate peroxidase. *Journal of Physical Chemistry B*, *103*, 227.
- Merlino, A., Krauss, I. R., Albino, A., Pica, A., Vergara, A., Masullo, M., De Vendittis, E., & Sica, F. (2011). Improving protein crystal quality by the without-oil microbatch method: Crystallization and preliminary X-ray diffraction analysis of glutathione synthetase from *Pseudoalteromonashaloplanktis*. *International Journal of Molecular Sciences*, *12*, 6312.
- Mildvan, A. S. (1997). Mechanisms of signaling and related enzymes. *Proteins: Structure, Function, and Genetics*, *29*, 401.
- Mintseris, J., Wiehe, K., Pierce, B., Anderson, R., Chen, R., Janin, J., & Weng, Z. (2005). Protein-protein docking benchmark 2.0: An update. *Proteins: Structure, Function, and Bioinformatics*, *60*, 214.
- Mohan, V., Gibbs, A. C., Cummings, M. D., Jaeger, E. P., & DesJarlais, R. L. (2005). Docking: Successes and challenges. *Current Pharmaceutical Design*, *11*, 323.
- Náray-Szabó, G. (Ed.). (2014). *Protein modelling*. Cham-Heidelberg/New York/Dordrecht/London: Springer.
- Náray-Szabó, G., Fuxreiter, M., & Warshel, A. (1997). Electrostatic basis of enzyme catalysis. In G. Náray-Szabó & A. Warshel (Eds.), *Computational approaches to biochemical reactivity*. Dordrecht: Kluwer.
- Náray-Szabó, G., Oláh, J., & Krámos, B. (2013). Quantum mechanical modeling: A tool for the understanding of enzyme reactions. *Biomolecules*, *3*, 662.

- Palmer, A. G., III, Kroenke, C. D., & Loria, J. P. (2001). Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods in Enzymology*, 339, 204.
- Petsko, G. A., & Ringe, D. (2015). Time-resolved crystallography, Petsko & Ringe Laboratories, Brandeis University, <http://www.bio.brandeis.edu/prLab/time.html>. Downloaded, 17 Mar 2015.
- Porollo, A., & Meller, J. (2007). Prediction-based fingerprints of protein interactions. *Proteins: Structure, Function, and Bioinformatics*, 66, 630.
- Protein Data Bank. (2015). <http://www.rcsb.org/pdb/results/results.do?qrId=BA1B8D74&tabtoshow=Current>. Retrieved on 13 Mar 2015.
- Redfield, C. (2004). NMR studies of partially folded molten globule states. In A. K. Downing (Ed.), *Protein NMR techniques* (2nd ed.). Totowa: Humana Press.
- Richter, B., Gsponer, J., Várnai, P., Salvatella, X., & Vendruscolo, M. (2007). The MUMO (Minimal Under-Restraining Minimal Over-Restraining) method for the determination of native state ensembles of proteins. *Journal of Biomolecular NMR*, 37, 117.
- Ritchie, D. W. (2008). Recent progress and future directions in protein-protein docking. *Current Protein & Peptide Science*, 9, 1.
- Rocchia, W., & Spagnuolo, M. (Eds.). (2015). *Computational electrostatics for biological applications*. Cham-Heidelberg/New York/Dordrecht/London: Springer.
- Rosetta@home. <http://boinc.bakerlab.org/rosetta/>. Retrieved 10 Mar 2011.
- Schmucki, R., Yokoyama, S., & Güntert, P. (2009). Automated assignment of NMR chemical shifts using peak-particle dynamics simulation with the DYNASSIGN algorithm. *Journal of Biomolecular NMR*, 43, 97.
- Schorn, C., & Taylor, B. J. (2004). *NMR-spectroscopy: Data acquisition* (2nd ed.). New York: Wiley-VCH.
- Schwieters, C. D., Kuszewski, J. J., & Clore, G. M. (2006). Using Xplor-NIH for NMR molecular structure determination. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 48, 47.
- SWISS-MODEL. A fully automated protein structure homology-modeling server, <http://swissmodel.expasy.org/>. Retrieved on 10 Mar 2011.
- Szente, B., Gáspári, Z., Nagy, A., Perczel, A., & Gráf, L. (2004). Same fold with different mobility: Backbone dynamics of small protease inhibitors from the desert locust, *Schistocerca gregaria*. *Biochemistry*, 43, 3376.
- Tory, K., Menyárd, D. K., Woerner, S., Nevo, F., Gribouval, O., Kerti, A., Stráner, P., Arrondel, P., Cong, E. H., Tulassay, T., Mollet, G., Perczel, A., & Antignac, C. (2014). Mutation-dependent recessive inheritance of NPHS2-associated steroid-resistant nephrotic syndrome. *Nature Genetics*, 46, 299.
- Vakser, I. A. (2014). Protein-protein docking: From interaction to intercome. *Biophysical Journal*, 107, 1785.
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, 26, 1701.
- Van Dijk, A. D. J., & Bonvin, A. M. J. J. (2006). Solvated docking: Introducing water into the modelling of biomolecular complexes. *Bioinformatics*, 22, 2340.
- Volk, J., Herrmann, T., & Wüthrich, K. (2008). Automated sequence-specific protein NMR assignment using the memetic algorithm MATCH. *Journal of Biomolecular NMR*, 41, 127.
- Wang, W., Donini, O., Reyes, C. M., & Kollman, P. A. (2001). Biomolecular simulations: Recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. *Annual Review of Biophysics and Biomolecular Structure*, 30, 211.
- Warshel, A. (2003). Computer simulations of enzyme catalysis: methods, progress, and insights. *Annual Review of Biophysics and Biomolecular Structure*, 32, 425.
- Wüthrich, K. (2002). NMR studies of structure and function of biological macromolecules, http://nobelprize.org/nobel_prizes/chemistry/laureates/2002/wutrich-lecture.pdf. Retrieved on 10 Mar 2011.

- Xu, D., & Guo, H. (2008). Ab initio QM/MM studies of the phosphoryl transfer reaction catalyzed by PEP mutase suggest a dissociative metaphosphate transition state. *Journal of Physical Chemistry B*, 112, 4102.
- Xu, Y., Xu, D., & Liang, J. (2007). *Computational methods for protein structure prediction and modeling* (Vol. 1–2). New York: Springer.
- Zsoldos, Z., Reida, D., Simona, A., Sadjada, S. B., & Johnson, A. P. (2007). eHiTS: A new fast, exhaustive flexible ligand docking system. *Journal of Molecular Graphics and Modelling*, 26, 198.

Wieslaw Nowak

Contents

Introduction	1628
Formalism of Molecular Mechanics and Molecular Dynamics Methods:	
A Short Presentation	1630
Some History	1630
On the Origin of Potential Energy Surface (PES) Concept	1631
Force Fields	1633
Structure Optimization or Energy Minimization	1635
General Molecular Dynamics Scheme	1635
Practical Aspects	1638
MD Codes	1638
Developments in Force Fields and Parametrization	1639
Visualization	1640
Review of Reviews	1640
Selected Examples of Applications of MD to Study Proteins	1643
Protein Folding Studies	1643
Intrinsically Disordered Proteins	1644
Protein-Drug Interactions and Docking	1644
Spectroscopy Experiments	1645
Functionally Important Motions (FIMs)	1645
Molecular Machines	1646
The Mechanism of Enzymatic Activity	1646
Transport Phenomena in Proteins	1647
Structure and Dynamics of Ion Channels and Porins	1647
Protein-DNA Interactions	1648
Origins of Molecular Diseases	1649
Simulations of Single-Molecule AFM Experiments	1649
Conclusions and Future Directions	1650
Bibliography	1652

W. Nowak (✉)

Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University, Toruń, Poland

e-mail: wiesiek@fizyka.umk.pl; Wieslaw.Nowak@fizyka.umk.pl

Abstract

The present advanced state of the computer hardware offers superb opportunities for further explorations of protein structure and dynamics. Sound and well-established theoretical models are successfully used for searching new biochemical phenomena, correlations, and protein properties. In this chapter, the fast-growing field of computer simulations of protein dynamics is panoramically presented. The principles of currently used computational methods are briefly outlined, and representative examples of their recent advanced applications are given. In particular protein folding studies, intrinsically disordered proteins, protein-drug interactions, ligand transport phenomena, ion channel activity, molecular machine mechanics, origins of molecular diseases, and simulations of single-molecule AFM experiments are addressed. Special attention is devoted to emerging methods of enhanced molecular dynamics.

Introduction

Proteins perform their functions through selective intermolecular interactions. These forces are exerted on external objects such as other proteins, inhibitors, nucleic acids, membranes, signal molecules, etc. The strength and temporal evolution of interactions do depend on proteins' amino acid composition and a particular spatial arrangement of these basic building blocks. Non-native conformations usually prevent proteins from performing their "natural" tasks. The definition of what is "a protein conformation" is somewhat blurry, since proteins are flexible molecules. The energy landscapes of such large systems are very rich, complex, and highly multidimensional. In other words, one may expect that in its native, functional form, a protein adopts not only one particular conformation, but rather numerous, closely lying related structures may participate in a given chemical or biological process. Quite often large-scale molecular motions are required for performing a given protein function (Bahar et al. 2010; van Oijen 2007). For example, hemoglobin increases its oxygen-binding capability via famous allosteric effects (Tekpinar and Zheng 2013), the enzyme polymerase augmented by PCNA DNA clamp (Rydzewski et al. 2015b) processes 1,000 base pairs per second during DNA replication, and G protein-coupled opioid receptor is activated by subtle conformational changes (Huang et al. 2015).

Capturing protein dynamical structures at work is not an easy task (Russel et al. 2009). Numerous experimental techniques have been developed: spectroscopy of all sorts, NMR, EPR, and even time-dependent X-ray crystallography. The computer era brought an excellent additional tool for studies of protein structures and dynamics. Since the pioneering ideas of A. Warshel et al. (1970) and the first application work by J.A. McCammon et al. (1977), empirical force fields have been used thousands of times to describe computationally proteins' dynamics and their interactions. At this point, one has to note that currently, quantum mechanics is the only physical theory having appropriate rigor for the description of molecular

systems and their interactions. Unfortunately, *ab initio* (or force field free) molecular dynamics (MD) simulations of proteins are still in their childhood and far from being available for every research group (Dal Peraro et al. 2007) (Rydzewski and Nowak, chapter “► [Molecular Dynamics in Excited State](#)”). On the other hand, a classical approximation, where individual atoms (or even groups of atoms in the so-called coarse-grained MD) are replaced by model material points interacting through analytically predetermined potentials, offers a very promising alternative for tedious, expensive, and difficult experimental studies of proteins. Computer modeling of protein structure and dynamics is a currently mature field of science (Karplus and McCammon 2002). Its role has been recognized by the 2013 Nobel Prize in chemistry awarded to Michael Levitt (2014), Martin Karplus (2014), and Arieh Warshel (2014) for “taking the chemical experiment into cyberspace” (Staffan Normark, Permanent Secretary of the Royal Swedish Academy of Sciences).

Why do we use computational molecular mechanics and/or molecular dynamics methods for studies of structure and protein dynamics? Here is a partial list of good reasons for their popularity and widespread applications:

1. Give models for mechanistic interpretation of protein functions.
2. Help to connect a protein structure and dynamics with its functions.
3. Provide time-resolved data on protein structures.
4. Generate visual models easy to manipulate (great pictures in papers).
5. Promise to determine intra- and intermolecular interactions quantitatively.
6. Help to discover subtle functional differences in protein architectures.
7. Are useful in drug design.
8. Provide an easy tool to study protein mutants.
9. Allow for checking catalytic properties.
10. Are easy to verify, extend, or modify.
11. Bring information on thermodynamics.
12. Provide data on mechanical strength.
13. Direct comparisons with experimental data such as NMR or X-ray structures, single-molecule AFM spectroscopy, etc. are possible.
14. Have a strong impact on biochemistry, biology, biophysics, and bioinformatics.
15. Cheap computers are available everywhere.
16. Good quality, free software is accessible via WWW.
17. A relatively short “learning curve.”
18. Allow for computer experiments in fully controlled conditions.
19. Provide “molecular big data” suitable for data mining.
20. The field of computer modeling is well established and has many experts, conferences, specialized journals, and university courses around.
21. Supercomputer power doubling every 18 months.
22. Strong support from the industry and governments, biopharma sector, medical agencies, computer makers, etc.

And, of course, the last but not the least reason for the great popularity of computational modeling of proteins is just human curiosity and interest in solving

scientific problems. For example, solving the protein folding problem is still the great challenge to scientific community. We know over 100 million protein sequences, but only some 100,000 3-D structures are available in the Protein Data Bank (as of year 2015). It is commonly believed that finding a realistic and sturdy method of converting 1-D protein sequence into 3-D structure should bring another Nobel Prize to the authors of such a discovery.

The purpose of this chapter is twofold: (a) we want to draw attention to major review papers and resources regarding studies of molecular dynamics of proteins and (b) we want to point out the most representative, current applications of MD methods which have been published in recent years.

Formalism of Molecular Mechanics and Molecular Dynamics Methods: A Short Presentation

In order to describe atomistic level chemical phenomena, one has to use quantum mechanics (Dahl 2001; Greenberger et al. 2009). This theory applied to molecules brought “an explosion” of quantum chemistry and together with the computer revolution has changed the way the chemistry is done in the twenty-first century. Despite great successes of quantum chemistry, this approach is still limited to small- or medium-size systems, having 20–300 atoms (Piela 2014). Proteins usually are composed of thousands of atoms; moreover, their realistic models should take into account a substantial number of water molecules. Thus, classical models of proteins are necessary. In such a simplified approach, electrons are neglected, quantum bonds are replaced by effective analytical potentials, and instead of atoms, we have carefully designed model balls. The motion of model atoms is treated in a classical way: validity of Newton equations of motion is assumed in the proteins’ nanoworld. Structures of molecules may be optimized using classical force fields and methods of molecular modeling; the time evolution of a position of each individual atom may be followed for a reasonably long period of time using the molecular dynamics formalism, but a possibility of forming/breaking chemical bonds is in classical simulations basically lost. Thus, the real chemistry, based on chemical reactions, needs the quantum theory. Nevertheless, the classical modeling provides so many useful hints into the nature of biomolecules that computational modeling of protein structure and dynamics is a fundamental part of the research in all major laboratories. Current technology allows for performing at least 35-ns classical MD simulations for 10,000,000 atoms in one computer day (Andoh et al. 2013).

Some History

The Electronic Numerical Integrator And Computer (ENIAC) was built at the University of Pennsylvania in Philadelphia in 1945. It weighed more than 27,000 kg (60,000 lb) and contained more than 18,000 vacuum tubes. It is regarded as the

first successful digital computer. Already in 1955, E. Fermi, J. Pasta, and S. Ulam calculated numerically the motions of a chain of coupled anharmonic oscillators. This is probably the first example that molecular dynamics simulation suggested analytical solution for the problem studied. In 1957, Alder and Wainwright studied dynamics of a system of hard, two-dimensional discs (Alder and Wainwright 1957). In 1971, Rahman and Stillinger used a more realistic Lennard-Jones interaction potential to study motions in water (Rahman and Stillinger 1971). The duration of those simulations was about 10 ps.

The concept of interaction force field has been applied to proteins by S. Lifson in 1969. This model was aimed at facilitating refinement process of protein structure determination based on X-ray diffraction experiments (Levitt and Lifson 1969). The role of quantum electronic effects in enzymatic reactions has been recognized later as well in the form of QM/MM approach (Warshel and Levitt 1976). Current force fields – sets of analytical formulas together with carefully chosen parameters – are based on infrared spectra, geometrical information from amide crystals, and quantum mechanical calculations. The famous paper by McCammon et al. (1977) on BPTI dynamics may be regarded as the first atomistic MD modeling of a protein. Interesting stories on the history and development of molecular dynamics simulation science may be found in the review published by M. Karplus (2003) and the article by W. Jorgensen (2013). At the end of 2015 in the PubMed Bibliographics database, nearly 20,000 papers were listed under the query “molecular dynamics simulation protein.” This number has doubled since 2011. The area of computational chemistry and biology is very strong and is fast growing.

On the Origin of Potential Energy Surface (PES) Concept

As we know, on the grounds of quantum mechanics, the complete information of the molecular system may be obtained from the wave function $|\Phi\rangle$, which is a solution of the Schrodinger equation:

$$\hat{H} |\Phi\rangle = \varepsilon |\Phi\rangle$$

where Hamiltonian \hat{H} contains both electron (i,j) and nuclear (A,B) degrees of freedom (in atomic units):

$$\begin{aligned} \hat{H} = & -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 & \text{atomic units (a.u.):} \\ & - \sum_{i=1}^N \sum_{j=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{i>j}^N \frac{1}{r_{ij}} + & \hbar = 1 \\ & + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} & "c = 1" \quad (137.) \\ & & m = 1 \\ & & e = 1 \end{aligned}$$

If we adopt Born-Oppenheimer approximation, a separation of electronic and nuclear degrees of freedom is possible:

$$\widehat{H}_{el} \Phi_{el} = \varepsilon_{el} \Phi_{el}$$

where electronic Hamiltonian is

$$\widehat{H}_{el} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{i>j}^N \frac{1}{r_{ij}}$$

and electronic wave function Φ_{el} depends in a parametric way on positions of all nuclei $\{\vec{R}_A\}$:

$$\begin{aligned} \Phi_{el} &= \Phi_{el} \left(\{\vec{r}_i\}; \{\vec{R}_A\} \right) \\ \varepsilon_{el} &= \varepsilon_{el} \left(\{\vec{R}_A\} \right) \end{aligned}$$

As one can see, the electronic energy ε_{el} depends on positions of nuclei as well. Thus, for each different arrangement of atoms, Φ_{el} is a different function of electron coordinates \vec{r}_i . For a fixed position of atoms (nuclei), the total energy reads

$$\varepsilon_{tot}^{fix N} = \varepsilon_{el} + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}$$

Such an approximation leads to the following nuclear Hamiltonian \widehat{H}_{nucl} :

$$\begin{aligned} \widehat{H}_{nucl} &= -\sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + \langle \widehat{H}_{nucl} \rangle_{\Phi_{el}} + \sum_A^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} = \\ &= -\sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + \underbrace{\varepsilon_{el} \{R_A\} + \sum_A^M \sum_B^M \frac{Z_A Z_B}{R_{AB}}}_{\varepsilon_{tot}^{fix N} \{R_A\}} = \\ &= -\sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + \varepsilon_{tot}^{fix N} \{R_A\} \end{aligned}$$

This total energy $\varepsilon_{tot}^{fix N}$ (with the fixed positions of atoms) provides a potential for nuclear motion. Such quantity is called potential energy function (called sometimes potential energy surface PES) and in molecular dynamics simulations is approximated by a force field. In terms of quantum mechanics solutions to the nuclear Schrödinger equation,

$$\widehat{H}_{nucl} \Phi_{nucl} = \varepsilon_{nucl} \Phi_{nucl}$$

provides information on vibrations, rotations, and translations of the molecule. But the dynamics of a molecule (protein) may be studied also classically, provided that the realistic molecular PES is known.

Force Fields

Classical force fields (FFs) have common assumption that the analytic potentials are good approximation to the “real” PES. It is also assumed that parameters pertaining to particular atom types (or groups of atoms) are transferable from one model biomolecular system to similar ones.

The general expression for potential energy V is the following:

$$V = \sum_{\text{bonds } i} V_i^B + \sum_{\text{bond angles } j} V_j^A + \sum_{\text{torsional angles } k} V_k^D + \sum_{\text{improper angles } l} V_l^E \\ + \sum_{\text{atom pairs } (r,s)} (V_{r,s}^C + V_{r,s}^P + V_{r,s}^{VdW})$$

The physical meaning of each term is described in the legend in Fig. 1. In general, an intuitive decomposition of energy is used in the above equation. Additive components correspond to energies of chemical bonds (V_i^B), energies related to angular deformation of the optimal molecular structure (V_j^A , V_k^D , V_l^E), and pairwise interactions of all atoms present in the system related to Coulomb energy of a system of partially charged model atoms (V^C) and so-called Lennard-Jones (or van der Waals) term ($V^P + V^{VdW}$).

One should distinguish between all-atom and united-atom force fields. In this second type, the groups of real atoms are used as basic model “atoms,” for example, the methyl group CH_3 may be treated as a special type of a united atom having molecular weight of 15. Such simplification saves some computer time and in certain cases does not affect the results of modeling. The drawback of the majority of force fields currently used is the lack of atomic polarization effects. Partial changes on atoms to some extent include polarization, but usually, parameters remain fixed through the whole simulation. But the reality is different. Induced dipole moments may change the dynamics of the molecule and may affect results of modeling. Charges are modified “on the fly.” Thus, the next generation of force fields shall include atomic polarizability (Warshel et al. 2007; Huang et al. 2014).

There are numerous classical force fields designed for the description of protein structure and dynamics. Among the most popular ones are CHARMM (MacKerell et al. 1998), AMBER (Weiner et al. 1984), GROMOS05 (Christen et al. 2005), and OPLS (Jorgensen and Tirado-Rives 1988). Perhaps, the best up-to-date account of various force fields and computer codes available for protein simulation is the Wikipedia on the Internet. The problem of the design of a potential energy function for proteins is discussed in Boas and Harbury (2007). Some papers concerning

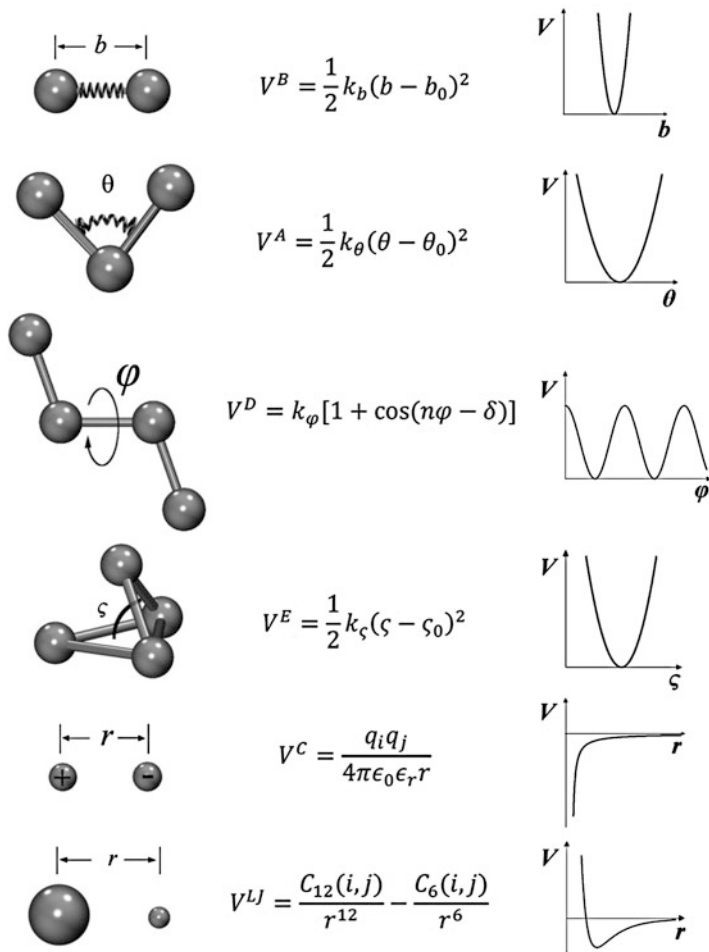


Fig. 1 Based on H. Grubmueller, “Proteins as Molecular Machines: Force Probe Simulations,” published in Computational Soft Matter: From Synthetic Polymers to Proteins, Lecture Notes, Norbert Attig, Kurt Binder, Helmut Grubmueller, Kurt Kremer (Eds.), John von Neumann Institute for Computing, Julich, Germany, NIC Series, Vol. 23, ISBN 3-00-012641-4, pp. 401–422, 2004. © 2004 by John von Neumann Institute for Computing. Interaction contributions to a typical force field. Bond stretch vibrations are described by a harmonic potential V^B , the minimum of which is at the equilibrium distance b_0 between the two atoms connected by chemical bond i (the indices i, j , etc. are not shown in the figure). Bond angles and out-of-plane (improper) angles are also described by harmonic potential terms, V^A and V^E , where Θ_0 and ζ_0 denote the respective equilibrium angles. Dihedral twists (torsional angles) are subjected to a periodic potential V^D ; the respective force constants are denoted by k 's with appropriate indices. Nonbonded forces are described by Coulomb interactions, V^C , and Lennard-Jones potentials, $V^{LJ} = V^P + V^{vdW}$, where the latter includes the Pauli repulsion, $V^P \sim r^{-12}$, and the van der Waals interaction, $V^{vdW} \sim -r^{-6}$, respectively

comparison of the quality of various force fields have appeared (Hornak et al. 2006; Guvench and MacKerell 2008; Lindorff-Larsen et al. 2012a). It is obvious that each force field was optimized for a different set of systems and they were tuned to reproduce different properties, so it is not clear how a fair comparison should be performed. On the other hand, in the protein modeling community, there is continuous tension and pressure that more effort should be devoted for careful scrutinizing of the force fields and for setting clear recommendations regarding “justified” error bars for results of computations (Aliev and Courtier-Murias 2010; Pantelopulos et al. 2015).

Structure Optimization or Energy Minimization

The most standard problem of molecular mechanics applied to proteins is the following: what is the best (optimum) structure of the studied protein? The common assumption is that the minimization of the energy (or sometimes other thermodynamic functions such as free energy) will give the answer. It is believed that such a structure is a good approximation to the native structure of the protein of interest.

There are many computational methods of finding local minima (Klepeis et al. 2003; Chou 2004; Kmiecik et al. 2007). Unfortunately, there is no general method known which might lead to the global minimum of PES of a given protein. Such techniques as simulated annealing (Kannan and Zacharias 2009) or replica exchange (Sugita and Okamoto 1999; Sugita 2009), supported by genetic algorithms, may help to sample the conformational space efficiently and may generate plausible “candidates” for the global minimum, but these approaches are rather heuristic recipes than rigorous procedures. Quite often, the energetics of transition from a conformation A to a conformation B is of interest and should be calculated (Schlegel 2003). There are numerous methods of finding classical reaction paths appropriate for proteins, for example, the self-penalty walk (Nowak et al. 1991). It is worth noting a tricky technique of milestoning (Kuczera et al. 2009). The interested reader may find details of reaction path calculations for large molecular systems in Elber et al. (2002) or Tao et al. (2012).

General Molecular Dynamics Scheme

The basic physics behind a simple molecular dynamics scheme seems to be trivial. We want to solve simultaneously M Newton’s equation of motion for each individual atom i (or a grain of atoms):

$$\mathbf{a}_i = \frac{1}{m_i} \mathbf{F}_i, \quad i = 1, \dots, M$$

The force \mathbf{F}_i acting on each atom may be calculated (locally) as gradient of the potential (PES):

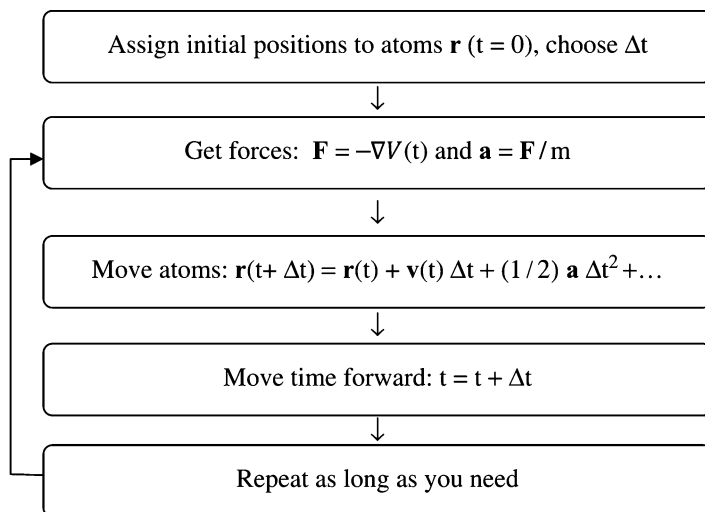


Fig. 2 A general scheme of generating MD trajectory

$$\mathbf{F}_i = -\nabla V$$

Having forces, one can easily have accelerations acting on each atom. We assume that the molecular system is deterministic (not always the case), and we may predict the position of all atoms by integrating the acceleration (twice) with respect to time. There are many numerical algorithms suitable for solving these problems, but due to its simplicity and numerical stability, Verlet algorithm is perhaps the most frequently used in real simulations (Frenkel and Smit 2002). Of course, one must assume the initial structure of the protein and has to plug in time into numerical algorithm. Time variable is discretized; the size of the time step depends on the timescale of the fastest motions we want to study. For real proteins in ambient temperatures, it is usually 1 fs (10^{-15} s). Figure 2 shows major steps in MD routines.

Here, we present typical steps in doing molecular dynamics simulations of proteins:

1. What is the scientific problem we want to solve?
2. Is classical MM/MD modeling an appropriate methodology for this type of problem? Would these calculations help to understand nature, indeed? How our results might be verified experimentally?
3. Do we have enough knowledge/experience/expertise or we prefer “black box” approach (“let’s calculate something, and we will see . . .”)?
4. What is the expected time of calculations required to obtain reasonable, publishable results? (Note: one trajectory usually is not enough.)
5. Do we have sufficient resources (licensed codes, computer time, storage space, manpower, etc.)?

6. Having answered points (1–5), we need to set up an initial model of the protein. Usually, an experimental structure (X-ray or NMR) downloaded from the PDB www.pdb.org (Berman et al. 2000) is a good starting point.
7. We should check whether the structure is complete. Is the resolution of the structure adequate for our purposes? Does it contain all amino acids? Are there any missing atoms? Do we have force field parameters for prosthetic groups, ligands, metal ions, and exotic stuff present in our favorite protein? One should always check the “biological unit” entry in PDB to avoid misunderstanding of the real structure of a protein of interest.
8. We need to decide in what particular environment simulation will proceed. Vacuum? Water? Continuum dielectric? Exotic solvent? pH? Should ions be added for charge neutralization?
9. In what conditions of temperature and pressure we plan simulations? How the temperature will be maintained? Shall we switch to the Langevin dynamics?
10. Do we need quasi-infinite model (periodic boundary conditions – PBC)? What is an appropriate shape and size of the solvation box? Should we use the Ewald summation technique to properly calculate electrostatic interactions?
11. Let us assume we perform a standard 50-ns MD simulation of a protein having a reasonable initial structure. We will add a box of model water molecules at least 6 (or 9) angstrom thick at each protein border region.
12. After this initial preparation, the first step will be an optimization of the protein structure. We may initially freeze the protein and allow for some steps of minimization of solvated water molecules (e.g., 500 steps of the steepest descent (SD) method).
13. Next, we may allow for some MD simulations of the solvent (water). We may gradually increase the temperature of waters from 0 K (minimized structure in principle is not related to any temperature) to 300 K, for 500–1,000 ps period of time. The time step in MD is usually 1 fs; thus, we will ask the MD code to perform up to 1,000,000 steps.
14. We may relax constraints and “unfreeze” the protein. Some 200–500 SD steps should be sufficient to transfer the protein (+water) from the “experimental” minimum to a “local force field-related” minimum. Too many steps result in overminimization (Rydzewski et al. 2015a)!
15. We may increase the temperature in a stepwise manner from 0 to 300 K for the whole system. 1–5 ns of such heating phase is often more than enough.
16. If we perform $T = \text{const}$ simulation, we need to equilibrate the system well before useful data may be collected. The equilibration time t_{eq} depends on the system studied. In the current papers, one can find t_{eq} from 1 to 10 ns. We should observe at least a few geometrical parameters of our protein. The RMS distance calculated for C_{α} atoms from the minimized (or PDB) structure to the current one may help to estimate whether the model is fully equilibrated. The RMS plot versus time should be flat.
17. Now, we may use the equilibrated system (protein + water) and launch a long production run (50 or more up to 1,000 ns).

18. The structures at selected time point (frames) should be stored. These structures are further analyzed using computer graphics and specialized software analysis tools. One may store structures every 100 steps (fs) or even every 1,000 steps or more; it depends on what data are needed.
19. In the theory, we should set up modeling for infinite time in order to sample all configuration space of the protein studied. This is obviously not possible; thus, we cross our fingers and believe in ergodicity of the trajectory obtained. Once calculated quantities do not depend on time of simulation, it is a reasonable signal that the longer calculation will not bring new information. Usually, several shorter trajectories, with different initial conditions, will provide better understanding of the protein than one but very long trajectory.

During MD simulations, temperature and pressure have to be controlled using special algorithms, and infinite systems may be mimicked by using periodic boundary conditions. Long-range electrostatic interactions are often accounted for by using the Ewald summation technique (Sagui and Darden 1999). Inclusion of stochastic character of atomic motion results in using Langevin dynamics instead of Newton dynamics (Schlick 2010).

One should note that solvation effects may be accounted for not only by including into the studied system explicit water molecules but also via implicit solvent models (Chen and Brooks 2008; Chen et al. 2008). Changes in free energy upon solvation are often estimated using the generalized Born model (Hou et al. 2011; Cumberworth et al. 2015).

Practical Aspects

Even the most sophisticated and advanced methods (Schwede and Peitsch 2008) need tools to perform computations. Both computer codes and hardware are required. The majority of papers published in this field utilized public domain codes to obtain data on protein dynamics. There are of course numerous commercial packages, such as Discovery Studio (Biovia), Sybyl-X (Certara), Yasara (YASARA Biosciences), and Desmond/Maestro (Schrödinger Inc.), popular in industry and certain research environments, but it seems that routine academic work is based on basically freely available software.

MD Codes

One of the first codes was CHARMM (commercial version is called CHARMM) developed originally at Harvard University. This suite of programs is very versatile and contains all major computational methods. It is parallelized and allows also for QM/MM simulations (Brooks et al. 2009). AMBER force field is popularly used for modeling of nucleic acids, but thousands of simulations of proteins have been published as well (Case et al. 2005; Salomon-Ferrer et al. 2013). The most recent,

14th, version of AMBER is heavily optimized with respect to performance (GPUs) and contains such advanced techniques as locally enhanced sampling and SCC-DFTB QM/MM methods.

In Europe, the GROMACS code has growing popularity due to its speed and good scaling on parallel clusters (Pronk et al. 2013; Van Der Spoel et al. 2005). The methods of essential dynamics, principal component analysis, and flooding (Lange et al. 2006) are available here. The most recent version – GROMACS 5.1 – takes advantage of multicore processors present in modern PCs and workstations and runs not only under many Unix distributions but Windows as well (Pronk et al. 2013). Both CUDA and OpenCL graphics card environments are supported (Abraham et al. 2015).

In our lab, we are quite happy with the NAMD code having 40,000 registered users (Phillips et al. 2005). It is well documented, relatively fast, well maintained, and often updated and scales nicely. It has certain flexibility in selection of the force field (CHARMM, AMBER). The authors have implemented locally enhanced sampling (LES), implicit ligand sampling (ILS), replica exchange, and steered molecular dynamics (SMD) schemes. New versions of NAMD (2.11 as of 2015) will run effectively on GPUs; there are also attempts to port this code to a computational grid environment.

This short presentation of major packages devoted to protein dynamics simulations is very far from being complete. Some WWW services, including Wikipedia, try to maintain the updated lists of available MD codes.

Developments in Force Fields and Parametrization

MD data may give insights into chemical or biological phenomena only if calculations are based on good-quality force fields and adequate parameters. Thus, much effort is devoted to testing and development of advanced force fields (Lindorff-Larsen et al. 2012a; Pantelopulos et al. 2015). Some researchers argue that NMR measurements provide objective measures of quality of FFs (Robustelli et al. 2010; Beauchamp et al. 2012; Huang and MacKerell 2013), while others warn that such an approach is not necessarily correct (Martin-Garcia et al. 2015).

Subtle intermolecular interactions often depend on deformations of atomic electronic densities. To account for such effects in a classical FF, atomic polarizability has to be added. Calculations with such polarizable FF are more time-consuming, but new phenomena may be studied using such more advanced protein models. The additive and polarizable variant of CHARMM FF has been very recently published (Vanommeslaeghe and MacKerell 2015). Other advances in polarizable force fields, with special attention devoted to longtime simulations, are discussed in Huang et al. (2014). Notably, a great deal of efforts is devoted to construction of new force fields suitable for modeling of medically important protein-surface interactions (Martin et al. 2015). Also efforts to optimize implicit solvent FF (Bottaro et al. 2013) are promising since longer and longer simulations are required, for example, in protein

folding studies. Comprehensive discussion of current status of FF developments may be found in the review by Lopes et al. (2015).

Programming user-friendly procedures of generating new parameters enhances area of MD applications. For example, a practical way to generate parameters (CHARMM, GROMACS) required for protein simulations with small organic ligands is to use the SwissParam server (Zoete et al. 2011). Another option of extending CHARMM MD simulations to new ligands offers the tool kit fTK implemented into newer versions of VMD (Mayne et al. 2013). For GROMACS users, an automated topology builder (Malde et al. 2011) may facilitate a lot of routine MD research on new proteins. Some hints of building new parameters may be found here (Wang et al. 2014).

Visualization

In the era of the Internet and efficient graphics cards, we have tens of programs designed to visualize a protein structure. Each year brings new players in this competition. Many researchers prefer to pay license fee and to use professional graphics for visualization of their structure and MD data: Chimera, PyMOL, Discovery Studio, Maestro, HyperChem, Yasara, etc. Open software is also available, for example, OVITO (Alexander 2010) or Avogadro (Hanwell et al. 2012). A comprehensive list of currently used molecular visualization codes may be found in the book by Gu and Bourne (2009).

User-friendly and robust code evolved from the early Visual Molecular Dynamics (VMD) software, created by K. Schulten group at UIUC in the USA (Humphrey et al. 1996). There are thousands of users of this package. An example of VMD visualization of the extracellular matrix synaptic protein reelin is presented in Fig. 3. Preparation of a similar picture is not difficult if one uses critically (Knapp and Schreiner 2009) the VMD guide (Hsin et al. 2008).

Review of Reviews

Since the very first proceedings of CECAM workshop in France devoted to models for protein dynamics (Berendsen 1976), details of computer simulations of physical and biological systems have been described in many comprehensive books. For example, a good starting point may be the book by Allen and Tildesley published already in 1987 (Allen and Tildesley 1987). It is devoted mainly to simulations of liquids, but many methodological aspects of computational modeling of physical systems are well covered. Details of molecular simulations are presented by Haile (1992). In a classical text by Rapaport (1995), explanations of basic software and algorithms may be found. The introductory but very informative book by Frenkel and Smit focuses on Monte Carlo and molecular dynamics methodologies. The authors analyze algorithms and present useful FORTRAN-based pseudocodes for basic steps of simulations (Frankel and Smit 2001). Broader aspects of molecular

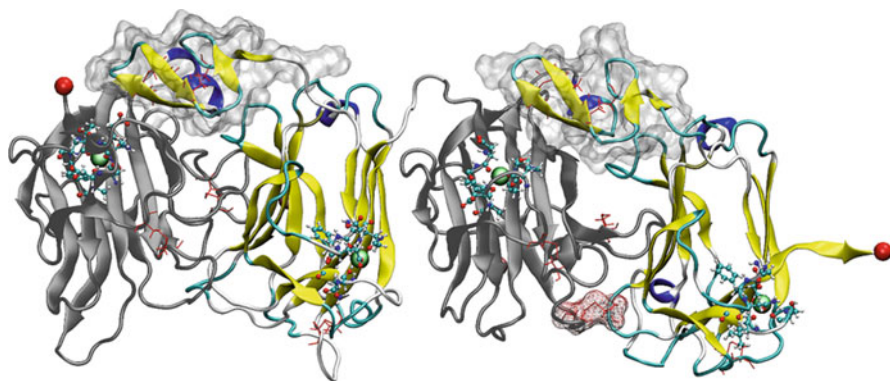


Fig. 3 A fragment of extracellular matrix protein reelin (2ee2 PDB code, two out of eight BNR-EGF-BNR repeats are shown) that plays a role in a synaptic plasticity and maintenance of synaptic function, visualized using the VMD code (Humphrey et al. 1996). NAG cofactors are shown in ball-and-stick representation, green spheres represent Ca^{+2} ions, the surface denotes EGR parts, and a disulfide bridge critical for mechanical resistance is shown in red mesh. Abnormal reelin expression is observed in autism spectrum disorders, schizophrenia, or Alzheimer's (author: K. Mikulska, UMK, Poland)

modeling are covered by A. Leach (2001). In this book, besides main algorithms and methods of computational chemistry and modeling methods of protein structure prediction, free energy calculations, solvation, and drug design applications are presented. The book coauthored by M. Karplus, one of the founding fathers of MD simulations of proteins, is a valuable source of information for everyone interested in protein dynamics (Becker and Karplus 2006).

Excellent reviews on various aspects of biomolecular modeling are published quite often. Here only a very concise, subjective, and limited review of the recent (i.e., published in the twenty-first century) reviews is presented, just to provide a handy reference to further search for relevant information.

A large body of proteins perform catalytic functions. MD calculations of enzymatic mechanisms are a great challenge to theory, and the best strategy for simulations is a matter of continuous debate. In reviews by A. Warshel et al. (Warshel 2002, 2003), main aspects of proper understanding of catalysis are described. Modeling of chemical reactions requires a special approach – some possibilities are outlined in van Speybroeck and Meier (2003), and a comprehensive review of computational enzymology, largely based on QM/MM methods, may be found in Lonsdale et al. (2010). Yet another class of problems arise when protein-protein interactions are modeled (Ritchie 2008).

Basic methods and main applications of computer modeling of biosystems are presented in a comprehensive work by Schlick (2010), a review by Goodfellow et al. (Moraitakis et al. 2003), a paper by van Gunsteren et al. (2002) (Hansson et al. 2002), or the review by K. Kremer (2003). Historical development of MD techniques is presented by Field (2015). Protein modeling field is outlined by

Naray-Szabo et al. in the previous 2012 edition of the *Handbook of Computational Chemistry* (Leszczynski 2012) and in more recent comprehensive and general review by Orozco (2014). Useful accounts on are presented in a recent review by Lorenz and Doltsinis (2012). Particularly interesting are studies of membrane proteins (Gumbart et al. 2005; Stansfeld and Sansom 2011a, b) since receptor proteins are common and attractive drug targets (Borhani and Shaw 2012). One can also find reviews dedicated to application of simulations in narrow subdisciplines, such as biotechnology (Aksimentiev et al. 2008). It is worth noting that more integrative approach, combining MD simulations, experimental data, and network analysis, is advocated as a tool to get an understanding of biological process in the cell (Papaleo 2015).

An account on classical methods, problems, and goals of biomolecular simulations is given in a comprehensive article by van Gunsteren et al. (2006). Focused mainly on proteins, a very informative paper by Adcock and McCammon provides an excellent description of methods and key results from MD (Adcock and McCammon 2006). Methodological advances were also reviewed at the same time: Chu et al. popularized multiscale simulations (Chu et al. 2007), Elber et al. critically commented the literature on longtime simulation methods (Dal Peraro et al. 2005), and Liwo et al. enumerated efficient methods of sampling of proteins' conformational space (Liwo et al. 2008). Newer review of enhanced sampling techniques has been published by Fujisaki et al. (2015). Coarse-grained MD simulations have a growing importance in computational biochemistry and biology. Numerous excellent reviews have been published recently (Marrink and Tieleman 2013; Saunders and Voth 2013; Kmiecik et al. 2014; May et al. 2014; Vicatos et al. 2014; Ingolfsson et al. 2014; Barnoud and Monticelli 2015).

Enhanced sampling of conformational space has been always promising keyword for progress in MD simulation methods (Spiwok et al. 2015; Fujisaki et al. 2015). Collective variables may reduce the space (Vashisth et al. 2014), but typically metadynamics, replica exchange, or Gaussian accelerated MD and their variants are extensively explored (Granata et al. 2013; Abrams and Bussi 2013; Comer et al. 2014; Andersen et al. 2015; Miao et al. 2015).

A lot of effort has been put in elaborating reliable and practical methods of calculating changes in free energies (Kholmurodov et al. 2003; Meirovitch 2007; Christ et al. 2010; Pohorille et al. 2010). Even "computational alchemy" term has been coined for some counterintuitive but physically valid methods (Straatsma and McCammon 1992; Aleksandrov et al. 2010). Such data are required, for instance, in drug design (Galeazzi 2009; Morra et al. 2008). Increasing role of taking into account entropic effects is now widely accepted (Noel and Whitford 2014; Rydzewski et al. 2015a). Enhanced MD helps now to access better ligand-protein docking (Andersen et al. 2015). Time-consuming free energy calculations may profit a lot from using enhanced sampling techniques (Bernardi et al. 2015; Miao et al. 2015) as well.

Advantages of using the computational approach to chemistry and biology are outlined in a comprehensive review of Mulholland et al. (van der Kamp et al. 2008). Supercomputers contribute a lot to the current blooming state of MD simulations

of large biosystems (Chipot 2015). One can find also more focused, specialized reviews of MD approach to biological problems (Dodson et al. 2008; Avila et al. 2011) or to future nanotechnology applications of self-assembling systems (Klein and Shinoda 2008), biotechnology (Aksimentiev et al. 2008), and drug design (Borhani and Shaw 2012). Quite often modern simulation techniques are called “computational microscope,” just to stress its role in molecular biology (Dror et al. 2012). Interestingly enough, with the petaflop computer system currently available, multimillion atom simulations are possible (Sanbonmatsu and Tung 2007). Promising efforts are made to construct computers dedicated to MD simulations on hardware level (Scarpazza et al. 2013). The success of ANTONs from D.E. Shaw Research is notable (Shaw et al. 2014).

Milliseconds in simulations of protein folding problems are currently sought (Bowman et al. 2011; Piana et al. 2014); thus, all means helping to reach this target are welcome (Lane et al. 2013), including optimization of MD codes (Krieger and Vriend 2015). Protein folding phenomena still pose more questions than answers (Dill and MacCallum 2012).

For sure, this field profits from the use of the Internet. Some groups have developed user-friendly and even charming portals to the simulation software (Miller et al. 2008). Folding@Home project has reached 20-Tflop computer power (Lane et al. 2013), and this software is available for other distributed computing environments now (Eastman and Pande 2015). Other researches, for example, the group of V. Daggett, invest in large-scale depositories of scientific data stemming from MD simulations (Van der Kamp et al. 2010). Sharing data, especially obtained using advanced computer resources and easy to expose in clouds, is always a good idea. Such an “ocean” of numbers deserves careful scrutiny. Hopefully, some scientific treasures will be fished out in the near future.

Selected Examples of Applications of MD to Study Proteins

In this short chapter, we point out selected, representative, widely discussed problems where the MD methods have been applied.

Protein Folding Studies

Proteins are synthesized as linear polymers, but for their function, precise 3-D structure is usually necessary. The process of folding to such a native structure may be conveniently studied using molecular dynamics simulations (Fersht and Daggett 2002; Zhang et al. 2009; Freddolino et al. 2010; Towse and Daggett 2013). The main obstacle in this research is still relatively limited (100–1,000 ns) timescale accessible for standard modeling (Scheraga et al. 2007). However, more and more benchmark calculations achieve microseconds (Piana et al. 2011), including classical studies for very small systems such as villin headpiece fragment, one of

the most stable and fastest-folding naturally occurring proteins (Duan and Kollman 1998; Ensign et al. 2007; Freddolino and Schulten 2009) or ubiquitin (Piana et al. 2013). Simulations help to propose universal folding mechanisms and to determine intermediates (Freddolino et al. 2008; Rizzuti and Daggett 2013; Piana et al. 2014). Better algorithms and huge computer power make performing meaningful folding simulations possible within a few days (Nguyen et al. 2014). Quite promising are new methodological developments adding core physical insights into a standard Replica Exchange MD (REMD) (Perez et al. 2015) or even mixing MD with experimental data (Sborgi et al. 2015). A lot of effort is devoted to finding an intelligent method of “ab initio” protein folding (Shakhnovich 2006; Blaszczyk et al. 2013). The progress may be checked by following the worldwide folding competition CASP (Kryshtafovych et al. 2016).

Intrinsically Disordered Proteins

Not all proteins are nicely folded, and numerous proteins remain disordered through the whole (or majority) of their life-span in a cell. The paradigm that only well-structured biomolecules may play significant physiological role has been abandoned at the beginning of the twenty-first century (Wright and Dyson 1999). The role of intrinsically disordered protein (IDP) is intensively studied (Wright and Dyson 2015) and opens very attractive field for MD simulations (Lindorff-Larsen et al. 2012b; Baker and Best 2014; Fu and Vendruscolo 2015). Notably, there is an ongoing discussion of what MD force field (Baker and Best 2013; Rauscher et al. 2015; Piana et al. 2015) and methodology (Do et al. 2014; Granata et al. 2015; Zerze et al. 2015) are the best for IDP modeling.

Protein-Drug Interactions and Docking

The pharmaceutical industry badly needs reliable theoretical methods for calculating ligand binding affinities (Aqvist et al. 2002; Gallicchio and Levy 2011; Zhao and Caflisch 2015; Yuriev et al. 2015). The problem is not easy, since many factors, for example, multiple binding sites, protein flexibility, and solvent model, affect a small value of the free energy of binding (Simonson et al. 2002; Deng and Roux 2009; Spyraakis et al. 2011). The chance of getting wrong results is high. However, in the literature, there are hundreds of papers devoted to protein-drug interactions. One of the main topics is HVI-related enzyme inhibitors (Monroe et al. 2014). The threat of a bird or swine flu pandemic triggered studies of interactions of neuraminidases or influenza A peptides with antiviral drugs (Le et al. 2011; Khurana et al. 2011). The role of induced fit effects during ligand docking to steroid hormone-binding receptors was analyzed by Cornell and Lam (2009). MD may be used in an anticancer drug development (Rosales-Hernandez et al. 2009; Lauria et al. 2010), GPCR inhibitors/activators (Dror et al. 2013; Tautermann et al. 2015), or studies of inhalational anesthetics' interactions with proteins (Vemparala et al. 2010). Studies

on microsecond timescale MD explained mechanisms of chemokine receptor CCR5 inactivation (Salmas et al. 2015), macrolide binding to the ribosome (Sothiselvam et al. 2014), or opioid receptor activation (Huang et al. 2015). The review of MD applications in drug design, especially in stability of drug-macromolecule complexes, has been recently published (Mortier et al. 2015).

Spectroscopy Experiments

Spectroscopy is an extremely useful analytical and diagnostic technique with wide applications in chemistry, physics, life sciences, industry, medicine, etc. Molecular dynamics helps to interpret experimental data; some examples are given below. Sen et al. used simulations to explain time-resolved Stokes-shift experiments with biopolymers (Sen et al. 2009). Fluorescent proteins, especially based on GFP, are often studied computationally. Sun et al. applied QM/MM and MD to explain dependence of red fluorescent protein on pH of the environment (Sun et al. 2010). Computational studies of energy transduction in photoactive yellow protein may give useful hints for spectroscopic studies (Gamiz-Hernandez and Kaila 2016) as well as do the modeling of visual pigments (Brunk and Rothlisberger 2015; Wanko et al. 2006). Quite often, the interpretation of NMR experiments profits from simulations, such as studies of lipid-binding sites to neurotoxin (Weber et al. 2015) or evaluation of rotational diffusion constants from MD (Wong and Case 2008). On the other hand, the NMR spin relaxation data help to improve force fields (Beauchamp et al. 2012).

Functionally Important Motions (FIMs)

Some motions of proteins are critical for their proper functioning (Henzler-Wildman and Kern 2007). MD simulations may identify such modes. Relation of protein's mechanics and function was reviewed in 2003 by Schulten (Tajkhorshid et al. 2003), but since that time, this group studied new problems, for example, plant phototropism (Freddolino et al. 2006b) and complete satellite *Tobacco mosaic virus* vibrations (Freddolino et al. 2006a). Other groups investigated large-scale motions in biosensors (Tatke et al. 2008), linker motions crucial for ligase (Liu and Nussinov 2010), or retinal release from opsin (Wang and Duan 2011). Extraction of information on FIM needs special methodology (Schuyler et al. 2009), such as essential dynamics (Amadei et al. 1993) or metadynamics (Biarnes et al. 2011); thus, new ways of MD data analysis are suggested (Hub and de Groot 2009; Vuillon and Lesieur 2015). Critical assessment of FIM analysis methods may be found in Moradi and Tajkhorshid (2014). The main global motions are encoded in protein structures, and these features may be explored in protein ligand studies using normal mode analysis (Bahar et al. 2015). A very interesting step toward systematic classification of dynamical properties of proteins called dynasome has

been proposed in paper (Hensen et al. 2012). Studies of over 100 distinct proteins show that internal mobility patterns exhibit rather continuous distribution. The strong correlation between structure and dynamics has been, however, noted.

Molecular Machines

Having such powerful computers at hand, we are ready to study molecular machines – proteins or bio-complexes that perform some mechanical work during their activity cycle (Scheres 2010; Elber and Kirmizialtin 2013). Rotations of parts of ATPase were studied by Ma et al. already in 2002 (Ma et al. 2002) and molecular rotation in ATP synthase by Aksimentiev et al. in 2004 (Aksimentiev et al. 2004). Solvent-induced lid opening in lipases was also analyzed computationally (Rehm et al. 2010). Even the whole cellular mechanics simulations were reviewed by Gao et al. (2006). In recent years, a flux on papers in this fascinating phenomenon of molecular machines has been observed (Kutzner et al. 2011; Sanbonmatsu 2012; Bock et al. 2013; Czub and Grubmuller 2014; Ito and Ikeguchi 2014; Mukherjee and Warshel 2012, 2015a, b; Ma and Schulten 2015).

Mechanoselective ion channels may be considered as molecular machines, too, and an ion gating is better understood now due to MD (Vasquez et al. 2008). Simulations were applied to the helicase motor (Dittrich and Schulten 2006; Yu et al. 2007; Flechsig and Mikhailov 2010) and perhaps to the most advanced (not counting more static virus capsids) object: the ribosome (Becker et al. 2009; Romanowska et al. 2008; Trylska 2010). Recent progress in ribosome simulations has been summarized by Sanbonmatsu (2012). The MD application in this field obviously profits from using 3-D animation; such tools bring new dimensions to protein science (Iwasa 2015). More information on large macromolecular complexes may be found in Perilla et al. (2015).

The Mechanism of Enzymatic Activity

Progress in computational enzymology is described in Lonsdale et al. (2010) and Carvalho et al. (2014). Reactions have to be treated using quantum mechanics, but some features of enzyme dynamics may be revealed by MD. For example, Peplowski et al. used the steered MD method to enforce a ligand transport within the biotechnological enzyme nitrile hydratase (Peplowski et al. 2008). In that way, residues that may change catalytic properties of this metalloprotein have been indicated. Carloni et al. used the MD approach to explain the mode of action of G proteins (Khafizov et al. 2009). Lid opening in proteasome has been studied in detail (Ishida 2014). Worth checking are papers devoted to computational studies of the model enzyme dihydrofolate reductase, since this enzyme is a target in cancer therapy and anti-malaria drug development (Kohen 2015). MD may serve also as a useful tool in designing new enzymatic functions (Damborsky and Brezovsky

2014). Free energy calculations in the context of enzymatic activity are easy to perform due to empirical valence bond (EVB) + MD software developments (Isaksen et al. 2015). Useful guidelines on how to perform effectively free energy simulations may be found in Klimovich et al. (2015).

Transport Phenomena in Proteins

Heme proteins are popular objects of MD simulations (Bikiel et al. 2006) since transport processes of small gaseous ligands (O_2 , NO, CO) are important for physiology, and therefore they serve as a “test ground” for new methods. Myoglobin is even sometimes called a “hydrogen atom” of MD simulations (Elber 2010). Other members of the heme globin family have been recently discovered, and studies of diffusion paths and free energy landscapes for neuroglobin (Orlowski and Nowak 2008), cytoglobin (Orlowski and Nowak 2007), or protoglobin (Forti et al. 2011a, b) were published. It seems that such systematic study allows for construction of uniform picture of ligand migration pathways and the evolution of transport protein structure (Cohen et al. 2008; Forti et al. 2011b). MD simulations are helpful in the explanation of complex enzymatic mechanisms, for example, of dehydrogenase/acetyl-CoA synthase (Wang et al. 2013) or cytochrome C oxidases (Oliveira et al. 2014), and interpretation of X-ray experiments (Tsuduki et al. 2012). However, determination of larger ligand diffusion paths is not trivial (Kingsley and Lill 2015). The standard MD protocol is not practical since simulation times required to collect reasonable statistics are long. Special variants of steered molecular dynamics are useful in such problems; see Fig. 4. The review of recent methodological progress and new ideas based on memetic algorithms are presented in a recent paper by Rydzewski and Nowak (2015).

Structure and Dynamics of Ion Channels and Porins

Dynamics and transport through ion channels and other pores in biological membranes is a subject of vigorous research (Khalili-Araghi et al. 2009; Sigg 2014). Among many papers, the work on channel gating by M. Sansom group (Fowler and Sansom 2013; Aryal et al. 2015) is worth mentioning as well as B. Roux et al. computational studies of ion channels (Horn et al. 2014; Li et al. 2014) and numerous papers on the ion conductance in a potassium channel (Boiteux et al. 2007; Kim and Warshel 2014; Delemotte et al. 2015; Linder et al. 2015) and sodium channels as well (Li and Gong 2015). The main research groups have tried to formulate a uniform picture on the voltage-dependent ion channel conductance mechanism (Vargas et al. 2012). One should note that external electric field has critical impact on ion channel activity and this effect may be effectively modeled (English and Waldron 2015).

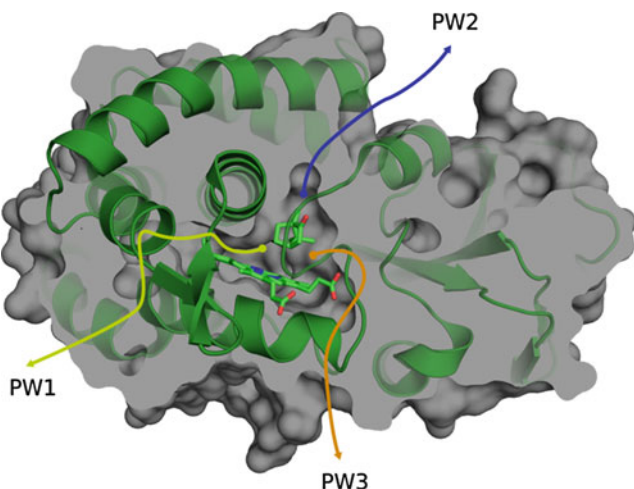


Fig. 4 A schematic view of complex access/exit paths PW1–PW3 of camphor ligand (licorice) in cytochrome P450cam enzyme calculated by our fast SMD memetic algorithm (Rydzewski and Nowak 2015). The figure was prepared using the PyMOL code (DeLano 2002) by J. Rydzewski

References to interesting computational studies of aquaporins may be found in a review by Hub et al. (2009). The water transport in eye lenses was investigated using large-scale simulations by D.E. Shaw group (Ikeguchi 2009). At the same time, K. Schulten's team has modeled nanomechanics of RNA in nanopores (Miao and Schulten 2009; Khalili-Araghi et al. 2009). More recently, a huge simulation of nanopore formation provided new data on nuclear pore gating (Gamini et al. 2014).

Sodium symporter modeling has been recently reviewed by Bisha and Magistrato (2016); other symporters are described in Kardos and Héja (2015) or Espinoza-Fonseca and Ramírez-Salinas (2015). Increasing computational resources open possibility of all-atom simulations of a protein translocation through a nanopore (Di Marino et al. 2015).

Protein-DNA Interactions

Transfer of information from DNA to proteins and its impact on the whole cell activity is determined by protein-DNA interactions. Such complexes are difficult to model due to their size and heterogeneity, but good understanding of these systems is a key to genetics (Mac Kerell and Nilsson 2008; Rohs et al. 2009). Good examples of successful MD applications are studies of p53 protein binding modes to DNA quadruplexes (Ma and Levine 2007), investigations of zinc-finger proteins binding to DNA (Lee et al. 2010), or modeling of DNA bending (van der Vaart 2015). Details of the mechanism of Lac repressor interactions (Villa et al. 2005) and sliding along DNA (Furini et al. 2010) have been discovered by the MD methods. There is a growing interest in studies of proteins related to clustered regularly interspaced

short palindromic repeats (CRISPRs) present in prokaryotic cells. A recent example of insightful study of CRISPR-related endoribonuclease may be found in Estarellas et al. (2015).

Origins of Molecular Diseases

Applications of computer simulations in the areas related to medical problems are numerous: funds come both from the public budgets and private charities. Many diseases have well-defined etiology, related to abnormalities in a protein structure, point mutations, etc. Such “molecular diseases” are popular objects of the theoretical modeling (Papaleo and Invernizzi 2011).

The epidemic of BSE spawned the great interest in the prion protein research. Simulations of folding (Sakudo et al. 2010; Kupfer et al. 2009) evolved into studies of membrane-bound complexes (DeMarco and Daggett 2009). The Alzheimer disease is related to amyloid fiber formation in the brain, and this process is successfully modeled (Urbanc et al. 2010; Kassler et al. 2010; Straub and Thirumalai 2010; Nasica-Labouze et al. 2015), too. Less known are studies of transthyretin (TTR) fibril formation (Ortore and Martinelli 2012; Rodrigues et al. 2010). However, the problem is serious since the aggregation of TTR leads to a lethal illness. It is necessary to notice some impressive studies of HIV virus (Carnevale et al. 2009; Zhao et al. 2013). There are examples of computational research on proteins involved in primary congenital glaucoma (Achary and Nagarajaram 2009), osteoporosis (Lee et al. 2011), epidermolysis bullosa simplex (Jankowski et al. 2014), autism spectrum disorders (Mikulska et al. 2011), or cancer (Kumar and Purohit 2014). An excellent example of MD exploratory study of point mutations in CFTR chloride channel involved in cystic fibrosis is the recent paper by Morion et al. (2015).

Simulations of Single-Molecule AFM Experiments

In the opinion of this author (and many others as well), non-equilibrium protein dynamics has an excellent overlap with single-molecule experiments performed by atomic force microscope (AFM) (Nowak and Marszalek 2005; Kumar and Li 2010; Galera-Prat et al. 2010). One of the first computational studies of ligand-antibody enforced dissociation has been published in 2001 by Karplus et al. (Paci et al. 2001). Stretching of individual molecules by the AFM cantilever gives force spectra, i.e., dependence of a force (in the range 0–3,000 pN) on protein extension (0–100 nm or even more) (Rief and Grubmuller 2002) that may be confronted with special type of MD simulations, for example, steered molecular dynamics (Nowak and Marszalek 2005) or even quantum-type SMD (Lu et al. 2004). An example of SMD stretching of reelin (see Fig. 2) – a protein important in neural system development and proper functioning of synapses – is presented in Fig. 5.

One can see that despite rather low stretching velocity, the computed SMD forces are much higher (1–2 nN) than that (100–300 pN) typically observed in AFM

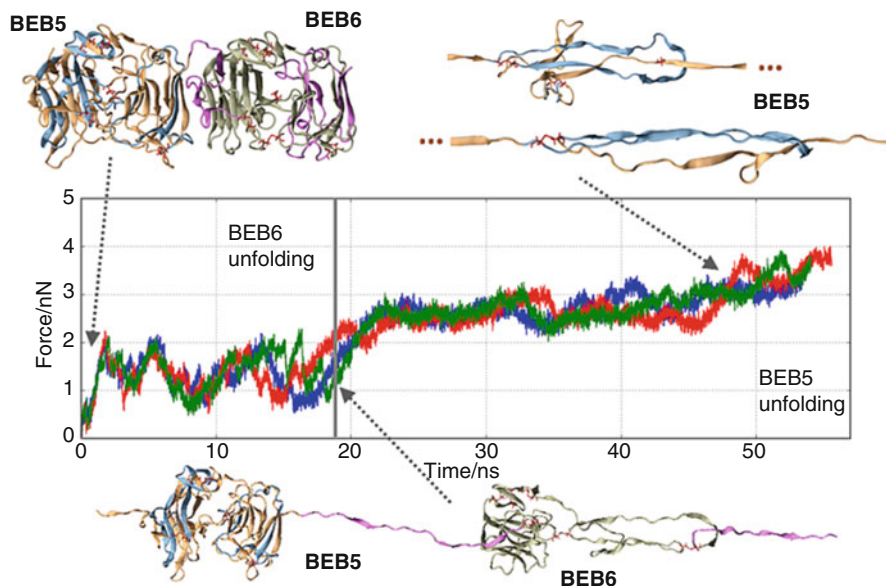


Fig. 5 An example of mechanical unfolding scenario of modular protein reelin recovered by SMD (CHARMM27 force field, NAMD code (Phillips et al. 2005)). K. Mikulska, W. Nowak (unpublished results). The figure was prepared using the VMD code (Humphrey et al. 1996)

experiments (Mikulska et al. 2012); this is related to high loading rates imposed by limited SMD time (G. Lee et al. 2004). Good progress toward SMD and high-speed AFM result consistency is observed (Sotomayor and Schulten 2007; Rico et al. 2013) especially if coarse-grained models of proteins are employed (Kumar and Li 2010; Galera-Prat et al. 2010; Mikulska et al. 2014; Chwastyk et al. 2014; Chen et al. 2016). Free energy profiles along a stretching coordinate may be calculated using Jarzynski theorem (Hummer and Szabo 2010; Rydzewski et al. 2015). Therefore, more and more laboratories use SMD modeling as a tool to interpret results of SMFS (He et al. 2012) and optical tweezers (Sieben et al. 2012) experiments. A computational microscope postulated long time ago (Lee et al. 2009) is at hand.

Conclusions and Future Directions

Computational chemistry is currently a well-established, fully functional branch of science (Akimov and Prezhdo 2015). Easy access to computers and high-quality, specialized computer codes results in myriad of applications. There is a large, well-trained, and active community of computational chemists. Results of calculations are useful and hard to obtain without the theoretical and computational approaches. Virtually all chemical and a large body of biological systems may be modeled using current facilities.

It looks that computer simulations of proteins have still very good prospects ahead. There are new, promising directions of the investigations. Better computational power offers possibility of real-time calculations of protein dynamics, including aggregates and membrane-embedded systems. MD simulations of proteins may help in rational drug design (Harvey and De Fabritiis 2012). Electronic excited states of proteins are unexplored area. The methodology of MD simulations of photoexcited states is still not well established (Kubiak and Nowak 2008) (Rydzewski and Nowak, chapter “► [Molecular Dynamics in Excited State](#)”), but future MD applications should include interactions of biological systems with light (Dittrich et al. 2005; Hayashi et al. 2009; Rossle and Frank 2009). For example, nice progress of understanding protonic gating in the popular green fluorescent protein chromophore has been achieved through dynamical simulations (Olsen et al. 2010). Similar studies related to optogenetics will appear soon. Every month brings new better software for extensive analysis of MD trajectories (Zwier et al. 2015) or new ways of doing MD (Kukic et al. 2015).

An interpretation of special experiments, such as high-pressure studies (Paci 2002) and especially cryo-electron microscopy (McGreevy et al. 2016) or X-ray data fitting (McGreevy et al. 2014), may be fruitfully augmented by an application of computational modeling. Computer scientists and physicists work hard to enlarge the maximum size of the simulated system. The whole virus all-atom simulation is not a record study anymore (Zink and Grubmuller 2009; Sanbonmatsu and Tung 2007; Zhao et al. 2013; Goh et al. 2015), computational virology field has been established (Reddy and Sansom 2016), and there are plans to use computational methods in the design of useful viruses {Zhang et al. 2015, p. 8316}. However, one should note that recent record of 64-million-atom NAMD simulations required 8,000 Cray XK7 nodes for a period of 2 months (Perilla et al. 2015).

Long time ago, it has been shown in Los Alamos National Laboratory that even 320-million-atom protein simulations are technically possible (2009). Large systems, such as big protein complexes in the living cell, will require coarse-grained approaches (Clementi 2008). A lot of insight comes from such modeling of membrane proteins (Sansom et al. 2008; Ayton et al. 2010; Stansfeld et al. 2015). Multiscale modeling is yet another line of methodological progress (Tozzini 2010; Nielsen et al. 2010; Goga et al. 2015). Such techniques expand timescale accessible for MD studies beyond 1 μ s limit (Chu et al. 2007; Ayton et al. 2007; Sherwood et al. 2008). Impressive simulations on microsecond timescale coarse-grained Martini FF molecular dynamics simulations of enveloped virions in explicit solvent (5,000,000 particles) were performed by Sansom group at Oxford (Reddy et al. 2015). The calculated properties of the influenza A virion were consistent with experimental measurements.

What is the best technological strategy for the optimum performance of MD simulations is a matter of debate (Borell 2008; Goga et al. 2015). For example, D.E. Shaw has invested a lot of resources in order to develop dedicated chips with record performance for specialized tasks (Klepeis et al. 2009). Other groups, such as K. Schulten's team, prefer improving algorithms and developing software for

running calculations on very powerful, relatively inexpensive, graphical processing units, mass produced for computer game fans (Hardy et al. 2009; Stone et al. 2007; Zhmurov et al. 2010).

Both approaches have difficulty with surpassing grid computing idea: “Folding@Home” project created perhaps the most powerful computational device ever and attracted a lot of young people to science (Pande et al. 2003; Belden et al. 2015). Hopefully, all this computational chemistry modeling effort will bring us a better understanding of nature and a better life for everybody.

Acknowledgment This work was supported in part by Polish funds for science NCN (N N202 262038 and 2012/05/N/ST3/03178). Infrastructure of ICNT UMK is also acknowledged.

Bibliography

- Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., et al. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1–2, 19–25.
- Abrams, C., & Bussi, G. (2013). Enhanced sampling in molecular dynamics using metadynamics, replica-exchange, and temperature-acceleration. *Entropy*, 16(1), 163.
- Achary, M. S., & Nagarajaram, H. A. (2009). Effects of disease causing mutations on the essential motions in proteins. *Journal of Biomolecular Structure and Dynamics*, 26(5), 609–624.
- Adcock, S. A., & McCammon, J. A. (2006). Molecular dynamics: Survey of methods for simulating the activity of proteins. *Chemical Reviews*, 106(5), 1589–1615.
- Akimov, A. V., & Prezhdo, O. V. (2015). Large-scale computations in chemistry: A bird’s eye view of a vibrant field. *Chemical Reviews*, 115(12), 5797–5890.
- Aksimentiev, A., Balabin, I. A., Fillingame, R. H., & Schulten, K. (2004). Insights into the molecular mechanism of rotation in the Fo sector of ATP synthase. *Biophysical Journal*, 86(3), 1332–1344.
- Aksimentiev, A., Brunner, R., Cohen, J., Comer, J., Cruz-Chu, E., Hardy, D., et al. (2008). Computer modeling in biotechnology: A partner in development. *Methods in Molecular Biology*, 474, 181–234.
- Alder, B. J., & Wainwright, T. E. (1957). Phase transition for a hard sphere system. *The Journal of Chemical Physics*, 27, 1208–1210.
- Aleksandrov, A., Thompson, D., & Simonson, T. (2010). Alchemical free energy simulations for biological complexes: Powerful but temperamental. *Journal of Molecular Recognition*, 23(2), 117–127.
- Alexander, S. (2010). Visualization and analysis of atomistic simulation data with OVITO—the open visualization tool. *Modelling and Simulation in Materials Science and Engineering*, 18(1), 015012.
- Aliev, A. E., & Courtier-Murias, D. (2010). Experimental verification of force fields for molecular dynamics simulations using Gly-Pro-Gly-Gly. *Journal of Physical Chemistry B*, 114(38), 12358–12375.
- Allen, M. P., & Tildesley, D. J. (1987). *Computer simulation of liquids*. Oxford: Clarendon.
- Amadei, A., Linssen, A. B., & Berendsen, H. J. (1993). Essential dynamics of proteins. *Proteins*, 17(4), 412–425.
- Andersen, O. J., Grouleff, J., Needham, P., Walker, R. C., & Jensen, F. (2015). Toward an enhanced sampling molecular dynamics method for studying ligand-induced conformational changes in proteins. *Journal of Physical Chemistry B*, 119(46), 14594–14603.
- Andoh, Y., Yoshii, N., Fujimoto, K., Mizutani, K., Kojima, H., Yamada, A., et al. (2013). MODYLAS: A highly parallelized general-purpose molecular dynamics simulation program

- for large-scale systems with long-range forces calculated by fast multipole method (FMM) and highly scalable fine-grained new parallel processing algorithms. *Journal of Chemical Theory and Computation*, 9(7), 3201–3209.
- Aqvist, J., Luzhkov, V. B., & Brandsdal, B. O. (2002). Ligand binding affinities from MD simulations. *Accounts of Chemical Research*, 35(6), 358–365.
- Aryal, P., Sansom, M. S., & Tucker, S. J. (2015). Hydrophobic gating in ion channels. *Journal of Molecular Biology*, 427(1), 121–130.
- Avila, C. L., Drechsel, N. J., Alcantara, R., & Ville-Freixa, J. (2011). Multiscale molecular dynamics of protein aggregation. *Current Protein and Peptide Science*, 21, 12(3), 221–234.
- Ayton, G. S., Noid, W. G., & Voth, G. A. (2007). Multiscale modeling of biomolecular systems: In serial and in parallel. *Current Opinion in Structural Biology*, 17(2), 192–198.
- Ayton, G. S., Lyman, E., & Voth, G. A. (2010). Hierarchical coarse-graining strategy for protein-membrane systems to access mesoscopic scales. *Faraday Discussions*, 144, 347–357.
- Bahar, I., Lezon, T. R., Yang, L.-W., & Eyal, E. (2010). Global dynamics of proteins: Bridging between structure and function. *Annual Review of Biophysics*, 39(1), 23–42.
- Bahar, I., Cheng, M. H., Lee, J. Y., Kaya, C., & Zhang, S. (2015). Structure-encoded global motions and their role in mediating protein-substrate interactions. *Biophysical Journal*, 109(6), 1101–1109.
- Baker, C. M., & Best, R. B. (2013). Matching of additive and polarizable force fields for multiscale condensed phase simulations. *Journal of Chemical Theory and Computation*, 9(6), 2826–2837.
- Baker, C. M., & Best, R. B. (2014). Insights into the binding of intrinsically disordered proteins from molecular dynamics simulation. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(3), 182–198.
- Barnoud, J., & Monticelli, L. (2015). Coarse-grained force fields for molecular simulations. *Methods in Molecular Biology*, 1215, 125–149.
- Beauchamp, K. A., Lin, Y. S., Das, R., & Pande, V. S. (2012). Are protein force fields getting better? A systematic benchmark on 524 diverse NMR measurements. *Journal of Chemical Theory and Computation*, 8(4), 1409–1414.
- Becker, O. M., & Karplus, M. (2006). *A guide to biomolecular simulations* (Vol. 4). Dordrecht: Springer.
- Becker, T., Bhushan, S., Jarasch, A., Armache, J. P., Funes, S., Jossinet, F., et al. (2009). Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. *Science*, 326(5958), 1369–1373.
- Belden, O. S., Baker, S. C., & Baker, B. M. (2015). Citizens unite for computational immunology! *Trends in Immunology*, 36(7), 385–387.
- Berendsen, H. J. C. E. (1976). In *Proceedings of the CECAM workshop on models for protein dynamics*, Orsay.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The protein data bank. *Nucleic Acids Research*, 28(1), 235–242.
- Bernardi, R. C., Melo, M. C., & Schulten, K. (2015). Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochimica et Biophysica Acta*, 1850(5), 872–877.
- Biarnes, X., Bongarzone, S., Vargiu, A. V., Carloni, P., & Rugggerone, P. (2011). Molecular motions in drug design: The coming age of the metadynamics method. *Journal of Computer-Aided Molecular Design*, 25(5), 395–402.
- Bikiel, D. E., Boechi, L., Capece, L., Crespo, A., De Biase, P. M., Di Lella, S., et al. (2006). Modeling heme proteins using atomistic simulations. *Physical Chemistry Chemical Physics*, 8(48), 5611–5628.
- Bisha, I., & Magistrato, A. (2016). The molecular mechanism of secondary sodium symporters under the lens of the computational microscope. *RSC Advances*, 6, 9522–9540.
- Blaszczak, M., Jamroz, M., Kmiecik, S., & Kolinski, A. (2013). CABS-fold: Server for the de novo and consensus-based prediction of protein structure. *Nucleic Acids Res*, 41(Web Server issue), W406–411.
- Boas, F. E., & Harbury, P. B. (2007). Potential energy functions for protein design. *Current Opinion in Structural Biology*, 17(2), 199–204.

- Bock, L. V., Blau, C., Schröder, G. F., Davydov, I. I., Fischer, N., Stark, H., et al. (2013). Energy barriers and driving forces in tRNA translocation through the ribosome. *Nature Structural & Molecular Biology*, 20(12), 1390–1396.
- Boiteux, C., Kraszewski, S., Ramseyer, C., & Girardet, C. (2007). Ion conductance vs. pore gating and selectivity in KcsA channel: Modeling achievements and perspectives. *Journal of Molecular Modeling*, 13(6–7), 699–713.
- Borell, B. (2008). Chemistry: Power play. *Nature*, 451, 240–243.
- Borhani, D. W., & Shaw, D. E. (2012). The future of molecular dynamics simulations in drug discovery. *Journal of Computer-Aided Molecular Design*, 26(1), 15–26.
- Bottaro, S., Lindorff-Larsen, K., & Best, R. B. (2013). Variational optimization of an all-atom implicit solvent force field to match explicit solvent simulation data. *Journal of Chemical Theory and Computation*, 9(12), 5641–5652.
- Bowman, G. R., Voelz, V. A., & Pande, V. S. (2011). Taming the complexity of protein folding. *Current Opinion in Structural Biology*, 21(1), 4–11.
- Brooks, B. R., Brooks, C. L., 3rd, Mackerell, A. D., Jr., Nilsson, L., Petrella, R. J., Roux, B., et al. (2009). CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*, 30(10), 1545–1614.
- Brunk, E., & Rothlisberger, U. (2015). Mixed quantum mechanical/molecular mechanical molecular dynamics simulations of biological systems in ground and electronically excited states. *Chemical Reviews*, 115(12), 6217–6263.
- Carnevale, V., Raugei, S., Neri, M., Pantano, S., Micheletti, C., & Carloni, P. (2009). Multi-scale modeling of HIV-1 proteins. *Journal of Molecular Structure (THEOCHEM)*, 898(1–3), 97–105.
- Carvalho, A. T. P., Barrozo, A., Doron, D., Kilshtain, A. V., Major, D. T., & Kamerlin, S. C. L. (2014). Challenges in computational studies of enzyme structure, function and dynamics. *Journal of Molecular Graphics and Modelling*, 54, 62–79.
- Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., et al. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16), 1668–1688.
- Chen, J., & Brooks, C. L., 3rd. (2008). Implicit modeling of nonpolar solvation for simulating protein folding and conformational transitions. *Physical Chemistry Chemical Physics*, 10(4), 471–481.
- Chen, J., Brooks, C. L., 3rd, & Khandogin, J. (2008). Recent advances in implicit solvent-based methods for biomolecular simulations. *Current Opinion in Structural Biology*, 18(2), 140–148.
- Chen, J., Xie, Z. -R., & Wu, Y. (2016). Study of protein structural deformations under external mechanical perturbations by a coarse-grained simulation method. *Biomechanics and Modeling in Mechanobiology*, 15, 317–329.
- Chipot, C. (2015). Applications to real size biological systems. In B. Engquist (Ed.), *Encyclopedia of applied and computational mathematics* (pp. 72–81). Berlin/Heidelberg: Springer.
- Chou, K. C. (2004). Structural bioinformatics and its impact to biomedical science. *Current Medicinal Chemistry*, 11(16), 2105–2134.
- Christ, C. D., Mark, A. E., & van Gunsteren, W. F. (2010). Basic ingredients of free energy calculations: A review. *Journal of Computational Chemistry*, 31(8), 1569–1582.
- Christen, M., Hunenberger, P. H., Bakowies, D., Baron, R., Burgi, R., Geerke, D. P., et al. (2005). The GROMOS software for biomolecular simulation: GROMOS05. *Journal of Computational Chemistry*, 26(16), 1719–1751.
- Chu, J.-W., Ayton, G. S., Izvekov, S., & Voth, G. A. (2007). Emerging methods for multiscale simulation of biomolecular systems. *Molecular Physics*, 105, 167–175.
- Chwastyk, M., Galera-Prat, A., Sikora, M., Gomez-Sicilia, A., Carrion-Vazquez, M., & Cieplak, M. (2014). Theoretical tests of the mechanical protection strategy in protein nanomechanics. *Proteins*, 82(5), 717–726.
- Clementi, C. (2008). Coarse-grained models of protein folding: Toy models or predictive tools? *Current Opinion in Structural Biology*, 18(1), 10–15.
- Cohen, J., Olsen, K. W., & Schulten, K. (2008). Finding gas migration pathways in proteins using implicit ligand sampling. *Methods in Enzymology*, 437, 439–457.

- Comer, J., Phillips, J. C., Schulten, K., & Chipot, C. (2014). Multiple-replica strategies for free-energy calculations in NAMD: Multiple-walker adaptive biasing force and walker selection rules. *Journal of Chemical Theory and Computation*, *10*(12), 5276–5285.
- Cornell, W., & Nam, K. (2009). Steroid hormone binding receptors: Application of homology modeling, induced fit docking, and molecular dynamics to study structure-function relationships. *Current Topics in Medicinal Chemistry*, *9*(9), 844–853.
- Cumberworth, A., Bui, J. M., & Gsponer, J. (2015). Free energies of solvation in the context of protein folding: Implications for implicit and explicit solvent models. *Journal of Computational Chemistry*, *37*(7), 629–640.
- Czub, J., & Grubmüller, H. (2014). Rotation triggers nucleotide-independent conformational transition of the empty beta subunit of F(1)-ATPase. *Journal of the American Chemical Society*, *136*(19), 6960–6968.
- Dahl, J. P. (2001). *Introduction to the quantum world of atoms and molecules*. Singapore: World Scientific.
- Dal Peraro, M., Ruggerone, P., Raugei, S., Gervasi, F., & Elber, R. (2005). Long-timescale simulation methods. *Current Opinion in Structural Biology*, *15*, 151–156.
- Dal Peraro, M., Ruggerone, P., Raugei, S., Gervasio, F. L., & Carloni, P. (2007). Investigating biological systems using first principles Car-Parrinello molecular dynamics simulations. *Current Opinion in Structural Biology*, *17*(2), 149–156.
- Damborsky, J., & Brezovsky, J. (2014). Computational tools for designing and engineering enzymes. *Current Opinion in Chemical Biology*, *19*, 8–16.
- DeLano, W. L. (2002). The PyMOL molecular graphics system.
- Delemotte, L., Kasimova, M. A., Klein, M. L., Tarek, M., & Carnevale, V. (2015). Free-energy landscape of ion-channel voltage-sensor-domain activation. *Proceedings of the National Academy of Sciences*, *112*(1), 124–129.
- DeMarco, M. L., & Daggett, V. (2009). Characterization of cell-surface prion protein relative to its recombinant analogue: Insights from molecular dynamics simulations of diglycosylated, membrane-bound human prion protein. *Journal of Neurochemistry*, *109*(1), 60–73.
- Deng, Y., & Roux, B. (2009). Computations of standard binding free energies with molecular dynamics simulations. *Journal of Physical Chemistry B*, *113*(8), 2234–2246.
- Di Marino, D., Bonome, E. L., Tramontano, A., & Chinappi, M. (2015). All-atom molecular dynamics simulation of protein translocation through an α -hemolysin nanopore. *Journal of Physical Chemistry Letters*, *6*(15), 2963–2968.
- Dill, K. A., & MacCallum, J. L. (2012). The protein-folding problem, 50 years on. *Science*, *338*(6110), 1042–1046.
- Dittrich, M., & Schulten, K. (2006). PcrA helicase, a prototype ATP-driven molecular motor. *Structure*, *14*(9), 1345–1353.
- Dittrich, M., Freddolino, P. L., & Schulten, K. (2005). When light falls in LOV: A quantum mechanical/molecular mechanical study of photoexcitation in Phot-LOV1 of *Chlamydomonas reinhardtii*. *Journal of Physical Chemistry B*, *109*(26), 13006–13013.
- Do, T. N., Choy, W. Y., & Karttunen, M. (2014). Accelerating the conformational sampling of intrinsically disordered proteins. *Journal of Chemical Theory and Computation*, *10*(11), 5081–5094.
- Dodson, G. G., Lane, D. P., & Verma, C. S. (2008). Molecular simulations of protein dynamics: New windows on mechanisms in biology. *EMBO Reports*, *9*(2), 144–150.
- Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H. F., & Shaw, D. E. (2012). Biomolecular simulation: A computational microscope for molecular biology. *Annual Review of Biophysics*, *41*(41), 429–452.
- Dror, R. O., Green, H. F., Valant, C., Borhani, D. W., Valcourt, J. R., Pan, A. C., et al. (2013). Structural basis for modulation of a G-protein-coupled receptor by allosteric drugs. *Nature*, *503*(7475), 295–299.
- Duan, Y., & Kollman, P. A. (1998). Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science*, *282*(5389), 740–744.

- Eastman, P., & Pande, V. S. (2015). OpenMM: A hardware independent framework for molecular simulations. *Computing in Science & Engineering*, 12(4), 34–39.
- Elber, R. (2010). Ligand diffusion in globins: Simulations versus experiment. *Current Opinion in Structural Biology*, 20(2), 162–167.
- Elber, R., & Kirmizialtin, S. (2013). Molecular machines. *Current Opinion in Structural Biology*, 23(2), 206–211.
- Elber, R., Ghosh, A., & Cardenas, A. (2002). Long time dynamics of complex systems. *Accounts of Chemical Research*, 35(6), 396–403.
- English, N. J., & Waldron, C. J. (2015). Perspectives on external electric fields in molecular simulation: Progress, prospects and challenges. *Physical Chemistry Chemical Physics*, 17(19), 12407–12440. doi:10.1039/C5CP00629E.
- Ensign, D. L., Kasson, P. M., & Pande, V. S. (2007). Heterogeneity even at the speed limit of folding: Large-scale molecular dynamics study of a fast-folding variant of the villin headpiece. *Journal of Molecular Biology*, 374(3), 806–816.
- Espinoza-Fonseca, L. M., & Ramírez-Salinas, G. L. (2015). Microsecond molecular simulations reveal a transient proton pathway in the calcium pump. *Journal of the American Chemical Society*, 137(22), 7055–7058.
- Estarellas, C., Otyepka, M., Koča, J., Banáš, P., Krepl, M., & Šponer, J. (2015). Molecular dynamic simulations of protein/RNA complexes: CRISPR/Csy4 endoribonuclease. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1850(5), 1072–1090.
- Fersht, A. R., & Daggett, V. (2002). Protein folding and unfolding at atomic resolution. *Cell*, 108(4), 573–582.
- Field, M. J. (2015). Technical advances in molecular simulation since the 1980s. *Archives of Biochemistry and Biophysics*, 582, 3–9.
- Flechsig, H., & Mikhailov, A. S. (2010). Tracing entire operation cycles of molecular motor hepatitis C virus helicase in structurally resolved dynamical simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 107(49), 20875–20880.
- Forti, F., Boechi, L., Bikiel, D., Martí, M. A., Nardini, M., Bolognesi, M., et al. (2011a). Ligand migration in *Methanosarcina acetivorans* protoglobin: Effects of ligand binding and dimeric assembly. *The Journal of Physical Chemistry B*, 115(46), 13771–13780.
- Forti, F., Boechi, L., Estrin, D. A., & Martí, M. A. (2011b). Comparing and combining implicit ligand sampling with multiple steered molecular dynamics to study ligand migration processes in heme proteins. *Journal of Computational Chemistry*, 32(10), 2219–2231.
- Fowler, P. W., & Sansom, M. S. (2013). The pore of voltage-gated potassium ion channels is strained when closed. *Nature Communications*, 4, 1872.
- Frankel, D., & Smit, B. (2001). *Understanding molecular simulation* (2nd ed.). San Diego: Academic.
- Freddolino, P. L., & Schulten, K. (2009). Common structural transitions in explicit-solvent simulations of villin headpiece folding. *Biophysical Journal*, 97(8), 2338–2347.
- Freddolino, P. L., Arhipov, A. S., Larson, S. B., McPherson, A., & Schulten, K. (2006a). Molecular dynamics simulations of the complete satellite tobacco mosaic virus. *Structure*, 14(3), 437–449.
- Freddolino, P. L., Dittich, M., & Schulten, K. (2006b). Dynamic switching mechanisms in LOV1 and LOV2 domains of plant phototropins. *Biophysical Journal*, 91(10), 3630–3639.
- Freddolino, P. L., Liu, F., Gruebele, M., & Schulten, K. (2008). Ten-microsecond molecular dynamics simulation of a fast-folding WW domain. *Biophysical Journal*, 94(10), L75–77.
- Freddolino, P. L., Harrison, C. B., Liu, Y., & Schulten, K. (2010). Challenges in protein folding simulations: Timescale, representation, and analysis. *Nature Physics*, 6(10), 751–758.
- Frenkel, D., & Smit, B. (2002). Chapter 4 – Molecular dynamics simulations. In *Understanding molecular simulation* (2nd ed., pp. 63–107). San Diego: Academic.
- Fu, B., & Vendruscolo, M. (2015). Structure and dynamics of intrinsically disordered proteins. *Advances in Experimental Medicine and Biology*, 870, 35–48.
- Fujisaki, H., Moritsugu, K., Matsunaga, Y., Morishita, T., & Maragliano, L. (2015). Extended phase-space methods for enhanced sampling in molecular simulations: A review. *Frontiers in Bioengineering and Biotechnology*, 3, 125.

- Furini, S., Domene, C., & Cavalcanti, S. (2010). Insights into the sliding movement of the lac repressor nonspecifically bound to DNA. *The Journal of Physical Chemistry B*, 114(6), 2238–2245.
- Galeazzi, R. (2009). Molecular dynamics as a tool in rational drug design: Current status and some major applications. *Current Computer-Aided Drug Design*, 5(4), 225–240.
- Galera-Prat, A., Gomez-Sicilia, A., Oberhauser, A. F., Cieplak, M., & Carrion-Vazquez, M. (2010). Understanding biology by stretching proteins: Recent progress. *Current Opinion in Structural Biology*, 20(1), 63–69.
- Gallicchio, E., & Levy, R. M. (2011). Advances in all atom sampling methods for modeling protein-ligand binding affinities. *Current Opinion in Structural Biology*, 21, 161–166
- Gamini, R., Han, W., Stone, J. E., & Schulten, K. (2014). Assembly of Nsp1 nucleoporins provides insight into nuclear pore complex gating. *PLoS Computational Biology*, 10(3), e1003488.
- Gamiz-Hernandez, A. P., & Kaila, V. R. (2016). Conversion of light-energy into molecular strain in the photocycle of the photoactive yellow protein. *Physical Chemistry Chemical Physics*, 18(4), 2802–2809.
- Gao, M., Sotomayor, M., Villa, E., Lee, E. H., & Schulten, K. (2006). Molecular mechanisms of cellular mechanics. *Physical Chemistry Chemical Physics*, 8(32), 3692–3706.
- Goga, N., Melo, M., Rzepiela, A., De Vries, A., Hadar, A., Marrink, S., et al. (2015). Benchmark of schemes for multiscale molecular dynamics simulations. *Journal of Chemical Theory and Computation*, 11(4), 1389–1398.
- Goh, B. C., Perilla, J. R., England, M. R., Heyrana, K. J., Craven, R. C., & Schulten, K. (2015). Atomic modeling of an immature retroviral lattice using molecular dynamics and mutagenesis. *Structure*, 23(8), 1414–1425.
- Granata, D., Camilloni, C., Vendruscolo, M., & Laio, A. (2013). Characterization of the free-energy landscapes of proteins by NMR-guided metadynamics. *Proceedings of the National Academy of Sciences of the United States of America*, 110(17), 6817–6822.
- Granata, D., Baftizadeh, F., Habchi, J., Galvagnion, C., De Simone, A., Camilloni, C., et al. (2015). The inverted free energy landscape of an intrinsically disordered peptide by simulations and experiments. *Scientific Reports*, 5, 15449.
- Greenberger, D., Hentschel, K., & Weinert, F. (2009). *Compendium of quantum physics*. Berlin/Heidelberg: Springer.
- Gu, J., & Bourne, P. E. (Eds.). (2009). *Structural bioinformatics* (2nd ed.). Hoboken: Wiley-Blackwell.
- Gumbart, J., Wang, Y., Aksimentiev, A., Tajkhorshid, E., & Schulten, K. (2005). Molecular dynamics simulations of proteins in lipid bilayers. *Current Opinion in Structural Biology*, 15(4), 423–431.
- Guvench, O., & MacKerell, A. D., Jr. (2008). Comparison of protein force fields for molecular dynamics simulations. *Methods in Molecular Biology*, 443, 63–88.
- Haile, M. (1992). *Molecular dynamics simulation: Elementary methods*. New York: Wiley.
- Hansson, T., Oostenbrink, C., & van Gunsteren, W. (2002). Molecular dynamics simulations. *Current Opinion in Structural Biology*, 12(2), 190–196.
- Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, 4(1), 17.
- Hardy, D. J., Stone, J. E., & Schulten, K. (2009). Multilevel summation of electrostatic potentials using graphics processing units. *Parallel Computing*, 35(3), 164–177.
- Harvey, M. J., & De Fabritiis, G. (2012). High-throughput molecular dynamics: The powerful new tool for drug discovery. *Drug Discovery Today*, 17(19), 1059–1062.
- Hayashi, S., Tajkhorshid, E., & Schulten, K. (2009). Photochemical reaction dynamics of the primary event of vision studied by means of a hybrid molecular simulation. *Biophysical Journal*, 96(2), 403–416.
- He, C., Genchev, G. Z., Lu, H., & Li, H. (2012). Mechanically untying a protein slipknot: Multiple pathways revealed by force spectroscopy and steered molecular dynamics simulations. *Journal of the American Chemical Society*, 134(25), 10428–10435.

- Hensen, U., Meyer, T., Haas, J., Rex, R., Vriend, G., & Grubmuller, H. (2012). Exploring protein dynamics space: The dynasome as the missing link between protein structure and function. *PLoS One*, *7*(5), e33931.
- Henzler-Wildman, K., & Kern, D. (2007). Dynamic personalities of proteins. *Nature*, *450*(7172), 964–972.
- Horn, R., Roux, B., & Aqvist, J. (2014). Permeation redux: Thermodynamics and kinetics of ion movement through potassium channels. *Biophysical Journal*, *106*(9), 1859–1863.
- Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., & Simmerling, C. (2006). Comparison of multiple AMBER force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, *65*, 712–725.
- Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling*, *51*(1), 69–82.
- Hsin, J., Arkhipov, A., Yin, Y., Stone, J. E., & Schulten, K. (2008). Using VMD: An introductory tutorial. *Current Protocols in Bioinformatics*, Chapter 5, Unit 5 7.
- Huang, J., & MacKerell, A. D., Jr. (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry*, *34*(25), 2135–2145.
- Huang, J., Lopes, P. E., Roux, B., & MacKerell, A. D., Jr. (2014). Recent advances in polarizable force fields for macromolecules: Microsecond simulations of proteins using the classical drude oscillator model. *Journal of Physical Chemistry Letters*, *5*(18), 3144–3150.
- Huang, W., Manglik, A., Venkatakrishnan, A. J., Laeremans, T., Feinberg, E. N., Sanborn, A. L., et al. (2015). Structural insights into micro-opioid receptor activation. *Nature*, *524*(7565), 315–321.
- Hub, J. S., & de Groot, B. L. (2009). Detection of functional modes in protein dynamics. *PLoS Computational Biology*, *5*(8), e1000480.
- Hub, J. S., Grubmuller, H., & de Groot, B. L. (2009). Dynamics and energetics of permeation through aquaporins. What do we learn from molecular dynamics simulations? *Handbook of Experimental Pharmacology*, *190*, 57–76.
- Hummer, G., & Szabo, A. (2010). Free energy profiles from single-molecule pulling experiments. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21441–21446.
- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, *14*(1), 33–38, 27–38.
- Ikeguchi, M. (2009). Water transport in aquaporins: Molecular dynamics simulations. *Frontiers in Bioscience*, *14*, 1283–1291.
- Ingolfsson, H. I., Lopez, C. A., Uusitalo, J. J., de Jong, D. H., Gopal, S. M., Periolo, X., et al. (2014). The power of coarse graining in biomolecular simulations. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *4*(3), 225–248.
- Isaksen, G. V., Andberg, T. A. H., Åqvist, J., & Brandsdal, B. O. (2015). Qgui: A high-throughput interface for automated setup and analysis of free energy calculations and empirical valence bond simulations in biological systems. *Journal of Molecular Graphics and Modelling*, *60*, 15–23.
- Ishida, H. (2014). Essential function of the N-termini tails of the proteasome for the gating mechanism revealed by molecular dynamics simulations. *Proteins*, *82*(9), 1985–1999.
- Ito, Y., & Ikeguchi, M. (2014). Molecular dynamics simulations of F1-ATPase. In *Protein conformational dynamics* (pp. 411–440). Cham: Springer.
- Iwasa, J. H. (2015). Bringing macromolecular machinery to life using 3D animation. *Current Opinion in Structural Biology*, *31*, 84–88.
- Jankowski, M., Wertheim-Tysarowska, K., Jakubowski, R., Sota, J., Nowak, W., & Czajkowski, R. (2014). Novel KRT14 mutation causing epidermolysis bullosa simplex with variable phenotype. *Experimental Dermatology*, *23*(9), 684–687.
- Jorgensen, W. L. (2013). Foundations of biomolecular modeling. *Cell*, *155*(6), 1199–1202.

- Jorgensen, W. L., & Tirado-Rives, J. (1988). The Opls potential functions for proteins – Energy minimizations for crystals of cyclic-peptides and crambin. *Journal of the American Chemical Society*, *110*(6), 1657–1666.
- Kannan, S., & Zacharias, M. (2009). Simulated annealing coupled replica exchange molecular dynamics – An efficient conformational sampling method. *Journal of Structural Biology*, *166*(3), 288–294.
- Kardos, J., & Héja, L. (2015). How membrane proteins work giving autonomous traverse pathways? *Structural Chemistry*, *26*(5–6), 1405–1410.
- Karplus, M. (2003). Molecular dynamics of biological macromolecules: A brief history and perspective. *Biopolymers*, *68*(3), 350–358.
- Karplus, M. (2014). Development of multiscale models for complex chemical systems: From H + H(2) to biomolecules (Nobel lecture). *Angewandte Chemie International Edition in English*, *53*(38), 9992–10005.
- Karplus, M., & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nature Structural Biology*, *9*(9), 646–652.
- Kassler, K., Horn, A. H. C., & Sticht, H. (2010). Effect of pathogenic mutations on the structure and dynamics of Alzheimer’s A beta(42)-amyloid oligomers. *Journal of Molecular Modeling*, *16*(5), 1011–1020.
- Khafizov, K., Lattanzi, G., & Carloni, P. (2009). G protein inactive and active forms investigated by simulation methods. *Proteins: Structure, Function, and Bioinformatics*, *75*(4), 919–930.
- Khalili-Araghi, F., Gumbart, J., Wen, P. C., Sotomayor, M., Tajkhorshid, E., & Schulten, K. (2009). Molecular dynamics simulations of membrane channels and transporters. *Current Opinion in Structural Biology*, *19*(2), 128–137.
- Kholmurodov, K. T., Altaisky, M. V., Puzynin, I. V., Darden, T., & Filatov, F. P. (2003). Methods of molecular dynamics for simulation of physical and biological processes. *Physics of Particles and Nuclei*, *34*(2), 244–263.
- Khurana, E., Devane, R. H., Dal Peraro, M., & Klein, M. L. (2011). Computational study of drug binding to the membrane-bound tetrameric M2 peptide bundle from influenza A virus. *Biochimica et Biophysica Acta*, *1808*(2), 530–537.
- Kim, I., & Warshel, A. (2014). Coarse-grained simulations of the gating current in the voltage-activated Kv1.2 channel. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(6), 2128–2133.
- Kingsley, L. J., & Lill, M. A. (2015). Substrate tunnels in enzymes: Structure-function relationships and computational methodology. *Proteins*, *83*(4), 599–611.
- Klein, M. L., & Shinoda, W. (2008). Large-scale molecular dynamics simulations of self-assembling systems. *Science*, *321*, 798–800.
- Klepeis, J. L., Pieja, M. J., & Floudas, C. A. (2003). Hybrid global optimization algorithms for protein structure prediction: Alternating hybrids. *Biophysical Journal*, *84*(2 Pt 1), 869–882.
- Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O., & Shaw, D. E. (2009). Long-timescale molecular dynamics simulations of protein structure and function. *Current Opinion in Structural Biology*, *19*(2), 120–127.
- Klimovich, P. V., Shirts, M. R., & Mobley, D. L. (2015). Guidelines for the analysis of free energy calculations. *Journal of Computer-Aided Molecular Design*, *29*(5), 397–411.
- Kmiecik, S., Gront, D., & Kolinski, A. (2007). Towards the high-resolution protein structure prediction. Fast refinement of reduced models with all-atom force field. *BMC Structural Biology*, *7*, 43.
- Kmiecik, S., Wabik, J., Kolinski, M., Kouza, M., & Kolinski, A. (2014). Coarse-grained modeling of protein dynamics. In *Computational methods to study the structure and dynamics of biomolecules and biomolecular processes* (pp. 55–79). Berlin/Heidelberg: Springer.
- Knapp, B., & Schreiner, W. (2009). Graphical user interfaces for molecular dynamics-quo vadis? *Bioinformatics and Biology Insights*, *3*, 103–107.
- Kohen, A. (2015). Dihydrofolate reductase as a model for studies of enzyme dynamics and catalysis [version 1; referees: 2 approved]. *F1000Research*, *4*(F1000 Faculty Rev), 1464. doi:10.12688/f1000research.6968.1.

- Kremer, K. (2003). Computer simulations for macromolecular science. *Macromolecular Chemistry and Physics*, 204(2), 257–264.
- Krieger, E., & Vriend, G. (2015). New ways to boost molecular dynamics simulations. *Journal of Computational Chemistry*, 36(13), 996–1007.
- Kryshtafovych, A., Monastyrskyy, B., & Fidelis, K. (2016). CASP11 statistics and the prediction center evaluation system. *Proteins: Structure, Function, and Bioinformatics*. doi:10.1002/prot.25005.
- Kubiak, K., & Nowak, W. (2008). Molecular dynamics simulations of the photoactive protein nitrile hydratase. *Biophysical Journal*, 94(10), 3824–3838.
- Kuczera, K., Jas, G. S., & Elber, R. (2009). Kinetics of helix unfolding: Molecular dynamics simulations with milestoning. *The Journal of Physical Chemistry. A*, 113(26), 7461–7473.
- Kukic, P., Kannan, A., Dijkstra, M. J., Abeln, S., Camilloni, C., & Vendruscolo, M. (2015). Mapping the protein fold universe using the CamTube force field in molecular dynamics simulations. *PLoS Computational Biology*, 11(10), e1004435.
- Kumar, S., & Li, M. S. (2010). Biomolecules under mechanical force. *Physics Reports*, 486(1–2), 1–74.
- Kumar, A., & Purohit, R. (2014). Use of long term molecular dynamics simulation in predicting cancer associated SNPs. *PLoS Computational Biology*, 10(4), e1003318.
- Kupfer, L., Hinrichs, W., & Groschup, M. H. (2009). Prion protein misfolding. *Current Molecular Medicine*, 9(7), 826–835.
- Kutzner, C., Czub, J., & Grubmuller, H. (2011). Keep it flexible: Driving macromolecular rotary motions in atomistic simulations with GROMACS. *Journal of Chemical Theory and Computation*, 7(5), 1381–1393.
- Lane, T. J., Shukla, D., Beauchamp, K. A., & Pande, V. S. (2013). To milliseconds and beyond: Challenges in the simulation of protein folding. *Current Opinion in Structural Biology*, 23(1), 58–65.
- Lange, O. E., Schafer, L. V., & Grubmuller, H. (2006). Flooding in GROMACS: Accelerated barrier crossings in molecular dynamics. *Journal of Computational Chemistry*, 27(14), 1693–1702.
- Lauria, A., Tutone, M., Ippolito, M., Pantano, L., & Almerico, A. M. (2010). Molecular modeling approaches in the discovery of new drugs for anti-cancer therapy: The investigation of p53-MDM2 interaction and its inhibition by small molecules. *Current Medicinal Chemistry*, 17(28), 3142–3154.
- Le, L., Lee, E., Schulten, K., & Truong, T. N. (2011). Molecular modeling of swine influenza A/H1N1, Spanish H1N1, and avian H5N1 flu N1 neuraminidases bound to Tamiflu and Relenza. *PLoS Currents*, 1, RRN1015. doi:10.1371/currents.RRN1015.
- Leach, A. (2001). *Molecular modelling: Principles and applications* (2nd ed.). Harlow: Prentice Hall.
- Lee, G., Nowak, W., Jaroniec, J., Zhang, Q., & Marszalek, P. E. (2004). Nanomechanical control of glucopyranose rotamers. *Journal of the American Chemical Society*, 126(20), 6218–6219.
- Lee, E. H., Hsin, J., Sotomayor, M., Comellas, G., & Schulten, K. (2009). Discovery through the computational microscope. *Structure*, 17(10), 1295–1306.
- Lee, J., Kim, J.-S., & Seok, C. (2010). Cooperativity and specificity of Cys2His2 Zinc finger protein – DNA interactions: A molecular dynamics simulation study. *The Journal of Physical Chemistry B*, 114(22), 7662–7671.
- Lee, K. H., Kuczera, K., & Banaszak Holl, M. M. (2011). The severity of osteogenesis imperfecta: A comparison to the relative free energy differences of collagen model peptides. *Biopolymers*, 95(3), 182–193.
- Leszczynski, J. (2012). *Handbook of computational chemistry*. Dordrecht/New York: Springer.
- Levitt, M. (2014). Birth and future of multiscale modeling for macromolecular systems (Nobel lecture). *Angewandte Chemie International Edition in English*, 53(38), 10006–10018.
- Levitt, M., & Lifson, S. (1969). Refinement of protein conformation using a macromolecular energy minimization procedure. *Journal of Molecular Biology*, 46, 269–279.

- Li, Y., & Gong, H. (2015). Theoretical and simulation studies on voltage-gated sodium channels. *Protein & Cell*, 6(6), 413–422.
- Li, Q., Wanderling, S., Paduch, M., Medovoy, D., Singharoy, A., McGreevy, R., et al. (2014). Structural mechanism of voltage-dependent gating in an isolated voltage-sensing domain. *Nature Structural & Molecular Biology*, 21(3), 244–252.
- Linder, T., Wang, S., Zangerl-Plessl, E.-M., Nichols, C. G., & Stary-Weinzinger, A. (2015). Molecular dynamics simulations of KirBac1.1 mutants reveal global gating changes of Kir channels. *Journal of Chemical Information and Modeling*, 55(4), 814–822.
- Lindorff-Larsen, K., Maragakis, P., Piana, S., Eastwood, M. P., Dror, R. O., & Shaw, D. E. (2012a). Systematic validation of protein force fields against experimental data. *PLoS One*, 7(2), e32131.
- Lindorff-Larsen, K., Trbovic, N., Maragakis, P., Piana, S., & Shaw, D. E. (2012b). Structure and dynamics of an unfolded protein examined by molecular dynamics simulation. *Journal of the American Chemical Society*, 134(8), 3787–3791.
- Liu, J., & Nussinov, R. (2010). Molecular dynamics reveal the essential role of linker motions in the function of cullin-RING E3 ligases. *Journal of Molecular Biology*, 396(5), 1508–1523.
- Liwo, A., Czaplewski, C., Oldziej, S., & Scheraga, H. A. (2008). Computational techniques for efficient conformational sampling of proteins. *Current Opinion in Structural Biology*, 18(2), 134–139.
- Lonsdale, R., Ranaghan, K. E., & Mulholland, A. J. (2010). Computational enzymology. *Chemical Communications*, 46(14), 2354–2372.
- Lopes, P. E., Guvench, O., & MacKerell, A. D., Jr. (2015). Current status of protein force fields for molecular dynamics simulations. *Methods in Molecular Biology*, 1215, 47–71.
- Lorenz, C., & Doltsinis, N. L. (2012). Molecular dynamics simulation: From “Ab Initio” to “Coarse Grained”. In J. Leszczynski (Ed.), *Handbook of computational chemistry* (pp. 195–238). Dordrecht: Springer.
- Lu, Z., Nowak, W., Lee, G., Marszalek, P. E., & Yang, W. (2004). Elastic properties of single amylose chains in water: A quantum mechanical and AFM study. *Journal of the American Chemical Society*, 126(29), 9033–9041.
- Ma, B., & Levine, A. J. (2007). Probing potential binding modes of the p53 tetramer to DNA based on the symmetries encoded in p53 response elements. *Nucleic Acids Research*, 35(22), 7733–7747.
- Ma, W., & Schulten, K. (2015). Mechanism of substrate translocation by a ring-shaped ATPase motor at millisecond resolution. *Journal of the American Chemical Society*, 137(8), 3031–3040.
- Ma, J., Flynn, T. C., Cui, Q., Leslie, A. G., Walker, J. E., & Karplus, M. (2002). A dynamic analysis of the rotation mechanism for conformational change in F(1)-ATPase. *Structure*, 10(7), 921–931.
- MacKerell, A. D., Jr., & Nilsson, L. (2008). Molecular dynamics simulations of nucleic acid-protein complexes. *Current Opinion in Structural Biology*, 18(2), 194–199.
- MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., et al. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B*, 102(18), 3586–3616.
- Malde, A. K., Zuo, L., Breeze, M., Stroet, M., Poger, D., Nair, P. C., et al. (2011). An automated force field topology builder (ATB) and repository: Version 1.0. *Journal of Chemical Theory and Computation*, 7(12), 4026–4037.
- Marrink, S. J., & Tieleman, D. P. (2013). Perspective on the Martini model. *Chemical Society Reviews*, 42(16), 6801–6822.
- Martin, L., Bilek, M. M., Weiss, A. S., & Kuyucak, S. (2015). Force fields for simulating the interaction of surfaces with biological molecules. *Interface Focus*, 6(1). [10.1098/rsfs.2015.0045](https://doi.org/10.1098/rsfs.2015.0045).
- Martin-Garcia, F., Papaleo, E., Gomez-Puertas, P., Boomsma, W., & Lindorff-Larsen, K. (2015). Comparing molecular dynamics force fields in the essential subspace. *PLoS One*, 10(3), e0121114.

- May, A., Pool, R., van Dijk, E., Bijlard, J., Abeln, S., Heringa, J., et al. (2014). Coarse-grained versus atomistic simulations: Realistic interaction free energies for real proteins. *Bioinformatics*, 30(3), 326–334.
- Mayne, C. G., Saam, J., Schulten, K., Tajkhorshid, E., & Gumbart, J. C. (2013). Rapid parameterization of small molecules using the force field toolkit. *Journal of Computational Chemistry*, 34(32), 2757–2770.
- McCammon, J. A., Gelin, B. R., & Karplus, M. (1977). Dynamics of folded proteins. *Nature*, 267(5612), 585–590.
- McGreevy, R., Singharoy, A., Li, Q., Zhang, J., Xu, D., Perozo, E., et al. (2014). xMDF: Molecular dynamics flexible fitting of low-resolution X-ray structures. *Acta Crystallographica Section D: Biological Crystallography*, 70(9), 2344–2355.
- McGreevy, R., Teo, I., Singharoy, A., & Schulten, K. (2016). Advances in the molecular dynamics flexible fitting method for cryo-EM modeling. *Methods*, 100, 50–60.
- Meirovitch, H. (2007). Recent developments in methodologies for calculating the entropy and free energy of biological systems by computer simulation. *Current Opinion in Structural Biology*, 17(2), 181–186.
- Miao, L., & Schulten, K. (2009). Transport-related structures and processes of the nuclear pore complex studied through molecular dynamics. *Structure*, 17(3), 449–459.
- Miao, Y., Feher, V. A., & McCammon, J. A. (2015). Gaussian accelerated molecular dynamics: Unconstrained enhanced sampling and free energy calculation. *Journal of Chemical Theory and Computation*, 11(8), 3584–3595.
- Mikulska, K., Peplowski, L., & Nowak, W. (2011). Nanomechanics of Ig-like domains of human contactin (BIG-2). *Journal of Molecular Modeling*, 17(9), 2313–2323.
- Mikulska, K., Strzelecki, J., Balter, A., & Nowak, W. (2012). Nanomechanical unfolding of α -neuexin: A major component of the synaptic junction. *Chemical Physics Letters*, 521, 134–137.
- Mikulska, K., Strzelecki, J., & Nowak, W. (2014). Nanomechanics of β -rich proteins related to neuronal disorders studied by AFM, all-atom and coarse-grained MD methods. *Journal of Molecular Modeling*, 20(3), 1–10.
- Miller, B. T., Singh, R. P., Klauda, J. B., Hodoscek, M., Brooks, B. R., & Woodcock, H. L. (2008). CHARMMing: A new, flexible web portal for CHARMM. *Journal of Chemical Information and Modeling*, 48, 1920–1929.
- Monroe, J. I., El-Nahal, W. G., & Shirts, M. R. (2014). Investigating the mutation resistance of nonnucleoside inhibitors of HIV-RT using multiple microsecond atomistic simulations. *Proteins*, 82(1), 130–144.
- Moradi, M., & Tajkhorshid, E. (2014). Computational recipe for efficient description of large-scale conformational changes in biomolecular systems. *Journal of Chemical Theory and Computation*, 10(7), 2866–2880.
- Moraitakis, G., Purkiss, A. G., & Goodfellow, J. M. (2003). Simulated dynamics and biological molecules. *Reports on Progress in Physics*, 66, 483–406.
- Mornon, J.-P., Hoffmann, B., Jonic, S., Lehn, P., & Callebaut, I. (2015). Full-open and closed CFTR channels, with lateral tunnels from the cytoplasm and an alternative position of the F508 region, as revealed by molecular dynamics. *Cellular and Molecular Life Sciences*, 72(7), 1377–1403.
- Morra, G., Meli, M., & Colombo, G. (2008). Molecular dynamics simulations of proteins and peptides: From folding to drug design. *Current Protein and Peptide Science*, 9(2), 181–196.
- Mortier, J., Rakers, C., Bermudez, M., Murgueitio, M. S., Riniker, S., & Wolber, G. (2015). The impact of molecular dynamics on drug design: Applications for the characterization of ligand-macromolecule complexes. *Drug Discovery Today*, 20(6), 686–702.
- Mukherjee, S., & Warshel, A. (2012). Realistic simulations of the coupling between the protomotive force and the mechanical rotation of the F₀-ATPase. *Proceedings of the National Academy of Sciences of the United States of America*, 109(37), 14876–14881.
- Mukherjee, S., & Warshel, A. (2015a). Bronsted slopes based on single-molecule imaging data help to unveil the chemically coupled rotation in F₁-ATPase. *Proceedings of the National Academy of Sciences of the United States of America*, 112(46), 14121–14122.

- Mukherjee, S., & Warshel, A. (2015b). Dissecting the role of the gamma-subunit in the rotary-chemical coupling and torque generation of F1-ATPase. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(9), 2746–2751.
- Nasica-Labouze, J., Nguyen, P. H., Sterpone, F., Berthoumieu, O., Buchete, N.-V., Coté, S., et al. (2015). Amyloid β protein and Alzheimer's disease: When computer simulations complement experimental studies. *Chemical Reviews*, *115*(9), 3518–3563.
- Nguyen, H., Maier, J., Huang, H., Perrone, V., & Simmerling, C. (2014). Folding simulations for proteins with diverse topologies are accessible in days with a physics-based force field and implicit solvent. *Journal of the American Chemical Society*, *136*(40), 13959–13962.
- Nielsen, S. O., Bulow, R. E., Moore, P. B., & Ensing, B. (2010). Recent progress in adaptive multiscale molecular dynamics simulations of soft matter. *Physical Chemistry Chemical Physics*, *12*(39), 12401–12414.
- Noel, J. K., & Whitford, P. C. (2014). How simulations reveal dynamics, disorder, and the energy landscapes of biomolecular function. *Israel Journal of Chemistry*, *54*(8–9), 1093–1107.
- Nowak, W., & Marszalek, P. (2005). *Molecular dynamics simulations of single molecule atomic force microscope experiments*. Singapore: World Scientific.
- Nowak, W., Czerminski, R., & Elber, R. (1991). Reaction path study of ligand diffusion in proteins: Application of the self penalty walk (SPW) method to calculate reaction coordinates for the motion of CO through leghemoglobin. *Journal of the American Chemical Society*, *113*(15), 5627–5637.
- Oliveira, A. S., Damas, J. M., Baptista, A. M., & Soares, C. M. (2014). Exploring O₂ diffusion in A-type cytochrome c oxidases: Molecular dynamics simulations uncover two alternative channels towards the binuclear site. *PLoS Computational Biology*, *10*(12), e1004010.
- Olsen, S., Lamothe, K., & Martinez, T. J. (2010). Protonic gating of excited-state twisting and charge localization in GFP chromophores: A mechanistic hypothesis for reversible photoswitching. *Journal of the American Chemical Society*, *132*(4), 1192–1193.
- Orlowski, S., & Nowak, W. (2007). Locally enhanced sampling molecular dynamics study of the dioxygen transport in human cytoglobin. *Journal of Molecular Modeling*, *13*(6–7), 715–723.
- Orlowski, S., & Nowak, W. (2008). Topology and thermodynamics of gaseous ligands diffusion paths in human neuroglobin. *Biosystems*, *94*(3), 263–266.
- Orozco, M. (2014). A theoretical view of protein dynamics. *Chemical Society Reviews*, *43*(14), 5051–5066.
- Ortore, G., & Martinelli, A. (2012). Computational studies on transthyretin. *Current Medicinal Chemistry*, *19*(15), 2380–2387.
- Paci, E. (2002). High pressure simulations of biomolecules. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, *1595*(1–2), 185–200.
- Paci, E., Caffisch, A., Pluckthun, A., & Karplus, M. (2001). Forces and energetics of haptent-antibody dissociation: A biased molecular dynamics simulation study. *Journal of Molecular Biology*, *314*(3), 589–605.
- Pande, V. S., Baker, I., Chapman, J., Elmer, S. P., Khaliq, S., Larson, S. M., et al. (2003). Atomistic protein folding simulations on the submillisecond time scale using worldwide distributed computing. *Biopolymers*, *68*, 91–109.
- Pantelopulos, G. A., Mukherjee, S., & Voelz, V. A. (2015). Microsecond simulations of mdm2 and its complex with p53 yield insight into force field accuracy and conformational dynamics. *Proteins*, *83*(9), 1665–1676.
- Papaleo, E. (2015). Integrating atomistic molecular dynamics simulations, experiments, and network analysis to study protein dynamics: Strength in unity. *Frontiers in Molecular Biosciences*, *2*, 28.
- Papaleo, E., & Invernizzi, G. (2011). Conformational diseases: Structural studies of aggregation of polyglutamine proteins. *Current Computer-Aided Drug Design*, *7*(1), 23–43.
- Peplowski, L., Kubiak, K., & Nowak, W. (2008). Mechanical aspects of nitrile hydratase enzymatic activity. Steered molecular dynamics simulations of *Pseudonocardia thermophila* JCM 3095. *Chemical Physics Letters*, *467*(1–3), 144–149.

- Perez, A., MacCallum, J. L., & Dill, K. A. (2015). Accelerating molecular simulations of proteins using Bayesian inference on weak information. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(38), 11846–11851.
- Perilla, J. R., Goh, B. C., Cassidy, C. K., Liu, B., Bernardi, R. C., Rudack, T., et al. (2015). Molecular dynamics simulations of large macromolecular complexes. *Current Opinion in Structural Biology*, *31*, 64–74.
- Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., et al. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, *26*(16), 1781–1802.
- Piana, S., Sarkar, K., Lindorff-Larsen, K., Guo, M., Gruebele, M., & Shaw, D. E. (2011). Computational design and experimental testing of the fastest-folding beta-sheet protein. *Journal of Molecular Biology*, *405*(1), 43–48.
- Piana, S., Lindorff-Larsen, K., & Shaw, D. E. (2013). Atomic-level description of ubiquitin folding. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(15), 5915–5920.
- Piana, S., Klepeis, J. L., & Shaw, D. E. (2014). Assessing the accuracy of physical models used in protein-folding simulations: Quantitative evidence from long molecular dynamics simulations. *Current Opinion in Structural Biology*, *24*, 98–105.
- Piana, S., Donchev, A. G., Robustelli, P., & Shaw, D. E. (2015). Water dispersion interactions strongly influence simulated structural properties of disordered protein States. *Journal of Physical Chemistry B*, *119*(16), 5113–5123.
- Piela, L. (2014). Chapter 1 – The magic of quantum mechanics. In *Ideas of quantum chemistry* (2nd ed., pp. 1–59). Oxford: Elsevier.
- Pohorille, A., Jarzynski, C., & Chipot, C. (2010). Good practices in free-energy calculations. *Journal of Physical Chemistry B*, *114*(32), 10235–10253.
- Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., et al. (2013). GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, *29*(7), 845–854.
- Rahman, A., & Stillinger, F. H. (1971). Molecular dynamics study of liquid water. *The Journal of Chemical Physics*, *55*, 3336–3359.
- Rapaport, D. C. (1995). *The art of molecular dynamics simulation*. Cambridge, UK: Cambridge University Press.
- Rauscher, S., Gapsys, V., Gajda, M. J., Zweckstetter, M., de Groot, B. L., & Grubmuller, H. (2015). Structural ensembles of intrinsically disordered proteins depend strongly on force field: A comparison to experiment. *Journal of Chemical Theory and Computation*, *11*(11), 5513–5524.
- Reddy, T., & Sansom, M. S. (2016). Computational virology: From the inside out. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1858*(7, Part B), 1610–1618.
- Reddy, T., Shorthouse, D., Parton, D. L., Jefferys, E., Fowler, P. W., Chavent, M., et al. (2015). Nothing to sneeze at: A dynamic and integrative computational model of an influenza A virion. *Structure*, *23*(3), 584–597.
- Rehm, S., Trodler, P., & Pleiss, J. (2010). Solvent-induced lid opening in lipases: A molecular dynamics study. *Protein Science*, *19*(11), 2122–2130.
- Rico, F., Gonzalez, L., Casuso, I., Puig-Vidal, M., & Scheuring, S. (2013). High-speed force spectroscopy unfolds titin at the velocity of molecular dynamics simulations. *Science*, *342*(6159), 741–743.
- Rief, M., & Grubmuller, H. (2002). Force spectroscopy of single biomolecules. *Chemphyschem*, *3*(3), 255–261.
- Ritchie, D. W. (2008). Recent progress and future directions in protein-protein docking. *Current Protein and Peptide Science*, *9*(1), 1–15.
- Rizzuti, B., & Daggett, V. (2013). Using simulations to provide the framework for experimental protein folding studies. *Archives of Biochemistry and Biophysics*, *531*(1), 128–135.
- Robustelli, P., Kohlhoff, K., Cavalli, A., & Vendruscolo, M. (2010). Using NMR chemical shifts as structural restraints in molecular dynamics simulations of proteins. *Structure*, *18*(8), 923–933.

- Rodrigues, J. R., Simoes, C. J. V., Silva, C. G., & Brito, R. M. M. (2010). Potentially amyloidogenic conformational intermediates populate the unfolding landscape of transthyretin: Insights from molecular dynamics simulations. *Protein Science*, *19*(2), 202–219.
- Rohs, R., West, S. M., Liu, P., & Honig, B. (2009). Nuance in the double-helix and its role in protein–DNA recognition. *Current Opinion in Structural Biology*, *19*(2), 171–177.
- Romanowska, J., Setny, P., & Trylska, J. (2008). Molecular dynamics study of the ribosomal A-site. *Journal of Physical Chemistry B*, *112*(47), 15227–15243.
- Rosales-Hernandez, M. C., Bermudez-Lugo, J., Garcia, J., Trujillo-Ferrara, J., & Correa-Basurto, J. (2009). Molecular modeling applied to anti-cancer drug development. *Anti-Cancer Agents in Medicinal Chemistry*, *9*(2), 230–238.
- Rossle, S. C., & Frank, I. (2009). First-principles simulation of photoreactions in biological systems. *Frontiers in Bioscience*, *14*, 4862–4877.
- Russel, D., Lasker, K., Phillips, J., Schneidman-Duhovny, D., Velazquez-Muriel, J. A., & Sali, A. (2009). The structural dynamics of macromolecular processes. *Current Opinion in Cell Biology*, *21*(1), 97–108.
- Rydzewski, J., & Nowak, W. (2015). Memetic algorithms for ligand expulsion from protein cavities. *The Journal of Chemical Physics*, *143*(12), 124101.
- Rydzewski, J., Jakubowski, R., & Nowak, W. (2015a). Communication: Entropic measure to prevent energy over-minimization in molecular dynamics simulations. *The Journal of Chemical Physics*, *143*(17), 171103.
- Rydzewski, J., Strzalka, W., & Nowak, W. (2015b). Nanomechanics of PCNA: A protein-made DNA sliding clamp. *Chemical Physics Letters*, *634*, 236–242.
- Sagui, C., & Darden, T. A. (1999). Molecular dynamics simulations of biomolecules: Long-range electrostatic effects. *Annual Review of Biophysics and Biomolecular Structure*, *28*(1), 155–179.
- Sakudo, A., Xue, G. A., Kawashita, N., Ano, Y., Takagi, T., Shintani, H., et al. (2010). Structure of the prion protein and its gene: An analysis using bioinformatics and computer simulation. *Current Protein & Peptide Science*, *11*(2), 166–179.
- Salmas, R. E., Yurtsever, M., & Durdagi, S. (2015). Investigation of inhibition mechanism of chemokine receptor CCR5 by micro-second molecular dynamics simulations. *Scientific Reports*, *5*, 13180.
- Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S., & Walker, R. C. (2013). Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle Mesh Ewald. *Journal of Chemical Theory and Computation*, *9*(9), 3878–3888.
- Sanbonmatsu, K. Y. (2012). Computational studies of molecular machines: The ribosome. *Current Opinion in Structural Biology*, *22*(2), 168–174.
- Sanbonmatsu, K. Y., & Tung, C. S. (2007). High performance computing in biology: Multimillion atom simulations of nanoscale systems. *Journal of Structural Biology*, *157*(3), 470–480.
- Sansom, M. S., Scott, K. A., & Bond, P. J. (2008). Coarse-grained simulation: A high-throughput computational approach to membrane proteins. *Biochemical Society Transactions*, *36*(Pt 1), 27–32.
- Saunders, M. G., & Voth, G. A. (2013). Coarse-graining methods for computational biology. *Annual Review of Biophysics*, *41*(42), 73–93.
- Sborgi, L., Verma, A., Piana, S., Lindorff-Larsen, K., Cerminara, M., Santiveri, C. M., et al. (2015). Interaction networks in protein folding via atomic-resolution experiments and long-time-scale molecular dynamics simulations. *Journal of the American Chemical Society*, *137*(20), 6506–6516.
- Scarpazza, D. P., Ierardi, D. J., Lerer, A. K., Mackenzie, K. M., Pan, A. C., Bank, J. A., et al. (2013). Extending the generality of molecular dynamics simulations on a special-purpose machine. In *Ieee 27th international parallel and distributed processing symposium (Ipdps 2013)*, Boston, pp. 933–945.
- Scheraga, H. A., Khalili, M., & Liwo, A. (2007). Protein-folding dynamics: Overview of molecular simulation techniques. *Annual Review of Physical Chemistry*, *58*, 57–83.
- Scheres, S. H. (2010). Visualizing molecular machines in action: Single-particle analysis with structural variability. *Advances in Protein Chemistry and Structural Biology*, *81*, 89–119.

- Schlegel, H. B. (2003). Exploring potential energy surfaces for chemical reactions: An overview of some practical methods. *Journal of Computational Chemistry*, 24(12), 1514–1527.
- Schlick, T. (2010). *Molecular modeling and simulation: An interdisciplinary guide: An interdisciplinary guide*. New York: Springer.
- Schuyler, A. D., Carlson, H. A., & Feldman, E. L. (2009). Computational methods for predicting sites of functionally important dynamics. *Journal of Physical Chemistry B*, 113(19), 6613–6622.
- Schwede, T., & Peitsch, M. C. (2008). *Computational structural biology: Methods and applications*. Hackensack: World Scientific.
- Sen, S., Andreatta, D., Ponomarev, S. Y., Beveridge, D. L., & Berg, M. A. (2009). Dynamics of water and ions near DNA: Comparison of simulation to time-resolved stokes-shift experiments. *Journal of the American Chemical Society*, 131(5), 1724–1735.
- Shakhnovich, E. (2006). Protein folding thermodynamics and dynamics: Where physics, chemistry, and biology meet. *Chemical Reviews*, 106(5), 1559–1588.
- Shaw, D. E., Grossman, J. P., Bank, J. A., Batson, B., Butts, J. A., Chao, J. C., et al. (2014). Anton 2: Raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. In *Paper presented at the proceedings of the international conference for high performance computing, networking, storage and analysis, New Orleans*.
- Sherwood, P., Brooks, B. R., & Sansom, M. S. (2008). Multiscale methods for macromolecular simulations. *Current Opinion in Structural Biology*, 18(5), 630–640.
- Sieben, C., Kappel, C., Zhu, R., Wozniak, A., Rankl, C., Hinterdorfer, P., et al. (2012). Influenza virus binds its host cell using multiple dynamic interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 109(34), 13626–13631.
- Sigg, D. (2014). Modeling ion channels: Past, present, and future. *The Journal of General Physiology*, 144(1), 7–26.
- Simonson, T., Archontis, G., & Karplus, M. (2002). Free energy simulations come of age: Protein-ligand recognition. *Accounts of Chemical Research*, 35(6), 430–437.
- Sothivelvam, S., Liu, B., Han, W., Ramu, H., Klepacki, D., Atkinson, G. C., et al. (2014). Macrolide antibiotics allosterically predispose the ribosome for translation arrest. *Proceedings of the National Academy of Sciences of the United States of America*, 111(27), 9804–9809.
- Sotomayor, M., & Schulten, K. (2007). Single-molecule experiments in vitro and in silico. *Science*, 316(5828), 1144–1148.
- Spiwok, V., Scur, Z., & Hosek, P. (2015). Enhanced sampling techniques in biomolecular simulations. *Biotechnology Advances*, 33(6 Pt 2), 1130–1140.
- Spyrakakis, F., BidonChanal, A., Barril, X., & Luque, F. J. (2011). Protein flexibility and ligand recognition: Challenges for molecular modeling. *Current Topics in Medicinal Chemistry*, 11(2), 192–210.
- Stansfeld, P. J., & Sansom, M. S. (2011a). From coarse grained to atomistic: A serial multiscale approach to membrane protein simulations. *Journal of Chemical Theory and Computation*, 7(4), 1157–1166.
- Stansfeld, P. J., & Sansom, M. S. (2011b). Molecular simulation approaches to membrane proteins. *Structure*, 19(11), 1562–1572.
- Stansfeld, P. J., Goose, J. E., Caffrey, M., Carpenter, E. P., Parker, J. L., Newstead, S., et al. (2015). MemProtMD: Automated insertion of membrane protein structures into explicit lipid membranes. *Structure*, 23(7), 1350–1361.
- Stone, J. E., Phillips, J. C., Freddolino, P. L., Hardy, D. J., Trabuco, L. G., & Schulten, K. (2007). Accelerating molecular modeling applications with graphics processors. *Journal of Computational Chemistry*, 28(16), 2618–2640.
- Straatsma, T. P., & McCammon, J. A. (1992). Computational alchemy. *Annual Review of Physical Chemistry*, 43, 407–435.
- Straub, J. E., & Thirumalai, D. (2010). Toward a molecular theory of early and late events in monomer to amyloid fibril formation. *Annual Review of Physical Chemistry*, 62, 437–463.
- Sugita, Y. (2009). Free-energy landscapes of proteins in solution by generalized-ensemble simulations. *Frontiers in Bioscience*, 14, 1292–1303.

- Sugita, Y., & Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, *314*, 141–151.
- Sun, Q., Doerr, M., Li, Z., Smith, S. C., & Thiel, W. (2010). QM/MM studies of structural and energetic properties of the far-red fluorescent protein HcRed. *Physical Chemistry Chemical Physics*, *12*(10), 2450–2458.
- Tajkhorshid, E., Aksimentiev, A., Balabin, I., Gao, M., Isralewitz, B., Phillips, J. C., et al. (2003). Large scale simulation of protein mechanics and function. *Advances in Protein Chemistry*, *66*, 195–247.
- Tao, P., Hodošček, M., Larkin, J. D., Shao, Y., & Brooks, B. R. (2012). Comparison of three chain-of-states methods: Nudged elastic band and replica path with restraints or constraints. *Journal of Chemical Theory and Computation*, *8*(12), 5035–5051.
- Tatke, S. S., Loong, C. K., D'Souza, N., Schoepfoerster, R. T., & Prabhakaran, M. (2008). Large scale motions in a biosensor protein glucose oxidase: A combined approach by DENS, normal mode analysis, and molecular dynamics studies. *Biopolymers*, *89*(7), 582–594.
- Tautermann, C. S., Seeliger, D., & Kriegl, J. M. (2015). What can we learn from molecular dynamics simulations for GPCR drug design? *Computational and Structural Biotechnology Journal*, *13*, 111–121.
- Tekpinar, M., & Zheng, W. (2013). Coarse-grained and all-atom modeling of structural states and transitions in hemoglobin. *Proteins*, *81*(2), 240–252.
- Towse, C.-L., & Daggett, V. (2013). Protein folding: Molecular dynamics simulations. In G. C. K. Roberts (Ed.), *Encyclopedia of biophysics* (pp. 2020–2025). Berlin/Heidelberg: Springer.
- Tozzini, V. (2010). Multiscale modeling of proteins. *Accounts of Chemical Research*, *43*(2), 220–230.
- Trylska, J. (2010). Coarse-grained models to study dynamics of nanoscale biomolecules and their applications to the ribosome. *Journal of Physics. Condensed Matter*, *22*(45), 453101.
- Tsuduki, T., Tomita, A., Koshihara, S.-Y., Adachi, S.-I., & Yamato, T. (2012). Ligand migration in myoglobin: A combined study of computer simulation and x-ray crystallography. *The Journal of Chemical Physics*, *136*(16), 165101.
- Urbanc, B., Betnel, M., Cruz, L., Bitan, G., & Teplow, D. B. (2010). Elucidation of amyloid beta-protein oligomerization mechanisms: Discrete molecular dynamics Study. *Journal of the American Chemical Society*, *132*(12), 4266–4280.
- Van Der Kamp, M. W., Shaw, K. E., Woods, C. J., & Mulholland, A. J. (2008). Biomolecular simulation and modelling: Status, progress and prospects. *Journal of the Royal Society Interface*, *5*, 173–190.
- van der Kamp, M. W., Schaeffer, R. D., Jonsson, A. L., Scouras, A. D., Simms, A. M., Toofanny, R. D., et al. (2010). Dynameomics: A comprehensive database of protein dynamics. *Structure*, *18*(4), 423–435.
- Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. (2005). GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, *26*(16), 1701–1718.
- van der Vaart, A. (2015). Coupled binding–bending–folding: The complex conformational dynamics of protein-DNA binding studied by atomistic molecular dynamics simulations. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1850*(5), 1091–1098.
- Van Gunsteren, W. F., Bakowies, D., Baron, R., Chandrasekhar, I., Christen, M., Daura, X., Gee, P., Geerke, D. P., Glättli, A., Hünenberger, P. H., Kastenholz, M. A., Oostenbrink, C., Schenk, M., Trzesniak, D., Van Der Vegt, N. F. A., & Yu, H. B. (2006). Biomacromolecular modeling: Goals, problems, perspectives. *Angewandte Chemie International Edition*, *45*, 4064–4092.
- van Oijen, A. M. (2007). Single-molecule studies of complex systems: The replisome. *Molecular BioSystems*, *3*(2), 117–125.
- van Speybroeck, V., & Meier, R. J. (2003). A recent development in computational chemistry: Chemical reactions from first principles molecular dynamics simulations. *Chemical Society Reviews*, *32*(3), 151–157.

- Vanommeslaeghe, K., & MacKerell, A. D., Jr. (2015). CHARMM additive and polarizable force fields for biophysics and computer-aided drug design. *Biochimica et Biophysica Acta*, 1850(5), 861–871.
- Vargas, E., Yarov-Yarovoy, V., Khalili-Araghi, F., Catterall, W. A., Klein, M. L., Tarek, M., et al. (2012). An emerging consensus on voltage-dependent gating from computational modeling and molecular dynamics simulations. *Journal of General Physiology*, 140(6), 587–594.
- Vashisth, H., Skiniotis, G., & Brooks, C. L., 3rd. (2014). Collective variable approaches for single molecule flexible fitting and enhanced sampling. *Chemical Reviews*, 114(6), 3353–3365.
- Vasquez, V., Sotomayor, M., Cordero-Morales, J., Schulten, K., & Perozo, E. (2008). A structural mechanism for MscS gating in lipid bilayers. *Science*, 321(5893), 1210–1214.
- Vemparala, S., Domene, C., & Klein, M. L. (2010). Computational studies on the interactions of inhalational anesthetics with proteins. *Accounts of Chemical Research*, 43(1), 103–110.
- Vicatos, S., Rychkova, A., Mukherjee, S., & Warshel, A. (2014). An effective coarse-grained model for biological simulations: Recent refinements and validations. *Proteins*, 82(7), 1168–1185.
- Villa, E., Balaeff, A., & Schulten, K. (2005). Structural dynamics of the lac repressor-DNA complex revealed by a multiscale simulation. *Proceedings of the National Academy of Sciences of the United States of America*, 102(19), 6783–6788.
- Vuillon, L., & Lesieur, C. (2015). From local to global changes in proteins: A network view. *Current Opinion in Structural Biology*, 31, 1–8.
- Wang, T., & Duan, Y. (2011). Retinal release from opsin in molecular dynamics simulations. *Journal of Molecular Recognition*, 24(2), 350–358.
- Wang, P.-H., Bruschi, M., De Gioia, L., & Blumberger, J. (2013). Uncovering a dynamically formed substrate access tunnel in carbon monoxide dehydrogenase/acetyl-CoA synthase. *Journal of the American Chemical Society*, 135(25), 9493–9502.
- Wang, L. P., Martinez, T. J., & Pande, V. S. (2014). Building force fields: An automatic, systematic, and reproducible approach. *Journal of Physical Chemistry Letters*, 5(11), 1885–1891.
- Wanko, M., Hoffmann, M., Frauenheim, T., & Elstner, M. (2006). Computational photochemistry of retinal proteins. *Journal of Computer-Aided Molecular Design*, 20(7–8), 511–518.
- Warshel, A. (2002). Molecular dynamics simulations of biological reactions. *Accounts of Chemical Research*, 35(6), 385–395.
- Warshel, A. (2003). Computer simulations of enzyme catalysis: Methods, progress, and insights. *Annual Review of Biophysics and Biomolecular Structure*, 32, 425–443.
- Warshel, A. (2014). Multiscale modeling of biological functions: From enzymes to molecular machines (Nobel lecture). *Angewandte Chemie International Edition in English*, 53(38), 10020–10031.
- Warshel, A., & Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology*, 103(2), 227–249.
- Warshel, A., Levitt, M., & Lifson, S. (1970). Consistent force field for calculation of vibrational spectra and conformations of some amides and lactam rings. *Journal of Molecular Spectroscopy*, 33(1), 84–99.
- Warshel, A., Kato, M., & Pisiakov, A. V. (2007). Polarizable force fields: History, test cases, and prospects. *Journal of Chemical Theory and Computation*, 3(6), 2034–2045.
- Weber, D. K., Yao, S., Rojko, N., Anderluh, G., Lybrand, T. P., Downton, M. T., et al. (2015). Characterization of the lipid-binding site of equinatoxin ii by nmr and molecular dynamics simulation. *Biophysical Journal*, 108(8), 1987–1996.
- Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., et al. (1984). A new force-field for molecular mechanical simulation of nucleic-acids and proteins. *Journal of the American Chemical Society*, 106(3), 765–784.
- Wong, V., & Case, D. A. (2008). Evaluating rotational diffusion from protein MD simulations. *Journal of Physical Chemistry B*, 112(19), 6013–6024.
- Wright, P. E., & Dyson, H. J. (1999). Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *Journal of Molecular Biology*, 293(2), 321–331.

- Wright, P. E., & Dyson, H. J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology*, *16*(1), 18–29.
- Yu, J., Ha, T., & Schulten, K. (2007). How directional translocation is regulated in a DNA helicase motor. *Biophysical Journal*, *93*(11), 3783–3797.
- Yuriev, E., Holien, J., & Ramsland, P. A. (2015). Improvements, trends, and new ideas in molecular docking: 2012–2013 in review. *Journal of Molecular Recognition*, *28*(10), 581–604.
- Zerze, G. I. H., Miller, C. M., Granata, D., & Mittal, J. (2015). Free energy surface of an intrinsically disordered protein: Comparison between temperature replica exchange molecular dynamics and bias-exchange metadynamics. *Journal of Chemical Theory and Computation*, *11*(6), 2776–2782.
- Zhang, J., Li, W., Wang, J., Qin, M., Wu, L., Yan, Z., et al. (2009). Protein folding simulations: From coarse-grained model to all-atom model. *IUBMB Life*, *61*(6), 627–643.
- Zhang, L., Lua, L. H. L., Middelberg, A. P. J., Sun, Y., & Connors, N. K. (2015). Biomolecular engineering of virus-like particles aided by computational chemistry methods. *Chemical Society Reviews*, *44*(23), 8608–8618. doi:[10.1039/C5CS00526D](https://doi.org/10.1039/C5CS00526D).
- Zhao, H., & Caffisch, A. (2015). Molecular dynamics in drug design. *European Journal of Medicinal Chemistry*, *91*, 4–14.
- Zhao, G., Perilla, J. R., Yufenyuy, E. L., Meng, X., Chen, B., Ning, J., et al. (2013). Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature*, *497*(7451), 643–646.
- Zhmurov, A., Dima, R. I., Kholodov, Y., & Barsegov, V. (2010). Sop-GPU: Accelerating biomolecular simulations in the centisecond timescale using graphics processors. *Proteins*, *78*(14), 2984–2999.
- Zink, M., & Grubmuller, H. (2009). Mechanical properties of the Icosahedral shell of southern bean mosaic virus: A molecular dynamics study. *Biophysical Journal*, *96*(4), 1350–1363.
- Zoete, V., Cuendet, M. A., Grosdidier, A., & Michielin, O. (2011). SwissParam: A fast force field generation tool for small organic molecules. *Journal of Computational Chemistry*, *32*(11), 2359–2368.
- Zwier, M. C., Adelman, J. L., Kaus, J. W., Pratt, A. J., Wong, K. F., Rego, N. B., et al. (2015). WESTPA: An interoperable, highly scalable software package for weighted ensemble simulation and analysis. *Journal of Chemical Theory and Computation*, *11*(2), 800–809.

Jeremy Curuksu, Srinivasaraghavan Kannan, and Martin Zacharias

Contents

Introduction	1672
Molecular Dynamics Simulations of Nucleic Acids	1673
Induced Conformational Changes During Molecular Dynamics Simulations	1677
Replica-Exchange Molecular Dynamics Simulations	1679
Combining Replica-Exchange and Umbrella Sampling Simulations	1682
Simulation Studies on DNA Bending	1683
Conformational Flexibility of the Kink-Turn Motif in RNA	1685
Conformational Transitions of Nucleic Acid Backbone States	1686
Conclusions	1691
Bibliography	1692

Abstract

Molecular dynamics (MD) simulations based on a classical force field are increasingly being used to study the structure and dynamics of nucleic acids. Simulation studies are limited by the accuracy of the force field description and by the time scale accessible by current MD approaches. In case of specific conformational transitions, it is often possible to improve the sampling of possible states by adding a biasing or umbrella potential along some coordinate describing the conformational transition. It is also possible to extract the associated free energy

J. Curuksu (✉)

Mathematics Department, Swiss Federal Institute of Technology, Lausanne (EPFL), Switzerland
e-mail: Curuksu@mit.edu

S. Kannan

Bioinformatics Institute (A*STAR), Singapore, Singapore
e-mail: raghavk@bii.a-star.edu.sg

M. Zacharias

Physik-Department T38, Technische Universität München, Garching, Germany
e-mail: martin.zacharias@mytum.de; martin.zacharias@ph.tum.de

change along the reaction coordinate. The development of advanced sampling methods such as the replica-exchange MD (REMD) approach allows significant enhancement of conformational sampling of nucleic acids. Recent applications of umbrella sampling and REMD simulation as well as combinations of both methodologies on nucleic acids will be presented. These approaches have the potential to tackle many open questions in structural biology such as the role of nucleic acid structure during recognition and packing and the function of nucleic acid fine structure and dynamics.

Introduction

Nucleic acids have essential functions in a variety of biological processes including storage of genetic information, replication of DNA, regulation of transcription, mRNA splicing, RNA interference, and protein synthesis. Most cellular DNA adopts primarily a double-stranded (B-form) helical structure and the interaction with proteins is mediated through major groove or minor groove recognition. The accessibility of DNA is also controlled by the condensation with nuclear proteins (histones) to compact structures (nucleosomes and chromatin in eukaryotes). Many cellular RNA molecules form more complex three-dimensional structures compared to DNA that consist only partially of double-stranded base-paired regions frequently interrupted by extra-unmatched nucleotides, mismatched base pairs, bulge structures, or hairpin loop structures that cap the end of helices (Zacharias 2000; Leontis and Westhof 2003; Al-Hashimi and Walter 2008; Huang and Lilley 2015).

In recent years, structural knowledge of RNA and RNA in complex with proteins, small organic ligands, and ions has increased dramatically since many new structures have been solved by X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy (Hall 2008; Steitz 2008; Furtig et al. 2003). The rapid increase of structural information has also helped to collect and better understand the types of interactions in folded nucleic acid structures and at nucleic acid–ligand interfaces (Hall 2008; Al-Hashimi and Walter 2008). Since RNA and DNA are of functional importance in many cellular structures, nucleic acids have also been increasingly recognized as possible drug targets. In addition, nucleic acids themselves may serve as drugs (DeJong et al. 2002; Zacharias 2003; Wong et al. 2009). In addition to X-ray crystallography and NMR spectroscopy, the structure determination using electron microscopy under low-temperature cryo conditions (CryoEM) has made dramatic progresses in recent years due to the development of new detectors and computational methods for evaluating EM images (Nogales and Scheres 2015). By averaging over many images in different orientations, it is possible for large nucleic acid molecules in complex with proteins (e.g., ribosome) to determine structures with near-atomic resolution without the requirement of a crystal (Knoops et al. 2012). Furthermore, it is also possible to classify images and define different structural states of a biomolecular assembly. However, this type of classification still results in the determination of static or partially averaged structures of a nucleic acid containing assembly. Possible intermediates or pathways of structural changes

cannot be covered by experimental methods alone at high spatial resolution. In addition, the static experimental structures of nucleic acids alone do not provide the energetic and entropic contributions for structure formation or association with respect to unfolded or unbound reference states. Despite the increased structural knowledge on nucleic acids and complexes with proteins, the structure formation process and the associated driving forces are still poorly understood. Since DNA is mostly in a regular B-form duplex geometry in a cell, the most urgent questions are how specific association with proteins is mediated and how the dynamics and the fine structure (meaning backbone conformations compatible with a B-helix) influence specific binding and packing of DNA. In contrast to DNA, RNA can form a variety of complex three-dimensional (3D) structures (Leontis and Westhof 2003). One can distinguish between secondary structure and tertiary structure, since relatively short duplex structures (minimum length of 4–6 base pairs) are thermodynamically stable and once formed typically do not refold into another structure upon tertiary structure formation of the RNA. Hence, in contrast to the small secondary structure elements in proteins, the secondary structure of RNA can be stable without stabilization by tertiary interactions. On the level of empirical rules and experimentally derived free energy contributions, the energetics and thermodynamics of RNA duplex formation are quite well understood (at least for duplex formation) (Liang and Schlick 2011). Tertiary structure formation corresponds then to the arrangement of secondary structure elements including the sterical constraints imposed by the structural motifs that connect stable secondary structure elements. The forces that drive the packaging of secondary structure elements into complicated and stable 3D structures (e.g., ribosome) are much less well understood (even on an empirical level). The packing of RNA into a compact structure leads to an accumulation of many negative phosphate groups with strong electrostatic repulsion balanced by a corresponding accumulation of counterions (Draper et al. 2005). Hence, 3D structure formation of RNA depends heavily on the surrounding ions and often also on the presence of proteins or other molecules that stabilize the structure. The term “metal core” in relation to the typical hydrophobic core in case of proteins has been introduced to characterize the interior of a large folded RNA molecule (Draper et al. 2005). Comparison with proteins also indicates that folded RNA structures are much less well packed than folded proteins, and numerous waters and ions are still present in the interior of the RNA tertiary structure (Draper 2008). To complement experimental structure determination, the application of computer simulation methods in particular of molecular dynamics (MD) approaches is increasingly being used to study the equilibrium dynamics and the structure formation process.

Molecular Dynamics Simulations of Nucleic Acids

For the theoretical investigation of nucleic acid structure and function, the application of quantum mechanical approaches is currently limited by the large size and complexity of nucleic acids. However, computer simulation studies based on

classical molecular mechanics force fields that use the position of whole atoms as variables (instead of electrons and nuclei in case of quantum mechanics) are used frequently to study the dynamics of nucleic acids (Cheatham 2004; McDowell et al. 2007; MacKerell and Nilsson 2008; Orozco et al. 2008; Kara and Zacharias 2013a). Simulations can complement experimental studies if the time scales or the molecular properties of interest are difficult to access experimentally. Detailed atomistic computational studies of nucleic acid dynamics and even folding of small structural motifs have become possible in recent years (Kara and Zacharias 2013a). In the current chapter, we will focus on the application of advanced sampling molecular dynamics (MD) methods to follow the dynamics of biomolecules at high spatial and time resolution. These methods could be useful to study the driving forces of structure formation, provided that the force field to describe the molecular interactions is sufficiently accurate.

Typically, MD simulation studies include the nucleic acid molecule at atomic resolution. During the simulation Newton's equations of motion are solved numerical in small time steps of $\sim 1-2$ fs ($1-2 \times 10^{-15}$ s). The interatomic interactions are described using a classical molecular mechanics force field of the following form:

$$V_{\text{tot}} = \sum_{i=1}^{\text{Nbonds}} \frac{1}{2} k_{b_i} (b_i - b_{oi})^2 + \sum_{i=1}^{\text{Nangles}} \frac{1}{2} k_{\theta_i} (\theta_i - \theta_{oi})^2 + \sum_{i=1}^{\text{Ndihedral}} \sum_{n=1}^{N\tau} k_n (1 + \cos(n\tau_i + \delta_{ni})) + \sum_{i \neq j}^{\text{Npairs}} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} + \frac{q_i q_j}{r_{ij}} \right) \quad (1)$$

The bonded terms (first three summations in the above equation) contain a sum over all chemical bonds, all bond angles, and dihedral angles of the nucleic acid structures. Usually, simple quadratic penalty terms with appropriate force constants (k_b and k_θ , respectively) are used to control the bond length (b) and bond angles (θ) variations of the molecules. A linear combination of periodic functions is typically employed to control dihedral torsion angles (τ). Additional nonbonded terms describe van der Waals and Coulomb interactions (as a double sum over all nonbonded pairs of atoms, last two terms in the above equation). The form of the energy function allows a rapid evaluation of the potential energy of a molecule and calculation of gradients necessary for energy optimization and molecular dynamics simulations based on a numerical solution of the classical equations of motion.

Simulation studies on biomolecules require the inclusion of surrounding aqueous solvent and ions. However, the inclusion of a sufficiently large number of explicit water molecules significantly increases the computational demand. The possibility to implicitly account for solvent effects in molecular mechanics calculations is of great interest because it would allow longer simulation times and a better convergence of calculated thermodynamic averages. However, in the present review,

we will focus on simulation studies on nucleic acids with an explicit consideration of surrounding ions and water molecules. In order to account for long-range electrostatic interactions, most simulations employ particle mesh Ewald (PME) methods where the long-range part of the electrostatic interaction is considered by an Ewald summation technique for periodic systems (Darden et al. 1993).

Several molecular mechanics force fields are available for simulation studies on nucleic acids. Among the most extensively tested force fields are the Charmm36 (Vanommeslaeghe and MacKerell 2015) and the Amber parmBsc0 (Perez et al. 2007a). In particular the parmBsc0 force field has been tested during very long simulations starting from B-DNA (Perez et al. 2007b). The simulations gave good agreement with the B-DNA starting structure for simulations that were extended to more than 1 μ s total simulation time. In addition, the combined effort of several groups (ABC consortium) has resulted in the characterization of the dynamics of all possible tetra base pair (bp) steps (136) embedded at the center of a longer (18 bp) DNA duplex, each extended for the time range of 1 μ s (Pasi et al. 2014). This gives a detailed picture of the available conformational space of duplex DNA and how it is influenced by all possible sequential arrangements of four consecutive base pairs (Pasi et al. 2014). Starting from standard B-DNA, these structures are equilibrated and simulated under identical conditions to allow direct comparison and extraction of sequence-dependent structural and dynamic features. This effort could eventually lead to a comprehensive description of the sequence-dependent flexibility of DNA.

An extension of the parmBsc0 force field also improves the performance for simulations of RNA molecules (Perez et al. 2007a, b). However, RNA molecules often contain noncanonical and non-regular structural motifs (such as bulges and loop structures) that require further optimization of force field parameters. An important aspect is here the realistic parameterization of conformational dihedral states around the *N*-glycosidic bond that connects nucleobases with the ribose unit (syn- vs. anti-conformations). Considerable improvement has been achieved in the force field description even for noncanonical motifs in recent efforts which requires, however, further evaluation. A further improved force field (extension of parmBsc0), the parmBsc1 force field shows even better agreement with experiment concerning the description of DNA but also allows an improved representation of folded non-helical RNA and DNA motifs (such as hairpin structures) (Ivani et al. 2015).

Also recently, force fields that account approximately for electronic polarization due to the strong electrostatic fields around nucleic acid molecules have become available (Babin et al. 2006). Additional terms often in form of field-induced dipoles on atoms are used to approximately describe the deformation of the electron cloud due to an electrostatic field. It has been shown that such force field models can improve MD simulations of DNA resulting in overall better agreement with available experimental data than the standard force field form with fixed partial charges on atoms and no additional polarization term (Babin et al. 2006).

Nucleic acids are highly solvated systems and are surrounded by nonspecifically and specifically bound ions that stabilize the polyelectrolyte character of the nucleic acid molecule. A number of biophysical techniques allow the investigation of the

effective charge of nucleic acids, and high-resolution X-ray crystallography can provide detailed spatial information on specifically bound solvent molecules and ions. However, the dynamics of nucleic acid hydration and ion binding is difficult to study experimentally at high spatial and time resolution. MD simulation studies have been extensively used to characterize the distribution and lifetimes of bound water molecules and ions in both DNA and RNA (reviewed in Hashem and Auffinger 2007).

In contrast to DNA, most RNA molecules are single stranded in a cell which results in the possible formation of a variety of secondary and tertiary structures. Many larger functional biological assemblies contain not only proteins but also structured RNA components (e.g., ribosomes, spliceosomes, and editosomes). The flexibility of several RNA structural motifs which form building blocks for larger RNA-containing biomolecules has been studied by MD simulations on the nanosecond time scale. Simulations on such motifs start typically from experimentally determined structures and can give important insights into the dynamics and how it is coupled to the function of RNA structural motifs.

An ultimate goal of simulation studies on nucleic acids is to follow the complete structure formation process starting from an unfolded RNA or DNA conformation. As described below this is indeed possible for small structural motifs such as short hairpin structures using atomistic MD simulation methods but still impossible for any larger structure of biological relevance. Since it involves large structures with dozens or even hundreds of nucleotides and involves time scales that are beyond current atomistic simulation approaches, often simplified (coarse-grained) models are used. In such simulations each nucleotide is represented by a reduced number of beads (interaction centers) and the solvent is not modeled explicitly (reviewed in Kara and Zacharias 2013a). However, even more extensive coarse-graining is possible. Coarse-grained representations have, for example, been successfully applied to model the fiber structure of chromatin (reviewed in Sim et al. 2012) with whole nucleosomes (or parts of nucleosomes) as coarse-grained centers. In the current chapter, we will focus on atomistic simulation studies on nucleic acids and methods to improve sampling such that structural transitions occurring on long time scales can still be sampled during advanced sampling MD simulations.

Unfortunately, the conformational sampling during conventional (c)MD simulations is strongly limited by the maximum simulation time that can currently be achieved. Even simulations of hundreds of nanoseconds or few microseconds that are currently possible are often too short to observe significant conformational transitions or the refolding of RNA or DNA structures. The poor sampling is due to kinetic trapping events in low-energy regions of the conformational free energy surface. The rugged nature of the energy hypersurface is in particular due to frequent atomic surface proximity inducing hard-core van der Waals repulsion and due to a restricted allowable set of rotamers in the nucleic acid backbone dihedral angles. Thus, many interesting questions about non-equilibrium dynamics of biomolecules cannot be resolved by current equilibrium cMD simulations.

Induced Conformational Changes During Molecular Dynamics Simulations

In cases where one is interested in specific conformational transitions, it is possible to overcome the sampling problem by inducing a desired conformational change during a simulation. This is possible if one can associate a specific parameter (reaction coordinate) with the conformational change of interest. The choice of such reaction coordinate is in general nontrivial and very task dependent. It is then possible to enforce structural transitions in MD simulations by adding a penalty potential (called umbrella potential) that biases the sampled conformations toward a desired state. The reaction coordinate can, for example, be the bending or stretching of a DNA molecule. Variants of this approach like targeted MD simulations where one enforces movement of atoms toward a desired target structure have also been developed (Sanbonmatsu and Tung 2007).

Typically, in umbrella sampling simulations, one adds a quadratic penalty term to the force field for keeping sampled states close to a desired interval along the reaction coordinate.

The consecutive sampling windows ensure a good probability of sampling along the generalized reaction coordinate. From the distributions of conformations within all the intervals along the reaction coordinate, it is possible to extract the associated free energy change for the transition (Fig. 1). Then an optimal unbiased probability density along the generalized coordinate can be recovered by statistical methods based on the analysis of weighted histograms (WHAM, see reference Kumar et al. 1992). The corrected probability density in turn equates the free energy along the generalized coordinate. The umbrella sampling methodology has been applied frequently to study structural transitions in nucleic acid molecules in recent years. Among these simulations are studies on the energetics of stacking and unstacking of nucleotides (Norberg and Nilsson 1995) and studies of the groove deformation of DNA (Zacharias 2006) which can play a major role for the binding and recognition by proteins. The specific looping out of single bases (base flipping) within a duplex DNA is another process that has been studied using umbrella sampling methods (Giudice et al. 2003; Giudice and Lavery 2003; Huang et al. 2003). Nucleo-base looping out events are of significant biological importance because it forms the mechanistic basis for repairing mispaired or damaged bases by a number of DNA repair enzymes (Dalhus et al. 2009). In addition, many DNA methylases induce looping out of nucleo-bases to perform sequence-specific methylation (Cheng and Blumenthal 2008).

Several umbrella sampling studies have been performed in order to calculate the free energy change associated with looping out a stacked and paired base in duplex DNA and RNA (Giudice et al. 2003; Giudice and Lavery 2003; Huang et al. 2003; Hart et al. 2005; Barthel and Zacharias 2006) and using either a pseudo angular or dihedral angle reaction coordinate. The calculated free energy in the range of 10–18 kcal mol⁻¹ indicates that looping out a paired base in DNA is unfavorable and corresponds to a rare event in case of free duplex DNA. In case of

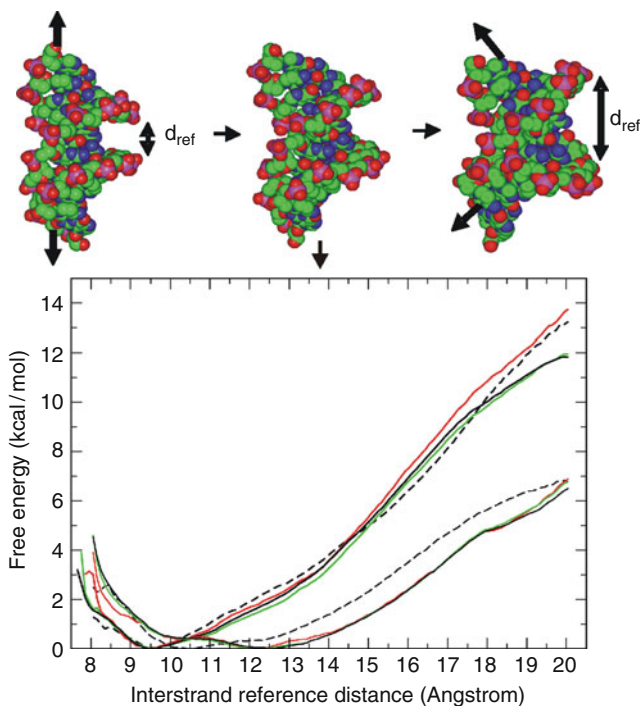


Fig. 1 Illustration of umbrella sampling simulations on induced minor groove opening in two DNA sequences (according to Zacharias 2006). The reaction coordinate for this simulation was the distance between two groups of backbone atoms on opposite DNA strands (*double arrow*). The distance was harmonically restrained to various values close to the reference distance d_{ref} (*upper panel*). Conformational snapshots observed during different stages of the umbrella sampling simulation are shown as van der Waals representations (*upper panel*, atom color code), and the d_{ref} distance and the helical axis directions (*bold arrows*) are indicated. In the *lower panel*, the calculated free energy changes versus distance extracted from the umbrella sampling simulations for a self-complementary 5'-CGCGAAATTCGCG sequence (upper set of curves) and for the 5'-CGCGTATATACGCG sequence (smaller free energy changes) are indicated. The different *line* colors correspond to different simulation lengths (1 and 2 ns per simulation window and combined data) in the forward direction (increasing d_{ref}), and the *dashed line* represents the result for a backward simulation starting from the DNA with the opened minor groove (Further details in Zacharias 2006)

looping out extra-unmatched bases stacked between neighboring base-paired duplex RNA, a significantly smaller barrier for looping out was found and it was possible to characterize stable conformational substates such as base triples that formed during the looping out process (Barthel and Zacharias 2006; Hart et al. 2005). The lower associated free energy of looping out extra-unmatched bases indicates that such conformations can be significantly populated and are in equilibrium with stacked intra-helical states. In addition to localized structural transitions such as looping out bases, the umbrella sampling method has also been used to study global deformations of DNA and RNA molecules. This includes induced minor groove

opening and closing (Zacharias 2006) as well as over- and under-twisting of duplex DNA (Kannan et al. 2006). The opening of the minor groove is essential for the binding of several DNA-binding proteins, among them the prominent example of the TATA box-binding transcription factors (Kim and Burley 1994; Nikolov et al. 1996). Simulation studies on duplexes with a central TATATA or AAATTT sequence motif showed significant differences in the calculated free energies associated with minor groove opening and a much smaller associated free energy change for the TATATA versus AAATTT case (see Fig. 1). The calculations indicated that the free energy change associated with the deformation of DNA toward a form with an open minor groove (as seen in several protein–DNA complexes) can be significant and depends on the sequence. It also allows one to estimate the free energy contribution to the indirect readout of DNA due to differences in deformability.

Replica-Exchange Molecular Dynamics Simulations

The application of classical molecular dynamics (cMD) simulations for studying the dynamics of biomolecules is limited by the accuracy of current force fields and the simulation time scale. Biomolecules like peptides and nucleic acids can adopt a variety of locally stable conformations that are separated by large energy barriers. Standard MD simulation at room temperature may be kinetically trapped in one of these local minima, and conformational transitions between stable states can therefore be rare events on currently accessible time scales. The replica-exchange (parallel tempering) technique is a widely used method to enhance conformational sampling in Monte Carlo (MC) (Swendsen and Wang 1986) and MD simulations (Gnanakaran et al. 2003; Sugita and Okamoto 1999; Kaihsu 2004; Ostermeir and Zacharias 2013; Luitz et al. 2015). In replica-exchange simulations, several copies (replicas) of the system are simulated independently and simultaneously using classical MD or MC methods at different simulation temperatures (or force fields: Hamiltonians). In standard REMD, copies or replicas of the system are simulated at different temperature ($T_0, T_1, T_2, \dots, T_N$). Each replica evolves independently, and after preset intervals (in the range of 500–1000 MD steps ~ 1 –2 ps), an exchange of pairs of neighboring replica is attempted according to a Metropolis acceptance criterion:

$$\begin{aligned} w(x_i \rightarrow x_j) &= 1 && \text{for } \Delta \leq 0; \\ w(x_i \rightarrow x_j) &= \exp(-\Delta) && \text{for } \Delta > 0 \end{aligned} \quad (2)$$

where

$$\Delta = (\beta_i - \beta_j) [E(r_j) - E(r_i)]$$

with $\beta = 1/RT$ (R , gas constant, and T , temperature) and $E(r)$ representing the potential energy of system for a given configuration.

The random walk in temperature allows conformations trapped in locally stable states (at a low simulation temperature) to escape by exchanging with replicas at higher simulation temperature. The REMD method has been successfully applied in

folding simulations of several peptides and mini-proteins (Kaihsu 2004; Kannan and Zacharias 2007a; reviewed in Ostermeir and Zacharias 2013; Luitz et al. 2015) and has more recently also being used for studying nucleic acids (Kannan and Zacharias 2007a; Zhuang et al. 2007; Villa et al. 2008; Kannan and Zacharias 2009; Garcia and Paschek 2008; Villa et al. 2008; Chen and García 2013).

Recently, Kannan and Zacharias (2007a) applied the REMD method to study the folding process of a DNA hairpin structure. Hairpin loops with a central GNA trinucleotide motif (G, guanine; A, adenine; N, any nucleotide) have been found to form particularly stable structures (Yoshizawa et al. 1997). For example, the 5'-GCGCAGC sequence in DNA forms a stable trinucleotide hairpin loop (with a sheared G:A closing pair) flanked by a two-base pair stem helix. The thermodynamic stability of the GCA trinucleotide loop, the influence of loop expansion, the influence of closing and flanking sequences, and also the contribution of individual hydrogen bonds and other nonbonded contacts to the folding stability have been studied extensively (Moody and Bevilacqua 2003).

Temperature-based REMD with 16 replicas in explicit solvent was used to study the structure formations of the 5'-GCGCAGC motif in DNA (Kannan and Zacharias 2007a) starting from single-stranded DNA. After an equilibration phase at constant pressure and temperature, the REMD simulations were conducted under constant volume using 16 replicas ranging in temperature from 315 K to 425 K. During REMD simulations (each 36 ns), conformations in very close agreement with the experimental hairpin structure appeared as dominant conformations during the final phase of the REMD simulations (~35 % at the lowest temperature replica). During REMD simulations already at a simulation time of 12 ns, conformations with an Rmsd of ~2 Å from experiment were sampled. After ~15–20 ns simulation time conformations as close as 1.2–1.6 Å (heavy atoms) with respect to the reference hairpin conformation were sampled as the dominant conformational states (Fig. 2). Simultaneous compaction and accumulation of folded structures were observed. Interestingly, at early stages of the simulations, a variety of central loop conformations but also arrangements close to experiment were sampled before the fully folded structure appeared. The analysis of intermediates at or shortly before the occurrence of fully folded hairpin structures indicated the formation of near-native trinucleotide loop conformations (without fully formed stem). Most of these intermediates included a stacking of the C₂ and G₃ bases which was further stabilized by hydrogen bonding to the A₅ base. Folding to the native hairpin structure appeared to occur almost simultaneously or quickly after the formation of the near-native trinucleotide loop. The folding of a similar DNA hairpin motif has also been successfully achieved in very long continuous MD simulations (Portella and Orozco 2010).

The REMD methodology has also been used for simulation studies of an RNA hairpin and for the folding of an RNA tetraloop r(5'-GCUUCGGC) in explicit solvent starting from a fully extended conformation using 52 replicas (Garcia and Paschek 2008). Apart from the folding process, REMD simulations were used to study the effect of changing the loop sequence and the closing base pair on the conformational distribution and on the thermostability of two RNA hairpins that

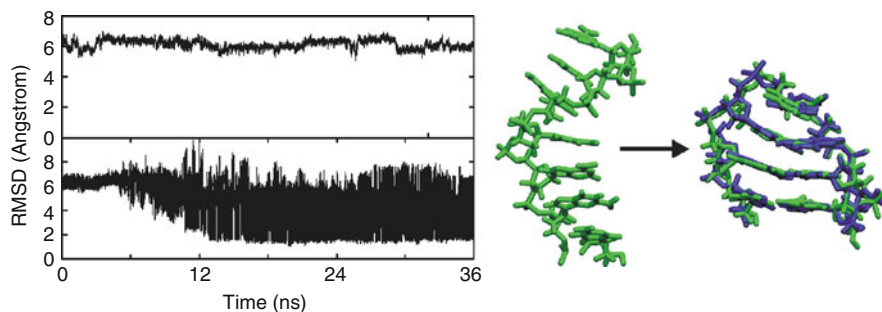


Fig. 2 (Left panel) Heavy atom root mean square deviation (*Rmsd*) of sampled DNA conformations (5'-GCGCAGC) from folded hairpin structure during a continuous MD simulation (*upper graph*) and a T-REMD simulation (*lower graph*) versus simulation time. (Right panel) Single-stranded start structure (*left*) and fully folded hairpin loop structure from the REMD simulation (*green stick model, right side*) superimposed onto the experimental hairpin structure (*blue color*)

have similar structures (Villa et al. 2008). A variation of REMD, called serial replica-exchange molecular dynamics (SREMD), has been employed to gain insight into the folding intermediates of another common RNA tetraloop hairpin with a central GCAA loop sequence (Bowman et al. 2008). Recently, also more complex motifs have been studied using T-REMD (Haldar et al. 2015).

T-REMD will fail if barriers are mostly entropic, and it has also been shown that sampling can even become worse compared to regular cMD for particular systems (Denschlag et al. 2008). The total number of exchanges within a fixed total simulation time can simply be increased by attempting more exchanges per time interval. This is expected to enhance the diffusion in the temperature ladder and in turn may improve the sampling (per simulation time interval). It has also been argued that a minimum time interval of ~ 1 ps should be respected to allow sufficient relaxation of the system after an exchange. However, recent studies suggest that increasing the exchange attempt rate may indeed improve the sampling of relevant states (Sindhikara et al. 2010). It has also been emphasized that the temperature control algorithm itself (the thermostat) of each replica run can influence the efficiency of T-REMD simulations and can create simulation artifacts (Rosta and Hummer 2010). Use of velocity scaling schemes, in particular, may not guarantee canonical sampling in each replica. Also, the schemes of performing exchanges between neighbors can affect the T-REMD performance (Rosta and Hummer 2010). Although the REMD methodology improves sampling in the above-discussed cases, a drawback of the conventional temperature (T-)REMD is the rapid increase of the number of replicas with increasing system size in order to cover a desired temperature range (Fukunishi et al. 2002). The ratio of the standard deviation of the system potential energy (a measure of the energy fluctuation) versus average energy decreases with the square root of the system size. Hence, to achieve sufficient overlap of the energy distributions between replica runs at different temperatures (required to achieve a reasonable exchange acceptance ratio),

the temperature “spacing” between neighboring replicas is required to decrease with system size. Another drawback of large numbers of replicas is the need to run longer simulations (or more exchanges) to allow sufficient “traveling (diffusion)” or exchanges between high- and low-temperature replicas compared to a small number of replicas. Especially in the case of simulations that include a large number of explicit water molecules, the rapid increase of the number of replicas in T-REMD simulations limits the applicability to small bimolecular systems.

Instead of using the simulation temperature as a replica coordinate, one can also use the force field or Hamiltonian of the system as a replica coordinate (Fukunishi et al. 2002; Jang et al. 2003; Liu et al. 2005; Affentranger et al. 2006; Kannan and Zacharias 2007b; Zacharias 2008). In H-REMD simulations it is possible to modify only a part of the Hamiltonian along the replica simulations. The advantage is that then exchanges between replicas depend only on the part of the Hamiltonian that differs between replicas. Consequently such approaches may require much fewer replicas for efficient sampling compared with T-REMD. Various H-REMD methods have been proposed mainly for peptides and proteins and demonstrated enhanced conformational sampling can be achieved with fewer replicas as compared to standard T-REMD (reviewed in Ostermeir and Zacharias 2013; Lutz et al. 2015).

Combining Replica-Exchange and Umbrella Sampling Simulations

Combining umbrella sampling with the replica exchanges is a variant of the H-REMD method (Sugita et al. 2000; Curuksu et al. 2009a) and consists of periodically swapping conformations between adjacent umbrella sampling windows i and j . The exchange criterion takes the form of an equation of micro-reversibility which takes into account the conformational probability density of the two windows i and j , i.e., Boltzmann factors based on the biased potential energy function of window i and window j . One advantage of coupling umbrella sampling with replica exchanges (REUS) is due to the fact that in general there is no obvious best possible coordinate for a given dynamical conformational transition (Sugita et al. 2000), especially when it involves many degrees of freedom (typically bending and folding of biomolecules). As a consequence it may not be possible to overcome all relevant barriers within the phase space associated to a given value of the generalized coordinate, i.e., within each umbrella window. The restriction of sampling only a narrow range along a reaction coordinate can cause conformational trapping of the sampled structures. Through periodic swaps between the sampling windows, the conformations trapped in a low-energy region of the conformational landscape easily diffuse to neighboring windows with a different associated umbrella energy function. Hence, the neighbor replicas provide an enhanced variety of conformations accessible to every sampling window in the limit of a high number of swaps between the windows. In the area of nucleic acids, this methodology has been applied, for example, in a study of adenine riboswitches (Di Palma et al. 2015)

and in simulation studies on the free energy of bending of the kink-turn motif in RNA (Curuksu et al. 2009a).

Simulation Studies on DNA Bending

Bent (i.e., curved) conformational states of DNA over several base pairs can be seen in many protein–DNA complexes (Zakrzewska 2003), in nucleosomal organizations of eukaryotic genetic material (Ong et al. 2007), in the prokaryotic nucleolus (Travers and Muskhelishvili 2005), and in viral DNA packed in capsids (Garcia et al. 2007). Several experimental studies employing cyclization assays (Cloutier and Widom 2004), molecular force sensors (Shroff et al. 2005), FRET (fluorescence resonance energy transfer) (Yuan et al. 2008), and AFM (atomic force microscopy) (Wiggins et al. 2006) experiments indicate that strong DNA bending may result in base pair kinking and a smaller deformation energy than expected from simple elastic models. The occurrence of sharp kinks in MD simulations of short minicircles was reported in (Lankas et al. 2006) with the same DNA sequences as studied experimentally in (Cloutier and Widom 2004).

DNA bending during MD simulations can be imposed by an appropriate restraint between regions flanking the DNA region of interest (Curuksu et al. 2009b) or, alternatively, circularizing this DNA sequence into minicircles (Lankas et al. 2006; Fig. 3). A full dynamical pathway of nucleic acids, global bending can also be characterized by MD simulations combined with enhanced sampling (Curuksu et al. 2008, 2009b), such as the umbrella sampling and replica-exchange (REUS) method introduced in the previous paragraph.

To gain insight into the full dynamical pathway of DNA global bending, our laboratory has carried out umbrella sampling by defining a generalized coordinate

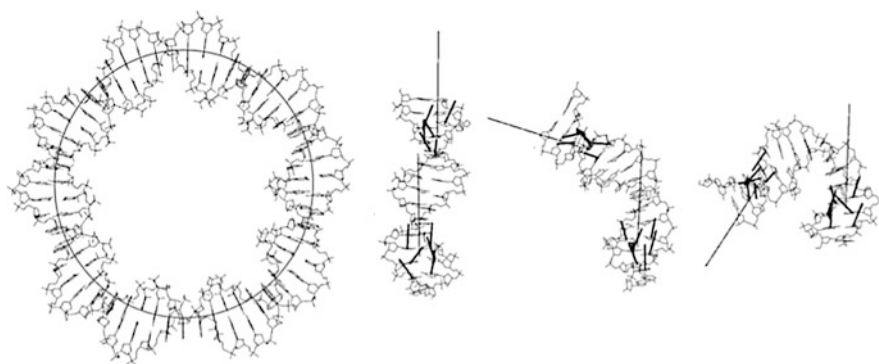


Fig. 3 Geometrical methods to induce bending of nucleic acids molecules. (*Left panel*): “self bent” DNA minicircle system. (*Right panels*): duplex DNA oligonucleotide bent by the load imposed at its two ends when restraining the angle between two vector handles (Curuksu et al. 2008). The vector handles (local helical axis) are indicated as *lines* and the local axis vectors as *short arrows*

for bending short DNA oligomers (Curuksu et al. 2008, 2009b). The bend angle is defined as an angle between two screw axes of the double helical segments referred to as handles. To calculate each screw axis, internucleotide rotation vectors are calculated for several neighboring dinucleotide steps within a fragment of n base pairs. Each rotation vector is based on a rotation matrix that is the total transformation from one nucleotide-fixed axis system into the coordinate frame of its neighboring nucleotide. The sum of the rotation vectors defines a vector “handle” oriented as a function of $(2n-2)$ specific internucleotide rotation angles (Curuksu et al. 2008).

During the simulations described in Curuksu et al. 2009b, four terminal base pairs at each end of a 15-base pair duplex DNA molecule were included to define two (terminal) handles. The global bend angle is given by the angle between the two handle vectors, and a restraining potential of the form $V(\theta) = k(\theta - \theta_{\text{ref}})^2$ was added to the force field (θ corresponds to the angle between the two handles). Since the bend angle restraining potential depends on the position of many atoms (because each handle is influenced by all the atoms that define the local axis system), the restraining force is optimally distributed over all atoms that affect the bending angle. Simulation of several nanoseconds per umbrella sampling window and periodical swapping of conformations between adjacent windows was carried out at regular intervals (every 2 ps). The global bending angle can be smoothly restrained to values around 0° and up to 150° (the most extreme DNA bending regimes as observed in nucleosomes are on the order of 90° for a 15 bp duplex DNA). Figure 4 illustrates the probability histograms of the rotation variable about the long base pair axis between two adjacent base pairs (helical “roll” variable) that characterizes kink motifs through which DNA fragments may relax the bending stress (Lankas et al.

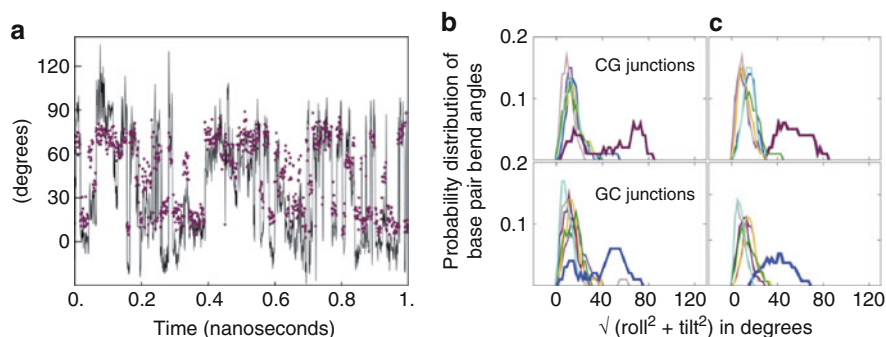


Fig. 4 (a) Time series of roll angle for base pair step C7pG8 (*dots*) and propeller angle at base pair C7:G24 (*line*) in the regime of large induced bending angle (150°) during umbrella sampling *with* replica exchange of the duplex $d(\text{CGCGCGCGCGCGCG})_2$. Roll values were averaged over four adjacent umbrella windows. Base pair kink density histogram observed for simulation windows with large DNA bending strain ($135\text{--}150^\circ$) during umbrella sampling, with (b) and without replica exchanges between neighboring sampling windows (c). Upper set of plots corresponds to roll fluctuations for every pyrimidine–purine base pair steps (CG) and the second set of plots to every purine–pyrimidine base pair steps GC in the oligomer (Unpublished results and Curuksu et al. 2009b)

2006; Curuksu et al. 2009). Reversible partial unstacking events associated with strong local bending kinks at central GC and CG steps were observed in several sampling windows of the REUS application (results for the 150° bending window are shown in Fig. 4a) characterized by a roll angle $>40^\circ$.

The base pair bend density histograms appeared as distributions with two maxima for the two junctions that form a kink (C7G8 in purple and G6C7 in blue) which is not the case in the umbrella sampling without replica exchange (Fig. 4b, c). This means that these two neighboring junctions are both kinked during a significant part of the trajectory as in conventional umbrella sampling [106] but can also adopt stacked conformations (no kink) during another significant part of the same trajectory. This is also indicated in Fig. 4a which illustrates the bend angle of the C7G8 junction and at the propeller twist of the base pair intermediate between G6C7 and C7G8 (C7:G24) as a function of simulation time. Propeller twist is the contrarotation about the long base pair axis within one base pair, and together with the roll deformations above, they were shown to characterize a specific kink motif called type II kink (Lankas et al. 2006). The time series for roll (purple dots) indicates a number of transitions between large bend angles ($\sim 60^\circ$) and small bend angles ($\sim 15^\circ$) corresponding to the two submaxima on the density plot (Fig. 4b) for the C7G8 step. Moreover these transitions occur in concert with propeller transitions (black line) between values typical for B-form DNA and values specific to the type II kink. Thus, the simulations suggested that type II kink is an elastic reversible deformation, which could not have been deduced from a standard umbrella sampling (Curuksu et al. 2009) where no exchange between the sampling windows took place. Clearly, type II kink appears as an important transient DNA motif at the regime of bending probed in the induced bending simulations (up to bend angles of 150°).

Conformational Flexibility of the Kink-Turn Motif in RNA

Recurrent structural motifs can stabilize the tertiary structure of large RNA-containing molecules such as ribosome, spliceosomes, and many other RNA-protein assemblies. Among the most common structures are kink-turn motifs (K-turns) that are characterized by a sharp bend ($\sim 120^\circ$, included angle of $\sim 60^\circ$, also frequently called “V-type” structure) of the phosphodiester backbone between a canonical helix (Watson-Crick base pairs, C-stem) and a helix that contains noncanonical (typically Hoogsteen) base pairs (termed NC stem) (Huang and Lilley 2015). The tip of the “V” structure forms a short asymmetric loop comprising nominally unpaired nucleotides between the two flanking rigid stems. K-turns can also form protein-binding regions and may mediate hinge motion in RNA-containing assemblies. The dynamics of such K-turn motifs have been studied extensively by continuous MD simulations (e.g., Razga et al. 2006) and also by a number of experimental approaches (Goody et al. 2004; Huang and Lilley 2015). However, continuous MD simulations may not cover the full range of possible conformational states for bending and kinking because of the existence of local minima separated by large barriers along a bending coordinate. In order to quantify

the full free energy landscape of K-turn bending, Curuksu et al. (2009) performed umbrella sampling simulations on this system (Fig. 5). A key structural element of the K-turn structure is the A-minor motif (illustrated in Fig. 5b). The A-minor motif stabilizes the characteristic global kinked (V-shaped) conformation of the K-turn. In order to characterize the relation between global structure and stability of the central A-minor motif, a reaction coordinate corresponding to the center-of-mass distance between the central adenine base and the adjacent G:C base pair (Fig. 5b) was employed (Curuksu et al. 2009). A quadratic penalty potential was used to open the A-minor motif, and in order to improve the sampling of relevant conformational states, the authors combined the US approach with H-REMD (H-REUS) allowing for exchanges between neighboring umbrella windows at frequent intervals during the simulations. The application of this H-REUS technique significantly improved the convergence of the calculated free energy landscape along the bending/kinking coordinate for the K-turn structural motif (Fig. 5c, d). It was also possible to identify intermediate states and to investigate the role of water molecules to mediate transitions or stabilize intermediate structures (Curuksu et al. 2009).

Conformational Transitions of Nucleic Acid Backbone States

Various phosphodiester backbone substates can be found in X-ray crystal structures of free nucleic acids and in complexes with proteins (Djuranovic and Hartmann 2004; Varnai et al. 2002). In particular, the coupled ε (rotation around C3'-O3' bond in C4'-C3'-O3'-P) and ζ (rotation around O3'-P bond in C3'-O3'-P-O5') in gauche-/trans and in trans/gauche-regimes, termed B_I and B_{II} substates, respectively, are frequent in crystal structures of free DNA (Varnai et al. 2002).

Since conformational transitions in dsDNA or dsRNA often require transitions of backbone dihedral angles, a generalized coordinate defined as a set of dihedral angles can in principle promote more global transitions. For example, we have designed a replica-exchange method that employs different levels of a specific biasing potential (illustrated in Fig. 6) to induce coupled transitions of the nucleic acid backbone dihedral angles (e.g., ε/ζ dihedral angles or pseudo-dihedral angles; see below) along the replicas (Curuksu and Zacharias 2009). The sampled conformations can exchange with a reference replica that is controlled by the original force field. The choice of biasing potentials in the replicas can be either fixed at the beginning of the simulation or optimized during an equilibration phase. The nucleic acid dihedral angle-dependent biasing potential that destabilizes a given ε/ζ combination takes the following form (penalty term added to the original force field):

$$v_i(\varepsilon, \zeta) = k \times \begin{cases} \left((x-r)^2 - (R-r)^2 \right)^2 & \text{if } r < x < R \\ (R-r)^4 & \text{if } x < r \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

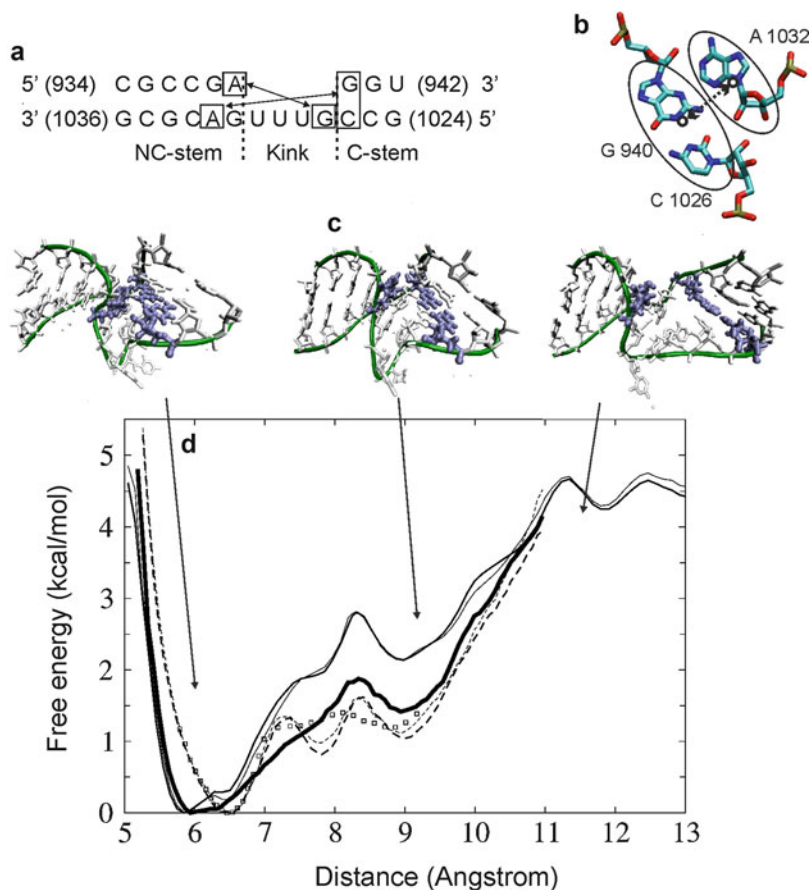


Fig. 5 (a) Sequence and secondary structure of a kink-turn motif (pdb-entry 1jj2) used for umbrella sampling (H-REUS) simulations along a reaction coordinate to separate the central A-minor motif (b) (Curuksu et al. 2009a). The reaction coordinate (distance between central adenine and a contacting guanine:cytosine base pair) was changed in steps of 0.5 Å using a quadratic biasing potential. (c) Illustration of intermediate states sampled during the simulations (A-minor motif in blue sticks) (d) Calculated free energy change along the reaction coordinate (bold line: H-REUS result, hair line and thin lines correspond to US simulations in the forward and backward directions, respectively, along the reaction coordinate)

where $x = \sqrt{(\epsilon - \epsilon_r)^2 + (\zeta - \zeta_r)^2}$ is the coordinate in the two-dimensional ϵ/ζ phase space defined by taking the center of the ceiling (ϵ_r , ζ_r) as origin (see Fig. 6), r is the radius of the ceiling, and R is the radius of the biasing potential. This function has the shape of a quasi-Gaussian with flat ceiling but does not use the computationally expensive exponential function. The width and height of the potential can be independently chosen, and different levels of the potential (with different values of k or E_{\max}) were used along the replicas to destabilize some

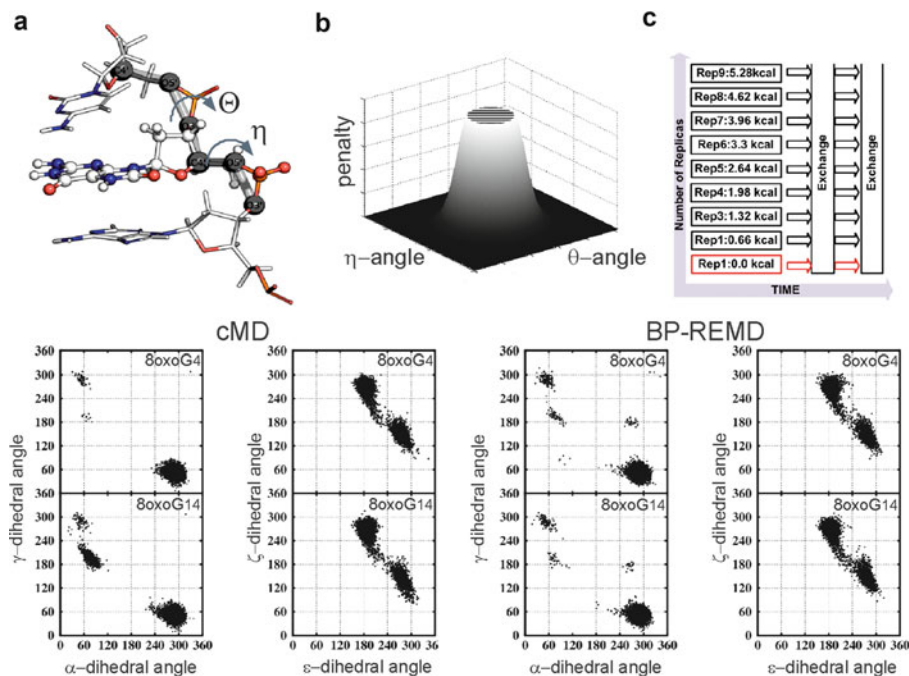


Fig. 6 (a) Illustration of the pseudo-dihedral angles η and θ used to define biasing potentials along replicas in biasing potential H-REMD simulations to promote dihedral transition in the nucleic acid backbone (Kara and Zacharias 2013b). The form of the two-dimensional biasing potentials is illustrated in (b) and the scheme of the BP-REMD simulations in (c). Note that one simulation (out of nine replicas) run without biasing under the control of the original force field. In the *lower panels*, the sampling of α/γ and ε/ζ backbone dihedral angle at the damaged 8oxoG nucleotides (at the two equivalent positions 4 and 14 of the duplex) is illustrated (Kara and Zacharias 2013b) for the cMD simulations (200 ns) and the BP-REMD simulations (10 ns). Each dot in the plots corresponds to sampled dihedral pairs of a conformation sampled every 0.75 ps. Only the sampling in the reference replica (at the original force field in BP-REMD) is recorded. Since the sampling at the two equivalent damaged sites (positions 4 and 14 in the duplex, Kara and Zacharias 2013b) is expected to be the same, the comparison especially of the α/γ states indicates a 10 ns BP-REMD simulation achieves much better convergence than a 200 ns cMD simulation (compare lower plots for the α/γ case)

given ε/ζ substates. Instead of employing a two-dimensional (2D) biasing potential specific for a given pair of ε/ζ dihedral angles, it is also possible to extend the method to employ a biasing potential that acts on pairs of pseudo-dihedral angles (Kara and Zacharias 2013b). The advantage is to cover the coupling of a larger set of dihedral angles embedded in the backbone segment defined by the pseudo-dihedral angles. Kara and Zacharias (2013b) used the nucleic acid backbone atoms $O3'-O5'-C4'-O3'$ in case of the η (here the first $O3'$ belongs to the previous nucleotide) and $C4'-O3'-O5'-C4'$ in case of the θ ($O5'$ and $C4'$ belong to the next nucleotide) pseudo-dihedral angle (illustrated in Fig. 6). Changes in these dihedral angles correlate with the two

most important conformational substates compatible with B-DNA: coupled changes of the α and γ dihedral angles (also called crank shift motions or α/γ -flip) and coupled changes of ϵ and ζ dihedral angles (Perez et al. 2007b).

The method was applied to study the conformational substates of nucleotides damaged by oxidation at the C8 atom of the guanine base (Kara and Zacharias 2013b). Oxidation of nucleobases is one of the most common sources of mutations in DNA and plays a crucial role in mutagenesis of human DNA, carcinogenesis, and aging (Dempfle and Harrison 1994). DNA damage can also alter the fine structure of DNA which may serve as a recognition signal for DNA repair enzymes. The most common and among the most mutagenic oxidative DNA modifications is the formation of 7,8-dihydro-8-oxoguanine, often called 8-oxoguanine. It is a result of the oxidation of the guanine C8, N7 atoms. The damaged guanine has the potential to pair not only with cytosine (C) but also with other bases (most commonly with adenine, A). An 8oxoG can incorrectly pair with A in a Hoogsteen mode (can form during the replication process) which during further replication rounds increases the number of G:C to T:A transversion mutations (Dempfle and Harrison 1994) These transversions are the second most common somatic mutations found in human cancers.

Biasing potential replica-exchange MD (BP-REMD) simulations as described above (see also Fig. 6) were started from regular B-form duplex start structures. The example DNA structures corresponded to palindromic sequences which contain 8oxoG sites in identical sequence context in the two equal strands. Hence, in case of convergence, the same sampling of conformational substates is expected to be identical for both damaged sites (or undamaged sites in case of the reference simulations with G instead of 8oxoG). The same start structures were used for cMD simulations and BP-REMD simulations. The use of nine replicas with biasing potentials on pseudo-dihedral pairs varying between 0 and 5.28 kcal·mol⁻¹ was sufficient to promote transitions between substates and still allowed sufficient exchanges and traveling along the replicas (acceptance rate $\sim 50\%$). The BP-REMD simulations were performed for 10 ns and continuous 200 ns cMD simulations were run for comparison (Kara and Zacharias 2013b). For the sampling of B_I and B_{II} substates, very similar results as for the cMD simulations were observed (Fig. 6) which are consistent with the above conclusion of converged sampling of the B_I/B_{II} ratio of substates during 100–200 ns of cMD simulation. The estimated free energy differences between the B_I and B_{II} substates (1.4 vs. 0.5 kcal·mol⁻¹) in case of damaged 8oxoG obtained from the probability distributions agreed quantitatively with the results from the direct free energy calculations for the dihedral angle transition (Kara and Zacharias 2013b). Also similar to the cMD simulations, the BP-REMD results did not indicate any alternative B_I/B_{II} substates at nucleotides 4 and 14 in case of undamaged DNA (G at positions 4 and 14, Fig. 6).

However, although 10 ns BP-REMD and 200 ns cMD indicated similar distributions for the B_I/B_{II} states at the two damaged nucleotides, the situation is different for the sampling of alternative α/γ -substates (Fig. 6). In case of the 8oxoG-containing duplex, not only substantial sampling of several possible substates at

the 8oxoG was observed but importantly also much closer agreement between the substate distributions at the two equivalent positions 4 and 14 in the duplex (Fig. 6). Due to the rapid exchanges between biased replicas and also frequent exchanges with the reference replica, a much more frequent sampling of alternative states in the reference replica was obtained. It is important to note that the total simulation length of the BP-REMD simulations (10 ns) is substantially shorter than the “lifetime” of alternative α/γ -states observed during the cMD simulations. The result indicates that cMD simulations even of 200 ns length are insufficient to yield converged results on possible nucleic acid α/γ -backbone substates, whereas much shorter BP-REMD simulations are better converged as judged by the close agreement in sampling of two equivalent sites in the self-complementary duplex.

From a stereochemical viewpoint, it is possible to qualitatively explain the observed substate preferences. The transition to a BII state results in an on average larger distance (by 0.6 Å) between the deoxyribose O4 atom of the nucleotide and the H8/O8 atoms of the G/8oxoG nucleo-base. While this is a small average distance change, it can result in a substantial decrease of the electrostatic repulsion in case of an 8oxoG in which case both oxygen atoms are partially negatively charged (in case of a G, the closer distance is electrostatically slightly stabilized because the H8 has partially positive charge). The pseudo-dihedral angle dependent 2D biasing potential specifically accelerates the sampling of ϵ/ζ and α/γ dihedral angle pairs but it was found that it also accelerates indirectly transitions of sugar pucker angles (Kara and Zacharias 2013b).

The BP-REMD method using a biasing potential in the replica runs for adjacent ϵ/ζ and α/γ dihedral angles (Curuksu and Zacharias 2009) was also applied to an experimental structure of an abasic site embedded in double-stranded DNA (seven central base pairs of a DNA abasic site deposited in the Protein Data Bank: PDB-entry: 2HSS; Chen et al. 2007). An abasic site occurs notably when chemical or radiation damage of DNA results in the loss of a base in one strand leaving an unpaired base in the opposite strand (Dempfle and Harrison 1994). From conventional MD simulations, it is known that abasic sites have significantly altered dynamics compared to regular DNA (Fujimoto et al. 2005).

In this study a variety of (ϵ/ζ) dihedral substates at the abasic site were sampled during cMD, T-REMD, and BP-REMD simulations. Both cMD and T-REMD approaches undersampled (or did not sample at all) the (*-gauche/trans*) and (*-gauche/+gauche*) regimes for the abasic site (Curuksu and Zacharias 2009). Alternatively, the application of the BP-REMD methodology to a DNA with a central abasic site significantly improved sampling of the ϵ and ζ dihedral angle substates at overall smaller computational costs. The coupling of the global structure of the DNA with a central abasic site to the (ϵ/ζ) dihedral substate sampling is illustrated in Fig. 7. Each of the sampled substate regions at the abasic site correlated with a different global bending geometry of the DNA molecule. This coupling may have important consequences for the recognition process of damaged DNA by repair enzymes.

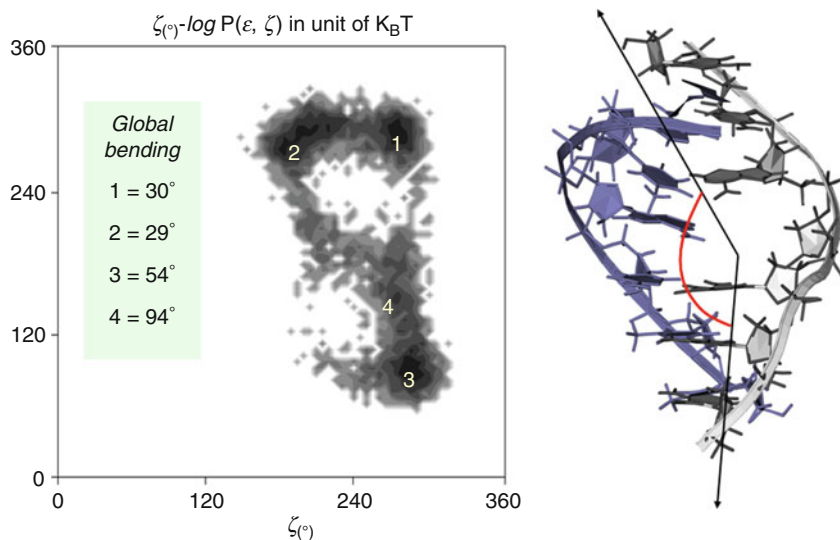


Fig. 7 Coupling of sampled ϵ/ζ dihedral angle substates and global bending of the abasic site [109]. (*Left panel*) Distribution of ϵ/ζ dihedral angles during BP-REMD (replica with the original force field) at the abasic site. The global bend angle of the DNA related to each ϵ/ζ dihedral angle regime is indicated. (*Right panel*) Illustration of the global bend angle in the BP-REMD simulation of a DNA oligonucleotide with central abasic site

Conclusions

The focus of this chapter contribution was the description of Hamiltonian replica-exchange methods to improve umbrella sampling and in combination with a specific biasing potential to promote dihedral transitions in nucleic acid molecules. Umbrella sampling can greatly benefit from frequent exchanges between neighboring sampling intervals to avoid trapping of conformations due to the restraining of the system with respect to a selected regime along a reaction coordinate. This can significantly improve the free energy convergence of such calculations. The technique has a wide range of possible applications to investigate the deformability of DNA and RNA molecules and to better understand its coupling to recognition by proteins and other ligands. The H-REMD technique based on specific biasing potentials to promote dihedral transitions described in the last paragraph can help to better characterize available backbone dihedral angle substates of nucleic acids. In future applications it could help to yield a better understanding of the relation between nucleic acid fine structure and how it is influenced by protein binding or could also be used to improve sampling during nucleic acid folding simulations.

Acknowledgments This work was supported by grants from DFG (Deutsche Forschungsgemeinschaft) to MZ.

Bibliography

- Affentranger, R., Tavernelli, I., & Di Iorio, E. E. (2006). A novel Hamiltonian replica exchange MD protocol to enhance protein conformational space sampling. *Journal of Chemical Theory and Computation*, 2, 217.
- Al-Hashimi, H. M., & Walter, N. G. (2008). RNA dynamics: Its about time. *Current Opinion in Structural Biology*, 18, 321.
- Babin, V., Baucom, J., Darden, T. A., & Sagui, C. (2006). Molecular dynamics simulations of DNA with polarizable force fields: Convergence of an ideal B-DNA structure to the crystallographic structure. *The Journal of Physical Chemistry. B*, 110, 11571.
- Barthel, A., & Zacharias, M. (2006). Conformational transitions in RNA single uridine and adenosine bulge structures: A molecular dynamics free energy simulation study. *Biophysical Journal*, 90, 2450.
- Bowman, G. R., Huang, X., Yao, Y., Sun, J., Carlsson, G., et al. (2008). Structural insight into RNA hairpin folding intermediates. *Journal of the American Chemical Society*, 130, 9676.
- Cheatham, T. E. (2004). Simulation and modeling of nucleic acid structure, dynamics and interactions. *Current Opinion in Structural Biology*, 14, 360.
- Chen, A. A., & Garcia, A. E. (2013). High-resolution reversible folding of hyperstable RNA tetraloops using molecular dynamics simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 16820.
- Chen, J., Dupradeau, F. Y., Case, D. A., Turner, C. J., & Stubbe, J. (2007). Nuclear magnetic resonance structural studies and molecular modeling of duplex DNA containing normal and 4'-oxidized abasic sites. *Biochemistry*, 46, 3096.
- Cheng, X., & Blumenthal, R. M. (2008). Mammalian DNA methyltransferases: A structural perspective. *Structure*, 16, 341.
- Cloutier, T. E., & Widom, J. (2004). Spontaneous sharp bending of double-stranded DNA. *Molecular Cell*, 14, 355.
- Curuksu, J., & Zacharias, M. (2009). Enhanced conformational sampling of nucleic acids by a new Hamiltonian replica exchange molecular dynamics approach. *The Journal of Chemical Physics*, 130, 104110.
- Curuksu, J., Zakrzewska, K., & Zacharias, M. (2008). Magnitude and direction of DNA bending induced by screw-axis orientation: Influence of sequence, mismatches and abasic sites. *Nucleic Acids Research*, 36, 2268.
- Curuksu, J., Sponer, J., & Zacharias, M. (2009a). Elbow flexibility of the kt38 RNA kink-turn motif investigated by free-energy molecular dynamics simulations. *Biophysical Journal*, 97, 2004.
- Curuksu, J., Zacharias, M., Lavery, R., & Zakrzewska, K. (2009b). Local and global effects of strong DNA bending induced during molecular dynamics simulations. *Nucleic Acids Research*, 37, 3766.
- Dalhus, B., Laerdahl, J. K., Backe, P. H., & Bjoras, M. (2009). DNA base repair-recognition and initiation of catalysis. *FEMS Microbiology Reviews*, 33, 1044.
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98, 10089.
- DeJong, E. S., Luy, B., & Marino, J. P. (2002). RNA and RNA-protein complexes as targets for therapeutic intervention. *Current Topics in Medicinal Chemistry*, 2, 289.
- Demple, B., & Harrison, L. (1994). Repair of oxidative damage to DNA: Enzymology and biology. *Annual Review of Biochemistry*, 63, 915.
- Denschlag, R., Lingenheil, M., & Tavan, P. (2008). Efficiency reduction and pseudo-convergence in replica exchange sampling of peptide folding-unfolding equilibria. *Chemical Physics Letters*, 458, 244.
- Di Palma, F., Bottaro, S., & Bussi, G. (2015). Kissing loop interaction in adenine riboswitch: Insights from umbrella sampling simulations. *BMC Bioinformatics*, 16, 1053.
- Djuranovic, D., & Hartmann, B. (2004). DNA fine structure and dynamics in crystals and in solution: The impact of BI/BII backbone conformations. *Biopolymers*, 73, 356.

- Draper, D. E. (2008). RNA folding: Thermodynamic and molecular descriptions of the roles of ions. *Biophysical Journal*, *95*, 5489–5495.
- Draper, D. E., Grilley, D., & Soto, A. M. (2005). Ions and RNA folding. *Annual Review of Biophysics and Biomolecular Structure*, *34*, 221–243.
- Fujimoto, H., Pinak, M., Nemoto, T., O'Neill, P., Kume, E., Saito, K., & Maekawa, H. (2005). Molecular dynamics simulation of clustered DNA damage sites containing 8-oxoguanine and abasic site. *Journal of Computational Chemistry*, *26*, 788.
- Fukunishi, H., Watanabe, O., & Takada, S. (2002). On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *The Journal of Chemical Physics*, *116*, 9058.
- Furtig, B., Richter, C., Wöhnert, J., & Schwalbe, H. (2003). NMR spectroscopy of RNA. *ChemBioChem*, *4*, 936.
- Garcia, A. E., & Paschek, D. (2008). Simulation of the pressure and temperature folding/unfolding equilibrium of a small RNA hairpin. *Journal of the American Chemical Society*, *130*, 815.
- Garcia, H. G., Grayson, P., Han, L., Inamdar, M., Kondev, J., Nelson, P. C., Phillips, R., Widom, J., & Wiggins, P. A. (2007). Biological consequences of tightly bent DNA: The other life of a macromolecular celebrity. *Biopolymers*, *85*, 115.
- Giudice, E., & Lavery, R. (2003). Nucleic acid base pair dynamics: The impact of sequence and structure using free-energy calculations. *Journal of the American Chemical Society*, *125*, 4998.
- Giudice, E., Várnai, P., & Lavery, R. (2003). Base pair opening within B-DNA: Free energy pathways for GC and AT pairs from umbrella sampling simulations. *Nucleic Acids Research*, *31*, 1434.
- Gnanakaran, S., Nymeyer, H., Portman, J., Sanbonmatsu, K. Y., & Garcia, A. E. (2003). Peptide folding simulations. *Current Opinion in Structural Biology*, *15*, 168.
- Goody, T. A., Melcher, S. E., Norman, D. G., & Lilley, D. M. J. (2004). The kink-turn motif in RNA is dimorphic, and metal ion-dependent. *RNA*, *10*, 254–264.
- Haldar, S., Kührová, P., Banáš, P., Spiwok, V., Šponer, J., Hobza, P., & Otyepka, M. (2015). Insights into stability and folding of GNRA and UNCG tetraloops revealed by microsecond molecular dynamics and well-tempered metadynamics. *Journal of Chemical Theory and Computation*, *11*, 3866–3877.
- Hall, K. B. (2008). RNA in motion. *Current Opinion in Chemical Biology*, *12*, 612.
- Hart, K., Nyström, B., Öhman, M., & Nilsson, L. (2005). Molecular dynamics simulations and free energy calculation of base flipping in dsRNA. *RNA*, *11*, 609.
- Hashem, Y., & Auffinger, P. (2007). Nucleic solvation: From outside to insight. *Current Opinion in Structural Biology*, *17*, 325.
- Huang, L., & Lilley, D. M. (2015). The kink turn, a key architectural element in RNA structure. *Journal of Molecular Biology*, *15*, 549.
- Huang, N., Banavali, N. K., & MacKerell, A. D., Jr. (2003). Protein facilitated base flipping in DNA by cytosine-5-methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 68.
- Ivani, I., et al. (2015). Parmbsc1: A refined force field for DNA simulations. *Nature Methods*, *13*, 55–58.
- Jang, S., Shin, S., & Pak, Y. (2003). Replica-exchange method using the generalized effective potential. *Physical Reviews Letters*, *91*, 58305.
- Kaihsu, T. (2004). Conformational sampling for the impatient. *Biophysical Chemistry*, *107*, 213.
- Kannan, S., & Zacharias, M. (2007a). Folding of a DNA hairpin loop structure in explicit solvent using replica-exchange molecular dynamics simulations. *Biophysical Journal*, *93*, 3218.
- Kannan, S., & Zacharias, M. (2007b). Enhanced sampling of peptide and protein conformations using replica exchange simulations with a peptide backbone biasing-potential. *Proteins*, *66*, 697.
- Kannan, S., & Zacharias, M. (2009). Simulation of DNA double-strand dissociation and formation during replica-exchange molecular dynamics simulations. *Physical Chemistry Chemical Physics*, *11*, 10589.

- Kannan, S., Kohlhoff, K., & Zacharias, M. (2006). B-DNA under stress: Over and un-twisting of DNA during molecular dynamics simulations. *Biophysical Journal*, *91*, 2956.
- Kara, M., & Zacharias, M. (2013a). Theoretical studies on nucleic acids folding. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *4*, 116.
- Kara, M., & Zacharias, M. (2013b). Influence of 8-oxoguanosine on the fine structure of DNA studied with biasing-potential replica exchange simulations. *Biophysical Journal*, *104*, 1089.
- Kim, J. L., & Burley, S. K. (1994). 1.9 Å resolution refined structure of TBP recognizing the minor groove of TATAAAAG. *Nature Structural Biology*, *1*, 638.
- Knoops, K., Schoehn, G., & Schaffitzel, C. (2012). Cryo-electron microscopy of ribosomal complexes in cotranslational folding, targeting, and translocation. *Wiley Interdisciplinary Reviews: RNA*, *3*, 429.
- Kumar, S. D., Bouzida, R., Swendsen, H., Kollman, P. A., & Rosenberg, J. M. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, *13*, 1011.
- Laing, C., & Schlick, T. (2011). Computational approaches to RNA structure prediction, analysis, and design Christian Laing and Tamar Schlick. *Current Opinion in Structural Biology*, *21*, 306–318.
- Lankas, F., Lavery, R., & Maddocks, J. H. (2006). Kinking occurs during molecular dynamics simulations of small DNA minicircles. *Structure*, *14*, 1527.
- Leontis, N. B., & Westhof, E. (2003). Analysis of RNA motifs. *Current Opinion in Structural Biology*, *13*, 300.
- Liu, P., Kim, B., Friesner, R. A., & Berne, B. A. (2005). Replica exchange with solute tempering: A method for sampling biological systems in explicit water. *Proceedings of the National Academy of Sciences*, *102*, 13749.
- Luitz, M., Bombliès, R., Ostermeir, K., & Zacharias, M. (2015). Exploring biomolecular dynamics and interactions using advanced sampling methods. *Journal of Physics: Condensed Matter*, *27*, 323101.
- Mackerell, A. D., Jr., & Nilsson, L. (2008). Molecular dynamics simulations of nucleic acid-protein complexes. *Current Opinion in Structural Biology*, *18*, 194.
- McDowell, S. E., Spacková, N., Sponer, J., & Walter, N. G. (2007). Molecular dynamics simulations of RNA: An in silico single molecule approach. *Biopolymers*, *85*, 169.
- Moody, E. M., & Bevilacqua, P. C. (2003). Folding of a stable DNA motif involves a highly cooperative network of interactions. *Journal of the American Chemical Society*, *125*, 16285.
- Nikolov, D. B., Chen, H., Halay, E. D., Hoffman, A., Roeder, R. G., & Burley, S. K. (1996). Crystal structure of a human TATA box-binding protein/TATA element complex. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 4862.
- Nogales, E., & Scheres, S. H. (2015). Cryo-EM: A unique tool for the visualization of macromolecular complexity. *Molecular Cell*, *58*, 677.
- Norberg, J., & Nilsson, L. (1995). Potential of mean force calculations of the stacking-unstacking process in single-stranded deoxyribonucleoside monophosphates. *Biophysical Journal*, *69*, 2277.
- Ong, M. S., Richmond, T. J., & Davey, C. A. (2007). DNA stretching and extreme kinking in the nucleosome core. *Journal of Molecular Biology*, *368*, 1067.
- Orozco, M., Noy, A., & Pérez, A. (2008). Recent advances in the study of nucleic acid flexibility by molecular dynamics. *Current Opinion in Structural Biology*, *18*, 185.
- Ostermeir, K., & Zacharias, M. (2013). Advanced replica-exchange sampling to study the flexibility and plasticity of peptides and proteins. *Biochimica et Biophysica Acta*, *1834*, 847–853.
- Pasi, M., Maddocks, J. H., Beveridge, D., Bishop, T. C., Case, D. A., Cheatham, T., 3rd, Dans, P. D., Jayaram, B., Lankas, F., Laughton, C., Mitchell, J., Osman, R., Orozco, M., Pérez, A., Petkevičiūtė, D., Spackova, N., Sponer, J., Zakrzewska, K., & Lavery, R. (2014). μ ABC: A systematic microsecond molecular dynamics study of tetranucleotide sequence effects in B-DNA. *Nucleic Acids Research*, *42*, 12272–12283.

- Perez, A., Marchan, I., Svozil, D., Spomer, J., Cheatham, T. E., III, Laughton, C. A., & Orozco, M. (2007a). Refinement of the AMBER force field for nucleic acids: Improving the description of $\alpha\gamma$ conformers. *Biophysical Journal*, *92*, 3817.
- Perez, A., Luque, F. J., & Orozco, M. (2007b). Dynamics of B-DNA on the microsecond time scale. *Journal of the American Chemical Society*, *129*, 14739–14745.
- Portella, G., & Orozco, M. (2010). Multiple routes to characterize the folding of a small DNA hairpin. *Angewandte Chemie International Edition in English*, *49*, 7673–7676.
- Razga, F., Zacharias, M., Reblova, K., Koca, J., & Spomer, J. (2006). RNA kink-turns as molecular elbows: Hydration, cation binding, and large-scale dynamics. *Structure*, *14*, 1–11.
- Rosta, E., & Hummer, G. (2010). Error and efficiency of simulated tempering simulations. *The Journal of Chemical Physics*, *132*, 034102.
- Sanbonmatsu, K. Y., & Tung, C. S. (2007). High performance computing in biology: Multimillion atom simulations of nanoscale systems. *Journal of Structural Biology*, *157*, 470.
- Shroff, H., Reinhard, B. M., Siu, M., Agarwal, H., Spakowitz, A., & Liphardt, J. (2005). Biocompatible force sensor with optical readout and dimensions of 6 nm. *Nano Letters*, *5*, 1509.
- Sim, A. Y. L., Minary, P., & Levitt, M. (2012). Modeling nucleic acids. *Current Opinion in Structural Biology*, *22*, 273–278.
- Sindhikara, D. J., Daniel, J., Emerson, D. J., & Roitberg, A. (2010). Exchange often and properly in replica exchange molecular dynamics. *Journal of Chemical Theory and Computation*, *6*, 2804–2808.
- Steitz, T. A. (2008). A structural understanding of the dynamic ribosome machine. *Nature Reviews. Molecular Cell Biology*, *9*, 242.
- Sugita, Y., & Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, *314*, 141.
- Sugita, Y., Kitao, A., & Okamoto, Y. (2000). Multidimensional replica-exchange method for free energy calculations. *The Journal of Chemical Physics*, *113*, 6042.
- Swendsen, R. H., & Wang, J. S. (1986). Replica Monte Carlo simulations of spin glasses. *Physical Reviews Letters*, *57*, 2607.
- Travers, A., & Muskhelishvili, G. (2005). Bacterial chromatin. *Current Opinion in Genetics and Development*, *15*, 507.
- Vanommeslaeghe, K., & MacKerell, A. D., Jr. (2015). CHARMM additive and polarizable force fields for biophysics and computer-aided drug design. *Biochimica et Biophysica Acta*, *1850*, 861.
- Varnai, P., Djuranovic, D., Lavery, R., & Hartmann, B. (2002). Alpha/gamma transitions in the B-DNA backbone. *Nucleic Acids Research*, *30*, 5398.
- Villa, A., Widjajakusuma, E., & Stock, G. (2008). Molecular dynamics simulation of the structure, dynamics, and thermostability of the RNA hairpins uCACGg and cUUCGg. *The Journal of Physical Chemistry. B*, *112*, 134.
- Wiggins, P. A., Van Der Heijden, T., Moreno-Herrero, F., Spakowitz, A., Phillips, R., Widom, J., Ceekers, C., & Nelson, P. C. (2006). High flexibility of DNA on short length scales probed by atomic force microscopy. *Nature Nanotechnology*, *1*, 137.
- Wong, H. M., Payet, L., & Huppert, J. L. (2009). Function and targeting of G-quadruplexes. *Current Opinion in Molecular Therapeutics*, *11*, 146.
- Yoshizawa, S., Kawai, G., Watanabe, K., Miura, K., & Hirao, I. (1997). GNA trinucleotide loop sequences producing extraordinarily stable DNA minihairpins. *Biochemistry*, *36*, 4761.
- Yuan, C., Chen, H., Lou, X. W., & Archer, L. A. (2008). DNA bending stiffness on small length scales. *Physical Reviews Letters*, *100*, 018102.
- Zacharias, M. (2000). Simulation of the structure and dynamics of nonhelical RNA motifs. *Current Opinion in Structural Biology*, *10*, 307.
- Zacharias, M. (2003). Perspectives of drug design that targets RNA. *Current Medicinal Chemistry*, *2*, 161.
- Zacharias, M. (2006). Minor groove deformability of DNA: A molecular dynamics free energy simulation study. *Biophysical Journal*, *91*, 882.

- Zacharias, M. (2008). Combining elastic network analysis and molecular dynamics simulations by Hamiltonian replica exchange. *Journal of Chemical Theory and Computation*, 4, 477.
- Zakrzewska, K. (2003). DNA deformation energetics and protein binding. *Biopolymers*, 70, 414.
- Zhuang, Z., Jaeger, L., & Shea, J. E. (2007). Probing the structural hierarchy and energy landscape of an RNA T-loop hairpin. *Nucleic Acids Research*, 35, 6995.

Model Systems for Dynamics of π -Conjugated Biomolecules in Excited States

39

Mario Barbatti, Matthias Ruckebauer, Jaroslaw J. Szymczak,
Bernhard Sellner, Mario Vazdar, Ivana Antol,
Mirjana Eckert-Maksić, and Hans Lischka

Contents

Introduction	1699
Mixed Quantum-Classical Dynamics Simulations	1700
Protonated Schiff Base Chains and Retinal	1702
The Primary Mechanism of Vision	1702
Heteroaromatic Rings and Nucleic Acid Bases	1709
Photostability of DNA and RNA	1709
Formamide and Peptide Bonds	1721
Probing Photoexcitation of Proteins	1721
Conclusions and Outlook	1727
Bibliography	1729

Abstract

Mixed-quantum classical dynamics simulations have recently become an important tool for investigations of time-dependent properties of electronically excited molecules, including non-adiabatic effects occurring during internal conversion

M. Barbatti (✉)

Institute for Theoretical Chemistry, University of Vienna, Vienna, Austria

e-mail: barbatti@kofo.mpg.de

M. Ruckebauer

Max-Planck-Institut für Kohlenforschung, Mülheim an der Ruhr, Germany

J.J. Szymczak

Research Lab Computational Technologies and Applications, University of Vienna, Vienna, Austria

B. Sellner • I. Antol • M. Eckert-Maksić

Laboratory for Physical–Organic Chemistry – Division of Organic Chemistry and Biochemistry, Rudjer Bošković Institute, Zagreb, Croatia

M. Vazdar • H. Lischka

Department of Chemistry and Biochemistry, Texas Tech University, Lubbock, TX, USA

processes. The high computational costs involved in such simulations have often led to simulation of model compounds instead of the full biochemical system. This chapter reviews recent dynamics results obtained for models of three classes of biologically relevant systems: protonated Schiff base chains as models for the chromophore of rhodopsin proteins; nucleobases and heteroaromatic rings as models for UV-excited nucleic acids; and formamide as a model for photoexcited peptide bonds.

List of Abbreviations

BLA	Bond length alternation
BP	Bicycle pedal
bR	Bacteriorhodopsin
CAS	Complete active space
CASSCF	CAS self-consistent field
CASPT2	CAS second-order perturbation theory
DFT	Density functional theory
DFTB	Density functional based tight binding
ESI	Electrospray ionization
FT	Folding table
GVB,PP	Generalized valence bond in the perfect pairing approximation
HT	Hula-twist
LIIC	Linear interpolation of internal coordinates
MALDI	Matrix-assisted laser-desorption ionization
MCSCF	Multiconfigurational self-consistent field
MCTDH	Multiconfigurational time-dependent Hartree
MRCI	Multireference configuration interaction
MR-CIS	MRCI with single excitations
MR-CISD	MRCI with single and double excitations
MXS	Minimum on the crossing seam
NRBP	Non-rigid bicycle pedal
OBF	One bond flip
OM2	Orthogonalization model 2
PSB	Protonated Schiff base
PSB _n	PSB with n double bonds
QM/MM	Quantum-mechanical/molecular-mechanical
Rh	Rhodopsin
ROKS	Restricted open-shell Kohn-Sham
RPSB	Retinal protonated Schiff base
SA	State average
TDSE	Time-dependent Schrodinger equation
UV	Ultraviolet
WC	Watson-Crick
ZPE	Zero point energy

Introduction

Photoelectronic excitation of organic molecules may trigger complex series of chemical and physical events, which are activated not only by the excess of energy provided by the photon but also by the reactive nature of the excited-state potential energy surfaces. While electronic ground states normally present relatively high barriers for interconversion between minima and strong adiabatic character, excited-state potential energy surfaces tend to show much lower energy barriers and numerous crossings between adiabatic states. The low energy barriers allow the molecule to explore geometric conformations that challenges the chemical intuition. The state crossings bring the molecule to the limit of validity of the Born–Oppenheimer approximation.

Between the photoabsorption and the formation of the final products, the electronic energy excess may be converted into mechanical energy of molecular vibration, isomerization, and dissociation, it may be dissipated as heat into the environment, or it may be reemitted with a red-shifted wavenumber. A measure of how diverse and complex these processes are can be seen from the fact that, depending on the molecule and on the excitation wavelength, they span a range of more than ten orders of magnitude between a few tens of femtoseconds of a radiationless decay and the milliseconds of the phosphorescent decay (Zewail 2000). Besides their intrinsic interest in basic sciences, molecular photoactivated processes are relevant for important natural phenomena and technological applications. In terms of the natural phenomena that will be discussed in this chapter, they constitute the first steps for the primary process of vision (Wald 1968) and the remarkable photostability of the five natural nucleic acid bases may have had important impact on life on Earth (Crespo-Hernandez et al. 2005).

In this chapter we will focus on dynamics simulations for photoexcited molecules and the understanding these simulations may provide. Naturally, dynamics simulations are tailored to answer questions related to time-dependent properties such as time constants of specific processes. In addition, dynamics simulations show which regions of the configuration space are explored and whether they are populated or not. For example, one may frequently find multiple reaction pathways departing from the Franck-Condon region that are all energetically accessible (Barbatti and Lischka 2008). While conventional static calculations with optimization of stationary structures can reveal such pathways but only give indirect information on the importance of each one based on relative energies, dynamics simulations can, in principle, tell much more precisely which one will be preferred and when it will be activated.

The applicability of excited-state dynamics simulations is still limited by its extreme computational demands. Different from ground-state dynamics simulations for which good force-fields and semi-empirical methods are available for computationally efficient simulations, the treatment of molecules in excited states depends very much on costly electronic structure calculations and on hybrid methods like QM/MM approaches. Moreover, when the molecule evolves into state-crossing regions of the configuration space, non-adiabatic corrections are required too

(Worth and Cederbaum 2004). Despite several attempts to develop semiempirical and DFT-based methods to treat the molecule in such regions (Fabiano et al. 2008b; Granucci et al. 2001; Levine et al. 2006; Tapavicza et al. 2007), the most reliable approaches to date are based on the multiconfiguration self-consistent field (MCSCF) and multireference configuration interaction (MRCI) *ab initio* levels.

Currently, most of the dynamics simulations for systems of chemical or biological interest are performed with the on-the-fly approach, in which energy gradients and other properties are computed as needed during the dynamics propagation. This avoids the cumbersome work of building a multidimensional potential energy surface or being forced to make educated guesses about a few important degrees of freedom. The price to pay for the on-the-fly approach is the lack of non-local information during the dynamics. Thus, the use of the on-the-fly, excited-state dynamics simulation approach should be done with an awareness of the risk of improper description of the zero-point energy and phase space distribution (Santoro et al. 2007b). Luckily, as soon as medium- and large-sized molecules move out of the Franck-Condon region, hot species are formed and such quantum effects tend to be reduced.

In view of the computational costs of the on-the-fly, excited-state dynamics simulations, these calculations are currently limited to few picoseconds of a 10- to 20-atom molecule. This is usually enough to investigate ultrafast internal conversion processes and will be the main topic of this chapter.

With the double aim of reducing computational costs as well as of understanding the most basic constituents in photodynamics processes, it has been common to investigate model systems, which are supposed to be the minimal unit that can still reproduce the photodynamics observed in the complete molecule. In the following sections, we shall review three classes of molecules and their model systems: (1) protonated Schiff bases, models for rhodopsin chromophores; (2) heteroaromatic rings, models for nucleic acid bases; and (3) formamide, model for peptide bonds.

Mixed Quantum-Classical Dynamics Simulations

Mixed quantum-classical approaches (Tully 1998) are the most popular methods to perform on-the-fly, excited-state dynamics simulations including non-adiabatic effects. In these approaches, which include the multiple spawning (Ben-Nun et al. 2000), the surface hopping (Tully 1990), and the Ehrenfest (Li et al. 2005) methods, the nuclear time evolution is treated classically by means of Newton's equations, while the time evolution of the population of each electronic state is treated separately, within different degrees of approximation. In the surface hopping method, the time evolution of the population is obtained in two steps: first, a non-adiabatic transition probability between each pair of states is computed and a stochastic algorithm is applied to decide in which state the classical trajectory is propagated in the next time step. Second, statistics over a large set of independently computed trajectories allow acquisition of the fraction of trajectories (occupation) in each state as a function of time.

Several procedures have been proposed to compute non-adiabatic transition probabilities in the surface hopping approach, from the most simple one, which just assumes that the probability is the unity if the energy gap between two states is smaller than some threshold (Hayashi et al. 2009), to more sophisticated approaches that take into account the variation of wavefunction coefficients (Fabiano et al. 2008a) or compute the Landau-Zener transition probability (Lasser and Swart 2008). One of the most reliable procedures for computing the non-adiabatic transition probability is Tully's fewest switches algorithm (Tully 1990). In this approach, the time-dependent Schrodinger equation (TDSE) is integrated simultaneously to the classical trajectory (Ferretti et al. 1996). To cope with the lack of non-local information introduced by the on-the-fly approach, non-local terms in the TDSE are neglected and the nuclear wavefunction is supposed to be entirely localized at the classical position determined by Newton's equation. Integration of this semi-classical version of the TDSE gives the adiabatic population of the states, which are then used to compute the probability using the fewest switches formula.

The integration of the TDSE depends on non-adiabatic coupling terms connecting different states. If an adiabatic representation is used to expand the molecular wavefunction, non-adiabatic coupling vectors need to be computed. Alternatively, if a diabatic representation is used, non-diagonal Hamiltonian elements are required. Either way, the computation of the non-adiabatic coupling terms are the bottleneck in non-adiabatic dynamics approaches. These terms are usually not available for most of quantum chemical methods and, when they are, they are computationally expensive with the cost increasing with the square of the number of states (Pittner et al. 2009). These difficulties have motivated, on one hand, the search for approximated hopping algorithms as those mentioned above, and on the other hand the computation of the coupling terms based on wavefunction overlaps (Fabiano et al. 2008a; Granucci et al. 2001; Hammes-Schiffer and Tully 1994; Pittner et al. 2009; Tapavicza et al. 2007; Werner et al. 2008). We should also point out that the availability of analytical gradients and non-adiabatic coupling vectors procedures at MRCI and MCSCF levels has been of great importance for dynamics simulations (Dallos et al. 2004; Lischka et al. 2004, 2006; Shepard 1995; Shepard et al. 1992).

One consequence of the hyperlocalization of the nuclear wavefunction in this approach is that non-diagonal terms in the density matrix do not vanish with time as they should do (Granucci and Persico 2007; Zhu et al. 2004). Decoherence can be imposed by applying an ad hoc correction to the adiabatic population every time step (Granucci and Persico 2007), which forces the non-diagonal terms in the density matrix to be reduced to zero within a certain time constant.

When a hopping event between two states takes place, this usually happens at a finite energy gap. To keep the total energy constant in the subsequent trajectory, it is necessary to correct the kinetic energy, for example, by rescaling the momentum or by adding more momentum to the direction of the non-adiabatic coupling vector (Ferretti et al. 1996; Muller and Stock 1997). It may also happen that the stochastic algorithm attempts to make a hop from a lower to an upper state in a region where there is not enough energy to do so. Such cases have been usually treated by forbidding the hopping occurrence (Ferretti et al. 1996). The momentum can be kept

or reversed afterwards. Another possibility is to take the time uncertainty principle to search for a geometry nearby where the hopping is allowed (Jasper et al. 2002).

Because of the stochastic nature of the fewest switches surface hopping approach, trajectories starting with the same initial conditions will give rise to different time development. Moreover, the initial conditions should reflect the initial phase space distribution. Therefore, the averages that define the state occupation should in principle be performed over this double ensemble of trajectories starting in different points of the phase space, several times in each one. Because of computational limitations, this procedure is usually reduced to a single ensemble of trajectories starting only once in different points of the phase space.

The ensemble of initial conditions can be generated in a diversity of ways. For instance, the simulation of an instantaneously excited wave packet into the Franck-Condon region may be achieved by selecting geometries and velocities obtained from a dynamics in the ground state, with this dynamics performed for a period long enough as to allow an adequate sampling of the phase space. Alternatively, each nuclear degree of freedom can be treated within the harmonic approximation and a Wigner distribution can be built. In most of the applications discussed here, the second alternative is used since it is expected to provide realistic initial conditions for single molecules.

Several applications will be reviewed and discussed in the following sections. Specific information about methods used will be given. Details can be found in the proper references. The simulations carried out by the present authors have been performed with the Newton-X program (Barbatti et al. 2007a, b) for mixed quantum-classical dynamics simulations, interfaced with the Columbus program (Lischka et al. 1981, 2001, 2006) for MRCI and MCSCF electronic structure calculations.

Protonated Schiff Base Chains and Retinal

The Primary Mechanism of Vision

Understanding the initial molecular mechanism of vision has been one of the most challenging subjects in the area of computational science during the last four decades (Andruniow et al. 2004; Ben-Nun et al. 1998; Birge 1981; Frutos et al. 2007; Garavelli et al. 1999b; González-Luque et al. 2000; Hayashi et al. 2009; Logunov and Schulten 1996; Polli et al. 2010; Rohrig et al. 2005; Saam et al. 2002; Send and Sundholm 2007; Szymczak et al. 2008, 2009; Wanko et al. 2005; Warshel 1976; Warshel and Barbov 1982; Warshel and Chu 2001; Weingart et al. 2007). The photoisomerization of the retinal protonated Schiff base (RPSB), the chromophore of the visual pigment rhodopsin (Rh) (Birge 1981, 1990; Palings et al. 1987; Schoenlein et al. 1991; Wang et al. 1994), followed by conformational changes in the protein is the primary event that initiates vision. This reaction, involving *cis-trans* isomerization of 11-*cis* RPSB (Fig. 1) to its all-*trans* form belongs to the fastest photochemical reaction in nature (Birge 1990). Many theoretical studies have been dedicated to the exploration of the nature of this process, however, the

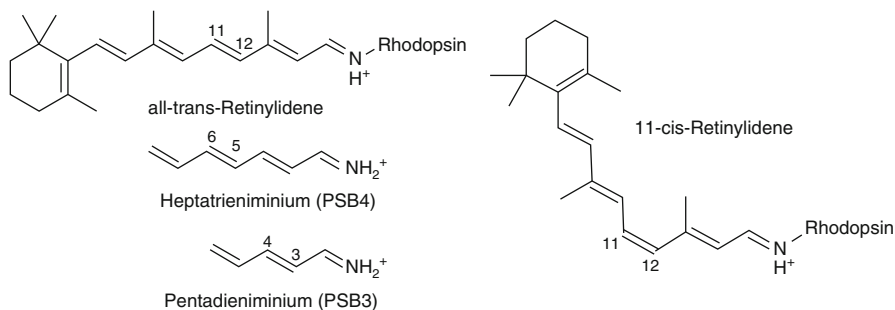


Fig. 1 Retinylidene, the chromophore of the rhodopsin proteins, and the Protonated Schiff Base (PSB) models

question of the detailed mechanism and assignment of experimentally obtained lifetimes remains open, mostly due to the fact that the structures of intermediates are unknown to a large extent. The proper description of the mechanism requires involvement of a high-level theoretical treatment, which is especially challenging due to the large size of RPSB, the multireference character of the electronic wavefunction in regions of interest, and the need of dynamics simulations. Besides vision, photoisomerization in rhodopsin proteins is also important for proton pump processes in archae microorganisms (Birge 1981).

The photochemical properties of RPSB have been studied theoretically, mostly by simulation of model systems. Various mechanisms of the process were proposed in the literature, related strongly to the model used in investigations. In the last two decades, large effort has been dedicated to the theoretical modeling of the Rh photo-behavior in the single-torsion motion called one-bond flip (OBF) as the main isomerization mechanism. These predictions, however, were based mostly on calculations of the minimum energy path or dynamics of the short PSB3 model (Garavelli et al. 1997; Migani et al. 2003; Szymczak et al. 2008; Vreven et al. 1997; Weingart et al. 2004, 2006). On the other hand, investigations on longer PSB chains in earlier works by Warshel (Warshel 1976; Warshel and Barboy 1982) or in more recent investigations (Andruniow et al. 2004; Frutos et al. 2007; Hayashi et al. 2009; Ishida et al. 2009; Polli et al. 2010; Saam et al. 2002; Szymczak et al. 2009) indicate that double-torsion mechanisms like bicycle pedal and nonrigid or asynchronous bicycle pedal should be the main isomerization mechanisms. Also our recent results (Szymczak et al. 2009), based on ab initio dynamics simulations of PSB4 model (discussed below) strongly indicate that RPSB isomerizes in the excited state by performing two concerted twist motions, from which only one tends to be continued during the relaxation into the ground state. Torsional motions of similar character were observed in the recent works by Buss and co-workers (Schapiro et al. 2009; Weingart et al. 2007), whose studies were based on similar four double-bond models augmented by two additional methyl groups in the C₉ and C₁₃ positions. It is also worth mentioning that QM/MM simulations of full RPSB employing the restricted open-shell Kohn-Sham (ROKS) method (Rohrig

et al. 2004) do not support double-torsions mechanism, pointing entirely to OBF as a way of performing photoisomerization in the Rh pocket. At the same time, QM/MM minimum energy path investigations of the same system (Gascon and Batista 2004) report non negligible involvement of four additional torsions to the main torsion in the isomerization process. The uncertainty among the scientific community about the nature of mechanism of RPSB photoisomerization clearly shows the need for further investigations in this field.

Model 1: PSB3

The penta-3,5-dieneiminium cation (PSB3) is the shortest of homologues that are investigated to understand the basic principles of the excited state behavior of retinal. It is the first molecule in the group reaching from methaniminium to retinal that features a double bond with further conjugated double bonds on both sides. At the same time the molecule is small enough to still be treated with high computational methods and medium-sized basis sets, even in dynamics. Therefore it has been subject of extensive theoretical investigations, especially by the groups of Robb, Buss, and Olivucci, which have contributed to the understanding of the excited state behavior and deactivation pathways of PSB3 in gas and solvated phases (Barbatti et al. 2007a, 2008; Garavelli et al. 1999a; Keal et al. 2009; Migani et al. 2003; Szymczak et al. 2008; Weingart et al. 2004, 2005).

Most dynamics investigations of PSB3 have been performed on a CASSCF level including all π orbitals in the active space with 6-31G/6-31G* or similar basis sets. The gas phase dynamics of PSB3 can be explained in a two-state, two-mode model (Garavelli et al. 1997; González-Luque et al. 2000; Migani et al. 2003) involving the ground and the first excited ($\pi\pi^*$) states and two dominating motions, a torsion around the central double bond, and the adjustment of the bond length alternation (BLA).

After vertical excitation into the first excited state the bond pattern adjusts to the changed electronic structure. The double/single bond distinction of the ground state is lifted and the BLA, being expressed as the difference between average single bond length and average double bond length, which is somewhat positive at the beginning, drops quickly to a larger negative value. This is accompanied by a shift of the highest occupied orbitals, π and π^* . They are separated from each other at the respective ends of the PSB3 molecule and, each occupied with one electron, induce a torsional motion around the central double bond of the molecule, in a mechanism known as one-bond-flip (Szymczak et al. 2008). This motion leads exactly to a conical intersection that is reached at a perpendicular orientation of the two halves of the molecule and small BLA. The excited state lifetime is about 100 fs. The high torsional momentum leads to a quantum yield of isomerization of PSB3 after photoexcitation of about 60%, regardless of the conformation the process started in, *cis* or *trans*. When restrictions are applied to the extremes of PSB3, another relaxation mechanism starts to compete with the one-bond-flip (Szymczak et al. 2008). This mechanism, named folding table, involves the torsion around the central bond occurring simultaneously to the semi-torsions around the other two formal double bonds.

Surface hopping dynamics simulations of PSB3 with the semiempirical OM2 method (Keal et al. 2009) show a picture very distinct from the CASSCF simulations. Depending on the choice of active space, the excited state relaxation shows a bi-exponential decay profile of the S_1 population, with a fast sub-picosecond time constant and a picosecond time constant. Overall, the relaxation process is predicted to be larger than 600 fs, much slower than the 100 fs predicted by CASSCF. Similar multi-exponential decay has also been described in wave packet propagation on a two-dimensional surface model for RPSB (Santoro et al. 2007b).

Inclusion of a solvent has the potential to change the dynamics drastically. A wide, space-requiring motion of a molecule would be hindered by collisions with the solvent molecules and different deactivation channels, which are normally not accessed, can become the primary path.

To assess the steric environmental effects on the photochemical deactivation of PSB3 it was embedded in a small sphere of n-hexane molecules at room temperature using a QM/MM scheme (Ruckebauer et al. 2010). The surface hopping dynamics simulations discussed here were computed at the MRCIS(4,5)/SA3CASSCF(6,6)/631G level, where 11 orbitals were frozen during the MR-CIS procedure. For solvation, the same quantum mechanical level was employed and 150 n-hexane molecules were introduced into the calculation using electrostatic embedding with the OPLSAA/TIP3P force field (Jorgensen and McDonald 1998).

The comparison of dynamics simulations in gas phase to simulations in n-hexane shows that the relaxation process is not too much affected by solvation. PSB3 is still too small a molecule to show steric hindering of the isomerization. Figure 2, right, shows a snapshot of one trajectory for the twisted PSB3 in n-hexane. It clearly demonstrates that PSB3 rotates in a quite volume conserving way, which minimizes the interactions with the solvent. As a consequence, the lifetime is roughly the same in gas phase and n-hexane (Fig. 2, left), in n-hexane slightly longer than in gas phase simulations, but the effect is limited to a difference of about 3 fs, which is statistically not significant.

Model 2: PSB4

The recent progress in computer capacity and in computational methods made it possible to address by high-level ab initio dynamics more realistic models than the just-described PSB3. In one of our latest investigations (Szymczak et al. 2009), the photodynamics of the hepta-3,5,7-trieniminium cation (PSB4) is investigated by performing extended on-the-fly surface hopping dynamics calculations in vacuo. This compound exhibits a significantly enhanced torsional variability while keeping the computational effort for the on-the-fly surface hopping calculations manageable so that good statistics within the framework of ab initio approaches can be achieved. Similar to our previous work on PSB3 (Szymczak et al. 2008), mechanical restrictions corresponding to the protein link on one PSB4 end and representing the remaining RPSB molecule on the other were applied by increasing masses of corresponding hydrogens to 1,000 amu. The constrained 5-cis isomer of PSB4 was investigated in order to model RPSB in the rhodopsin; dynamics results for doubly

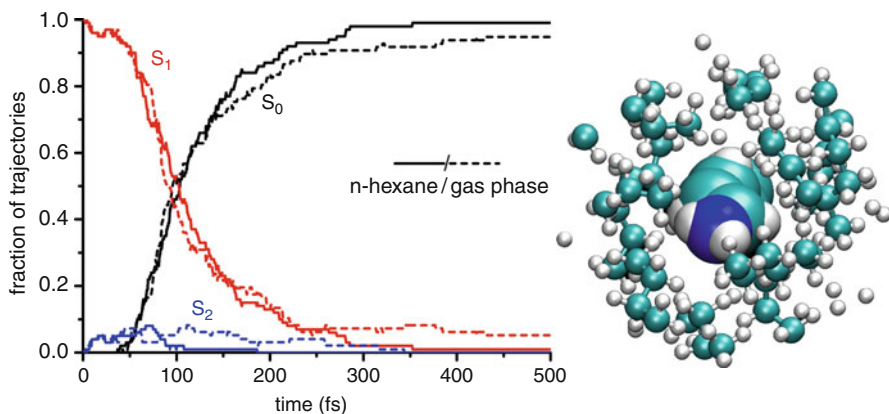


Fig. 2 Average adiabatic populations of PSB3 in gas phase and n-hexane (*left*) and twisted PSB3 in n-hexane (*right*)

restricted PSB4 were then compared to time-dependent femtosecond-resolved experiments (Kochendoerfer and Mathies 1996; Kukura et al. 2005; Peteanu et al. 1993; Schoenlein et al. 1991).

Dynamic simulations were performed at CASSCF level. Eight electrons were included in the calculations within an active space consisting of eight orbitals (four π and the corresponding four π^* orbitals) and averaged over two states [SA-2-CASSCF(8,8)]. To validate the correctness of the method employed, calculations of the vertical excitation energies and conical intersections (MXS) were performed also at MRCI level. Additionally, calculations of LIIC paths from the S_1 minimum to the MXSs give similar, very satisfactory accord between different methods. Thus, the applied theoretical level has been carefully chosen and verified to balance quality of the approach and the computational cost of few hundred thousands of quantum chemical calculations (energies, gradients, and non-adiabatic coupling vectors).

The photorelaxation of PSBn chains is well described by the two-state, two-mode model discussed in the previous section (Garavelli et al. 1997; González-Luque et al. 2000; Migani et al. 2003). After photoexcitation, in the initial stage of the dynamics the PSBn system relaxes by adjusting the bonds lengths, elongating the double bonds and shortening the single bonds. This phase, common to all currently studied PSBn systems, has been described in the previous section. After adjusting all bonds, PSBn proceeds to the second step of the dynamics. The crossing seam is reached by skeletal torsions around one or more bonds. When the molecular system switches from the excited state to the ground state it can either continue or reverse its motion that led to the crossing seam and further relax to the final product geometry.

The kind of torsional motion taking place in PSB4 was assigned to one of five main groups (see Fig. 3): one-bond flip (OBF), bicycle pedal (BP) (Warshel 1976),

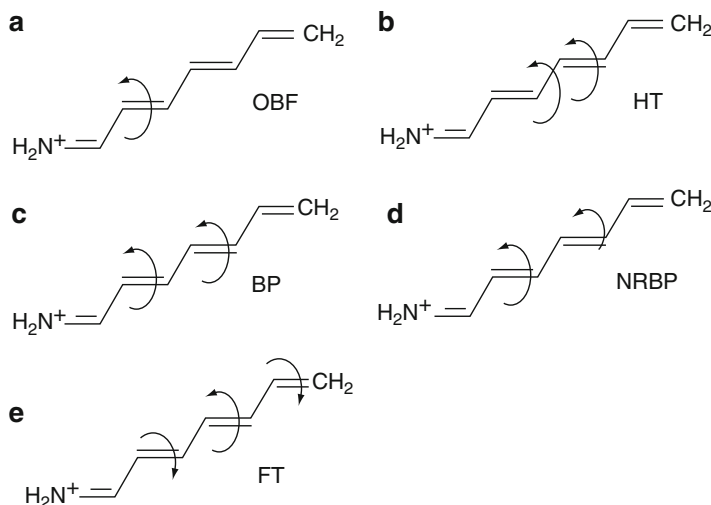


Fig. 3 Torsional motions involved in the main isomerization mechanisms: (a) OBF; (b) HT; (c) BP; (d) NRBP; (e) FT

non-rigid bicycle pedal (NRBP) (Warshel and Barboy 1982), hula-twist (HT) (Liu and Asato 1985), and the recently proposed folding table (FT) (Szymczak et al. 2008) mechanisms. A classification scheme was developed (Szymczak et al. 2008, 2009) facilitating the automatic assignment of these structures. It divides the analysis of the torsional dynamics into two stages: motions driving the system to the conical intersection and motions after decaying to the ground state, with independent possibility of investigation of the product formation. The advantage of the proposed scheme is that this classification algorithm is general enough to be applied for the massive amount of information coming from dynamics simulations. It can also be applied for investigations of longer PSBs and can even be further used to study behavior of complete RPSB.

The analysis of the first stage of the process, the mechanisms leading 5-cis-PSB4 to the seam of conical intersections, shows the coexistence of two groups of mechanisms. The first one includes OBF and FT mechanisms and proceeds as such to produce a complete torsion around only one bond, mostly around the one equivalent to the $C_{11}C_{12}$ bond in RPSB. The other rotation proceeds around the $C_{13}C_{14}$ bond, however, its importance is significantly smaller (less than 20% of cases featuring single torsion mechanism). All together, OBF and FT constitute around one third of motions driving 5-cis-PSB4 to the conical intersection. The second group, double-torsion mechanisms, is dominating and appears about twice as frequent as the single-torsion group. It includes BP and NRBP, which are characterized by performing concerted twists around two bonds. Within this group twice as many trajectories feature the NRBP than the BP mechanism. Not a single case of hula-twist motion was observed during the dynamics simulations of 5-cis-PSB4, no matter what was the degree of restrictions (Szymczak et al. 2009).

When returning to the ground state, the excess electronic energy is converted into mechanical energy. In this vibrationally hot ground state there will be freedom to move into different conformations, which in normal thermal situations are usually confined by high potential energy barriers. Since the transfer of the mechanical energy to the protein will take at least several picoseconds (Birge 1981), the motion in the hot ground state may be decisive for the actual quantum yields. The behavior of PSB4 after decaying to the ground state was analyzed from two points of view: firstly, whether the mechanism bringing PSB4 to the intersection seam is reversed or not after hopping; secondly, whether the initial motion results in the associated products or not. Here we concentrate only on the double-torsion mechanisms. The most abundant way of continuation of concerted double-torsion motions is not a complete continuation or reversion, but partial continuation. Interestingly, the second torsion involved in concerted motion is the one equivalent to the $C_{13}C_{14}$ bond in RPSB. The same pattern was observed by Buss and co-workers (Schapiro et al. 2009; Weingart et al. 2007), whose studies were based on a similar four double-bond model PSB4 augmented by two additional methyl groups in the C_9 and C_{13} positions. On the hand, available studies on complete RPSB (Hayashi et al. 2009; Ishida et al. 2009; Warshel 1976; Warshel and Barboy 1982) point to C_9C_{10} as the other torsion involved in the concerted twist mechanism. This difference may suggest that either the chain lengths of the systems studied is still too short for representing complete RPSB or that the methyl group at the C_{13} plays important role in blocking the given torsion in the protein pocket. The full continuation (rotation around both involved double-bonds) of double-torsion group of mechanisms appears only in 22 % of NRBP cases and 10 % of the BP cases. But even when full continuation is observed, the associated photoproduct is formed in half or less of the cases because of the hot ground state isomerization processes. The overall picture that can be drawn from our simulations is the one in which 5-cis-PSB4 performs concerned twists around two double bonds to reach the region of conical intersection of which only one is continued after decaying to the ground state.

Although we expect that inclusion of at least one more double bond to PSB4 model may affect the relative importance of mechanisms or may even reveal new mechanisms, one can clearly see qualitative advance in comparison to the PSB3 case for revealing the role of multiple-bond torsional mechanisms. The next natural step in this sequence of models of Rh and bR chromophores investigations is to perform a full nonadiabatic dynamics for RPSB.

Current Status of the Field

The large improvement of computational capabilities opened new possibilities in the field of studies of the excited state behavior of retinal and retinal model systems. Recent studies on the subject of dynamics are basically going in two directions. First, they combine more extended and complete models of retinal with use of higher theoretical levels. Second, they focus on studies of environmental influence on the electronic structure and excited state behavior. For the latter, two main features are of highest interest: the influence of charged groups on the spectral sensitivity of

retinal and the change in the mechanism leading to the conical intersection when the molecule is restricted in a confined and restricted space. The problem of including environmental effects on the static and dynamic properties of retinal and its models is mostly addressed by using hybrid QM/MM methods. Other studies, as discussed in the previous section, simulated the mechanical constraints for the molecules by means of explicitly constraining the movement of single atoms or groups.

As was previously mentioned, recent work by Schapiro et al. (2009), who employed the PSB4 model augmented by two methyl groups, reports bicycle pedal mechanism involving the equivalent of $C_{13}C_{14}$ bond as a second torsion. Although, these findings are in line with others based on PSB4 (Szymczak et al. 2009) and modified RPSB models (Ishida et al. 2009), simulations involving complete RPSB suggest that the C_9C_{10} should be the secondary bond performing concerted torsion with $C_{11}C_{12}$. Hayashi et al. (2009), have performed CASSCF QM/MM surface hopping dynamics for the complete retinal mechanically embedded in rhodopsin. Their analysis gave a picture in which isomerization around the $C_{11}C_{12}$ double bond is accompanied by coupled torsions around C_9C_{10} and C_6C_7 . It is concluded (Hayashi et al. 2009) that the excited state dynamics proceeds in such a way as to minimize interaction with the protein. On the other hand, scaled CASSCF QM/MM dynamics simulations of a single trajectory of retinal in bovine rhodopsin by Frutos et al. (2007) suggested the action of steric constraints for the torsion around C_9C_{10} . A quenching influence of protein was also proposed in the recent semiempirical QM/MM multiple spawning dynamics by Martinez and coworkers (Virshup et al. 2009). There, in contrast to RPSB simulations in vacuo, environmental effect was predicted to create highly directed dynamics leading to high yields of specific photoproducts. The same was concluded in simulations employing the ROKS method (Rohrig et al. 2004), which predict OBF as the only mechanism responsible for formation of all-*trans* RPSB.

As one can see in this brief review, despite important advances in the studies of retinal and its models based on dynamics simulation methods achieved in the last years, there is still an ongoing discussion about the mechanism of retinal isomerization in rhodopsin proteins, which calls for further investigations.

Heteroaromatic Rings and Nucleic Acid Bases

Photostability of DNA and RNA

The five natural nucleobases, adenine, guanine, cytosine, thymine, and uracil, absorb UV radiation in the range of 250–280 nm (UVC band) and at higher energies of about 200 nm (Abouaf et al. 2003; Clark and Tinoco 1965; Clark et al. 1965; Duggan et al. 1957; Middleton et al. 2009). This radiation, mainly the UVC, which is not blocked by the ozone in the air, causes DNA damage that, when not repaired by enzymatic mechanisms, may lead to mutagenesis and carcinogenesis (Cadet and Berger 1985).

The five nucleobases are non-fluorescent species (Longworth et al. 1966), which implies that the photo-energy is dissipated by means of ultrafast internal conversion. Time-resolved spectroscopy experiments have shown that either in gas phase or in solution the time constant for deactivation of the nucleobases excited at the 260 nm $\pi\pi^* \leftarrow S_0$ band is in the range of 0.5–6 ps (Canuel et al. 2005; Crespo-Hernández et al. 2004; Kang et al. 2002; Ullrich et al. 2004). This means that pathways to conical intersections where radiationless decay occurs are easily available from the Franck-Condon region.

The fact that nucleobases remain in reactive excited states only for a short time has been taken as an indication of enhanced photostability against UV irradiation (Middleton et al. 2009; Serrano-Andres and Merchan 2009). It may have happened that in the prebiotic era or in the early stages of evolution of the genetic code, nucleobases showing ultrafast decay had a competitive advantage against other chemical species more prone to undergo photochemical reactions. Indeed, one of the problems in understanding how life originated is to understand how chemical reactions not controlled by biogenetic mechanisms could lead to a few key substances instead of a multitude of chemicals, all of them with minor concentrations. It has been recently shown (Powner et al. 2009) that one possible route for prebiotic formation of nucleotides possesses as one of the steps an UV bath that could eliminate all but the nucleotides themselves, which reinforce the hypothesis that the intrinsic photostability of the nucleobases played an important role in the origin of life on Earth.

The mechanisms by which the nucleobases return to the ground state are subtly connected to the molecular structure of these bases. For instance, simple tautomerization of adenine into 2-aminopurine results in an increase of excited-state lifetime from 1.1 to 30 ps (Canuel et al. 2005). Methyl-substitution of cytosine or cytidine at C₅ position increases the lifetime in water solution by a factor of seven (Malone et al. 2003). Fluoro-substitution at the same site or acetyl-substitution at N₄ have even larger effects, increasing the lifetime from 1 ps to, respectively, 88 and 280 ps (Malone et al. 2003). Double methyl substitution in positions N₁ and N₃ of thymine increases its lifetime in gas phase from 5 ps (Canuel et al. 2005) to about 150 ps (He et al. 2003). Similar effects were also reported for single and double methyl substitutions in uracil (He et al. 2003).

Protonation and deprotonation of the bases depending on the environmental conditions has also influence on the excited-state lifetimes, although not so strong to make them fluorescent (Longworth et al. 1966). For cytosine, for example, the lifetime in water solution is raised from 1 to 13 ps when the pH is changed from 6.8 to 13 (Malone et al. 2003). For cytidine, the lifetime in water solution is reduced from 1 to 0.6 ps when the pH is changed from 6.8 to 0.08 (Malone et al. 2003). The excited-state lifetimes of nucleobases are dependent on the excitation wavenumber as well. For adenine, it increases from 1.1 ps to about 1.7 ps when the pump wavelength is changed from 276 to 267 nm (Chin et al. 2007). For 1-3-dimethyl thymine, it increases from about 60 to 200 ps when the pump wavelength is changed from 240 to 280 nm (He et al. 2003).

A great deal of theoretical work has been dedicated to nucleobases and their derivatives to identify conical intersections and reaction paths leading to them. In summary it has been found that these conical intersections can be formed by four different mechanisms:

1. Ring-puckering (Marian 2005; Perun et al. 2005a). In these conical intersections, the pyrimidine ring is deformed into conformations that produce twisted bonds (Barbatti et al. 2008). These conical intersections, which normally connect the $\pi\pi^*$ to the ground state, have the same origin as the twisted conical intersection in the methaniminium cation or the twisted pyramidalized conical intersection in ethylene (Barbatti et al. 2008). They are usually connected to the $\pi\pi^*$ state at the Franck-Condon region by barrierless pathways (Merchan et al. 2006). Out-of-plane amino and methyl deformations observed in conical intersections of adenine, cytosine, and thymine are caused by ring puckering as well.
2. Carbonyl out-of-plane deformation (Matsika 2004). In these conical intersections the ring remains planar, while the CO bond becomes almost perpendicular to the ring plane. The paths connecting the Franck-Condon region to them show an up-hill profile, which should decrease their efficiency for internal conversion (Zechmann and Barbatti 2008). This kind of conical intersection is analogous to formamide (Antol et al. 2007) and formaldehyde (Araujo et al. 2008) and connects the $n_O\pi^*$ state to the ground state.
3. Ring-opening (Perun et al. 2005b). These conical intersections are caused by stretching of CN ring bonds and connect $\pi\sigma_{CN}^*$ states to the ground state. Usually they appear in low-energy regions but are separated from the Franck-Condon by high barriers. They are analogous to the stretched-pyramidalized conical intersection in the methaniminium cation (Barbatti et al. 2008).
4. NH stretching (Perun et al. 2005b). These conical intersections connect the $\pi\sigma_{CH}^*$ and the ground states. They are accompanied by small barriers to the $\pi\pi^*$ state in the Franck-Condon region. Several of such conical intersections are not available in the nucleotides because of the hydrogen substitution by the sugar linkage.

Internal conversion may take place not only at the conical intersections of isolated nucleobases but also at intersections formed along the hydrogen bonds connecting base pairs. It has been shown that a charge-transfer state crosses the ground state along the inter-pair hydrogen transfer (Schultz et al. 2004). When this mechanism is compared between guanine-cytosine Watson-Crick (WC) pair and other non-WC pairs formed by the same bases, it turns out that the WC pair possesses the lowest energy crossings (Sobolewski et al. 2005). This has been taken as an indication that inter-pairs internal-conversion may have constituted an evolutionary advantage for WC conformations. This hypothesis, however, has been disputed based on experiments showing that single and double stranded DNA decay follow a similar kinetics (Crespo-Hernandez et al. 2005). According to these experiments, initially excited state decays to an exciplex state in less than 1 ps.

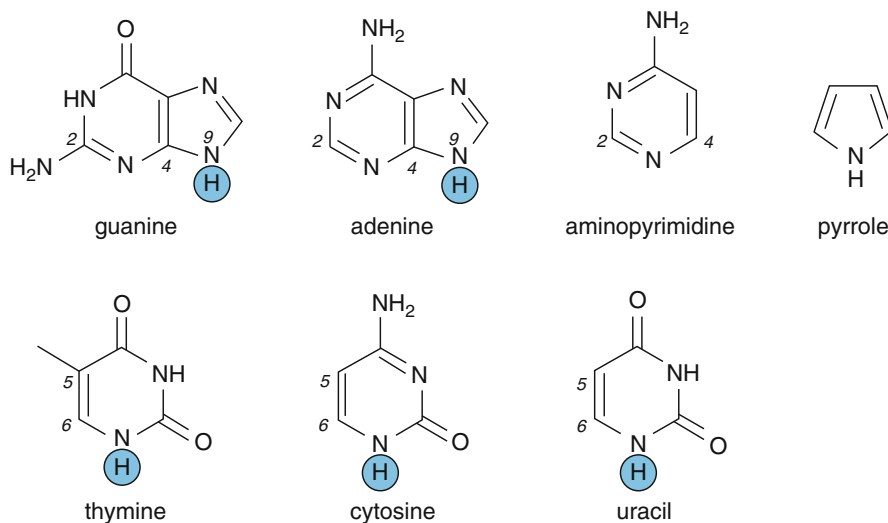


Fig. 4 The five nucleobases and the heterocycle models aminopyrimidine and pyrrole. The connections to the sugar in the nucleotides are indicated by the detached hydrogen atom

The exciplex state involves stacked nucleobases and lives for 10–100 ps before returning to the ground state via charge recombination. The stacking character of the exciplex state limits the damage to one single strand. The undamaged strand may then be used as a template for the repair mechanisms (Crespo-Hernandez et al. 2005; Middleton et al. 2009).

In the next sections we will focus on recent investigations of internal conversion in nucleobases and nucleobase models using dynamics simulations. First, we will explore aminopyrimidine (Fig. 4), a model system that can help to understand how pyrimidine rings deactivate after UV excitation. Then, we will focus on pyrrole, to see how five-membered rings deactivate. Finally, in section “Current Status of the Field,” we will discuss dynamics simulations of the nucleobases themselves.

Model 1: Aminopyrimidine

Pyrimidine rings are structural subunits of all five nucleobases. In fact, the strong UV absorption in the 260 nm band is due to the $\pi\pi^*$ transition in the pyrimidine ring of both pyrimidine and purine nucleobases. For this reason, it is especially relevant to understand how internal conversion takes place in these systems.

Conical intersections in aminopyrimidine are very similar to those in the pyrimidine ring of adenine. The same is true for reaction paths connecting the Franck-Condon region to these conical intersections (Barbatti and Lischka 2008; Barbatti et al. 2008). As can be expected from simple inspection, main differences arise in the two carbon sites where the imidazole group connects with pyrimidine to form the purine structure (Fig. 4). While in aminopyrimidine a low energy conical intersection is formed by puckering of the C_4 atom, this kind of deformation is hindered in the C_4 site of adenine. (Note that the numbering for aminopyrimidine

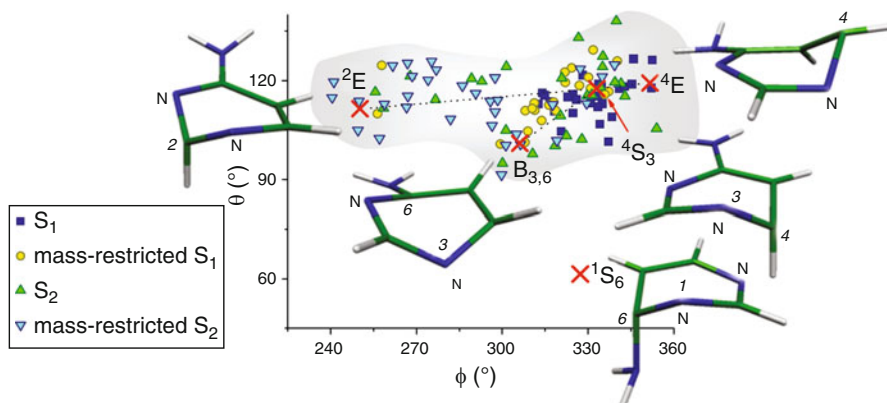


Fig. 5 $S_1 \rightarrow S_0$ hopping position in the $\theta - \phi$ Cremer–Pople space for aminopyrimidine (Data from Barbatti and Lischka 2007; Barbatti et al. 2008)

indicated in Fig. 4 is not the conventional one, but that it is chosen to make the comparison to adenine straightforward.)

A series of investigations have focused on dynamics simulations for aminopyrimidine (Barbatti and Lischka 2007; Barbatti et al. 2008). As we discuss in this section, these results show that in spite of the strong similarities between the electronic excited states of adenine and aminopyrimidine, the restrictions imposed by the imidazole group make the dynamics of the two species essentially distinct.

Surface-hopping dynamics simulations was performed at SA-3-CASSCF(8,7)/6-31G* level (Barbatti and Lischka 2007; Barbatti et al. 2008). First, 30 trajectories were started in the S_1 state. The molecule returned to the ground state in 0.4 ps. A second batch of 30 trajectories was started in the S_2 state. In this case, the $S_2 \rightarrow S_1$ deactivation occurred in only 9 fs. The return to the ground state took place again in about 0.4 ps. These time constants are too fast compared to adenine, which has an excited-state lifetime of 1 ps (Canuel et al. 2005). In order to understand the reason for the difference, we can look at the nuclear conformation at the $S_1 \rightarrow S_0$ hopping time. These conformations are show for all trajectories in Fig. 5 projected on the $\theta - \phi$ Cremer–Pople space (Boeyens 1978; Cremer and Pople 1975). Each point in this space represents a different kind of puckering of a six-membered ring, constituting a very convenient way to describe puckered conformations in heterocycles. As we can see in the figure, both sets of trajectories (starting in S_1 and starting in S_2 states) return to the ground state preferentially at conical intersections showing ring deformation at N_3 and C_4 sites, mainly at the screw-boat conformation 4S_3 . The hopping distribution is more broadly spread when the dynamics starts in S_2 , which can be explained by the energy excess. In this case, we can observe trajectories returning to ground state at the boat conformation $B_{3,6}$ and even a few of them at the envelope 2E conformation.

Since puckering at the C_4 site is hindered in adenine by the imidazole group, the concentration of hopping events involving deformation of this site in

aminopyrimidine explains the difference between the lifetimes of both species. Apparently, aminopyrimidine is decaying at conical intersections not available in adenine. This hypothesis can be tested by restricting puckering deformation at sites C_4 and C_5 of aminopyrimidine in a similar way as they are restricted in adenine (Barbatti and Lischka 2007). To do so, the mass of the hydrogen atoms at these sites is increased to 45 amu. as to emulate the moment of inertia imposed by an imidazole group. Two new sets of 30 trajectories were run. One set starting in the S_1 state and another set starting in the S_2 state. Results of this second set are shown for the first time in this work.

Mass-restricted trajectories starting in S_1 returned to ground state in 0.96 ± 02 ps (Barbatti et al. 2008). In spite of the good agreement with the 1 ps measured for adenine, the hopping conformations are still concentrated in N_3 and C_4 deformations (Fig. 5). This occurs because the mass restrictions at C_4 -H and C_5 -H are decoupled. This is different in adenine, where the ring imposes a strong correlation in the motion of these sites. The long lifetime of mass-restricted aminopyrimidine in comparison to the non-restricted case is due to the longer period that the molecule needs to tune the conical intersection coordinates. Figure 5 also shows that the hopping events of mass-restricted aminopyrimidine starting in the S_1 state are displaced in the direction of $B_{3,6}$ conformations.

Curiously, when mass-restricted dynamics is started in the S_2 state, the lifetime is increased even more, to 1.4 ps (Fig. 6). Hopping events occur along the whole region connecting the envelope 4E conformations to envelope 2E conformations indicating that a branch of crossing seam exists in this region. Different from the other three sets of dynamics simulations, a substantial amount of hopping events occurs at the 2E conformation, which, as we shall review in section “Current Status of the Field,” is the main conformation accessed for adenine deactivation (Barbatti and Lischka 2008). Note that two conditions were necessary to trigger internal conversion in this region of the crossing seam, first the restriction of the C_5 and C_4 sites and second the excess of energy of the S_2 initial state.

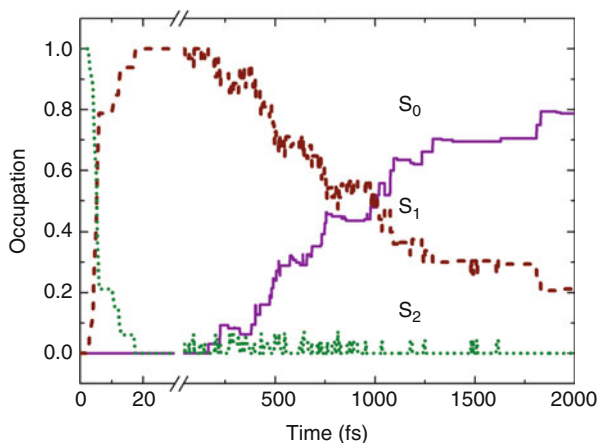


Fig. 6 State occupation of mass-restricted aminopyrimidine starting in the S_2 state

Hopping events at 6S_1 conical intersections, which are energetically accessible from the Franck-Condon region, were not observed in any of the four sets of trajectories (Fig. 5). Characteristically, 6S_1 conformations imply a strong out-of-plane deformation of the amino group. A plausible explanation for the lack of internal conversion in this region of the crossing seam is that trajectories are trapped in other regions of the S_1 surface before the molecule has time to undergo this strong deformation. Results for aminopyrimidine starting in S_2 state without mass-restriction, but increasing the active space from (8,7) to (10,8) by addition of one more n orbital to the active space show that a fraction of trajectories (about 15 %) can access this conical intersection (Szymczak et al. 2011).

The effect of stacking interactions in DNA conformation has also been investigated in the dynamics of aminopyrimidine interacting with guanine nucleobases (Nachtigallova et al. 2010). The results show that the ring puckering process observed in gas phase is not restricted by the neighbor molecules.

Model 2: Pyrrole

Pyrrole cannot be directly regarded as a model for nucleobases. However, the understanding of its UV-excited dynamics may bring insights of the role of the imidazole ring in the pyrimidine bases adenine and guanine. Different from the pyrimidine subunit, which is responsible for the UV absorption in the near UV region, both pyrrole and imidazole absorb at much shorter wavelengths of about 200 nm (Daura et al. 1999; Devine et al. 2006). Because of this, it may be expected that purine bases absorbing in the far UV region may present an excited state dynamics similar to that of five-membered rings.

Pyrrole is a simple five-membered heterocyclic compound which undergoes a rapid non-adiabatic deactivation (Lippert et al. 2004) with very low quantum yields (Shin 2004). It has been proposed (Sobolewski and Domcke 2000) that deactivation occurs via a NH-stretching mechanism along a ${}^1\pi\sigma^*$ excited state surface. Although it is certain that fast H fragments are formed through this mechanism, it cannot explain experimental findings of slow H atoms and other fragments such as HCN and C_3H_3 (Blank et al. 1994; van den Brom et al. 2005; Wei et al. 2004). This means that another mechanism should occur in competition with the NH stretching, which was proposed to involve ring-puckering and ring-opening of the heterocyclic ring (Barbatti et al. 2006; Poterya et al. 2007).

Non-adiabatic dynamics simulations for the pyrrole radical cation and other five-membered rings were performed in reduced dimensionality by means of wave packet propagation using the multiconfigurational time-dependent Hartree approach (MCTDH) together with the linear vibronic model, allowing a good reproduction of spectral features (Köppel et al. 2004; Venkatesan et al. 2007). Wave packet dynamics simulations have also been performed focusing on the NH-stretching reaction coordinate (Lan and Domcke 2008; Vallet et al. 2004, 2005) revealing details of the reaction along this coordinate, especially of the dependence of the dissociation branching ratios and time scales on the initially excited modes. These simulations in reduced dimensionality predict a quantum yield for NH dissociation near to unity. Dynamics simulations of pyrrole in its full dimensionality have been performed

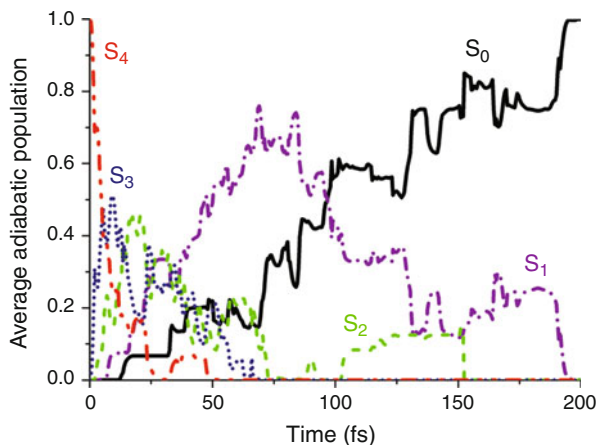


Fig. 7 Average adiabatic populations of trajectories for five states as a function of time after photoexcitation of pyrrole into the S_4 state (Data from Vazdar et al. 2009)

using surface hopping approach (Vazdar et al. 2009). And, as we are going to discuss in more detail next, it revealed that other deactivation channels beyond the NH-stretching are activated. These channels involve strong ring deformation. The adiabatic dynamics in the vibrationally excited ground state, also discussed below (Sellner et al. 2009), showed that these ring deformation channels result in photochemical products consistent with the heavy fragments experimentally observed.

Excited-state non-adiabatic dynamics simulations were performed for 90 trajectories in a total simulation time of 200 fs with a time step of 0.5 fs (Vazdar et al. 2009). The dynamics simulations have been calculated at the MR-CISD/SA-5-CASSCF(4,5) level of theory including five electronic states (ground state, two $\pi\pi^*$ states and two Rydberg π -3s (σ^*) states).

After excitation to the S_4 state, pyrrole depopulates via a series of conical intersections between valence and Rydberg states in a very short time. The average adiabatic population of ground and excited states as a function of time is presented in Fig. 7.

The analysis shows that the valence S_4 state depopulates very quickly. After around 10 fs, the average population of S_4 and S_3 states is identical. After 50 fs, the S_4 state is completely depopulated, while S_3 and S_2 state depopulate after ca 75 fs. The S_1 state starts to depopulate after ca 100 fs. After 100 fs, the dynamics is basically a two-state dynamics between ground and first excited state. The complete population transfer to the ground state occurs within 200 fs.

The statistical analysis presented in Fig. 8 reveals that 80 % of trajectories follow the NH stretching mechanism for deactivation into the ground state. The remaining 13 % of trajectories undergo a ring deformation mechanism (10 % is ring puckering and 3 % is ring opening). Finally, 7 % of trajectories do not deactivate during 200 fs

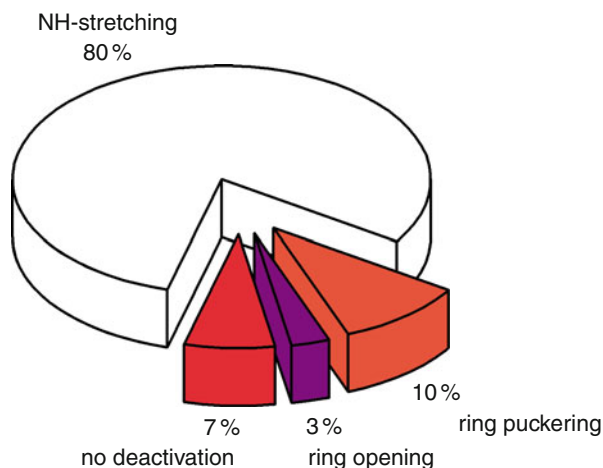


Fig. 8 Statistical analysis of different trajectory types in excited-state dynamics of pyrrole (Data from Vazdar et al. 2009)

of the dynamics simulation. Dynamics performed with TDDFT method showed a very similar distribution of mechanisms (Barbatti et al. 2010b).

The observation of the ring deformation deactivation mechanism for a substantial number of trajectories confirms that it can be the source of experimentally observed heavy fragments. To verify this hypothesis, dynamics in the ground state should be performed. This is a specially challenging task because after deactivating, the photoenergy is converted into vibrational energy. In this very hot vibrational ground state, pyrrole can reach strongly distorted geometries and dissociate through several channels. Most of conventional quantum chemical methods are not tailored to work under such extreme conditions.

This problem has been addressed by starting the hot ground state dynamics at a ring-puckered ring-opened conical intersection (see Fig. 9a) using a wavefunction formed by a combination of the generalized valence bond method in the perfect pairing approximation (GVB-PP) and complete active space (CAS) (Sellner et al. 2009). The GVB-PP-CAS(6,5)/6-31G* level applied to the gradient calculation for the dynamics has proven not only to be more reliable concerning photoproducts but also much more stable during dynamics compared to conventional SA-2-CAS(6,6) calculations. The results reported here have been calculated with this GVB-PP-CAS(6,5) space. A hundred trajectories have been propagated for a maximal time of 1 ps using a step size of 0.5 fs.

As the dynamics is started on a point in the crossing seam, special care has to be taken to generate initial velocities. Two approaches have been employed, sampling initial velocities only in the branching space and sampling random velocities in the complete space using weighted Gaussian random numbers, both showing good correlation. The results discussed here are based on the second approach. Translational and rotational contributions of these velocities were removed, which

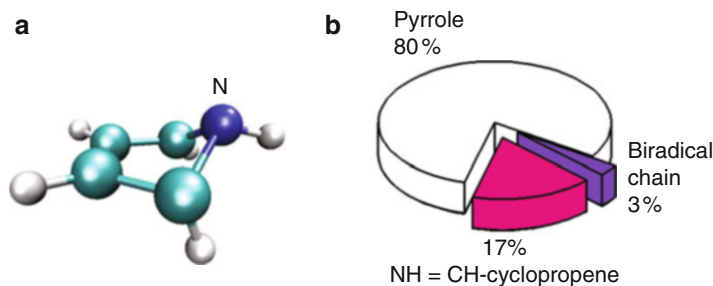


Fig. 9 (a) Ring-puckered ring-opened S_0/S_1 conical intersection (initial geometry of adiabatic ground state dynamics for pyrrole); and (b) quantum yield within 1 ps of simulation (Data from Sellner et al. 2009)

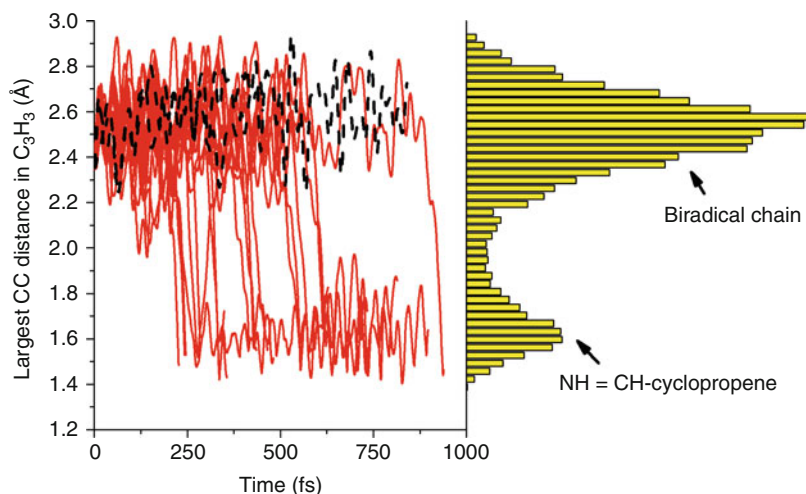


Fig. 10 Largest CC distance in C_3H_3 as a function of time showing the dynamics of the formation of $NH=CH$ -cyclopropene out of the biradical chain (Data from Sellner et al. 2009)

have been rescaled to an energy of 3.01 eV. This energy is the sum of half of the ZPE (2.24 eV) and the energy difference between the Frank Condon point and the MXS which is assumed to be converted from potential to kinetic energy.

The results show that besides pyrrole itself, two main photoproducts were obtained within the simulation time (see Fig. 9b). These are $NH=CH-C_3H_3$ and a biradical chain ($NH=CH-CH-CH=CH$). A detailed analysis of the trajectories ending at these photoproducts shows that the biradical chain is mainly a precursor to the substituted cyclopropene fragment. The cyclization starts after 200 fs and the new ring formation occurs in an average time of only 17 fs (see Fig. 10). Regarding the consequences on the experimentally observed (cyclo)propene fragment, this shows that the ring formation can not only happen before the dissociation into the HCN and C_3H_3 fragments takes place, but seems to be the dominating process.

Dissociation into H_xCN ($x = 1, 2$) and C_3H_3 did not take place within the 1 ps simulation time and it should occur on a longer timescale. The systematic investigation of the average bond distances during the dynamics (Sellner et al. 2009), clearly reveals that cleavage of those bonds yielding the experimentally observed fragments are more likely to occur than the cleavage of bonds leading to non-observed fragments.

Summarizing, non-adiabatic excited state dynamics of pyrrole showed two main deactivation mechanisms, the first one being the NH stretching explaining the (fast) hydrogen atoms and the second one the ring deformation (including opening and puckering). Dynamics from the puckered conical intersection showed that the second pathway is responsible for $NH=CH-C_3H_3$ and the biradical chain which can be seen as precursors to the experimentally observed heavier fragments regarding their electronic structure and bonding situation.

Current Status of the Field

A relatively large number of investigations reporting dynamics simulations for nucleobases, nucleobase tautomers, nucleotides, derivatives, WC pairs and quartets in different solvation conditions has been published in the last years. Recent overviews of the field can be found in Alexandrova et al. (2010) and Barbatti et al. (2010a). In this section, we shortly review the dynamics of isolated bases focusing on which relaxation mechanism these works have proposed for each molecular system.

Adenine: In spite of a large number of different conical intersections that has been described for this nucleobase (see Barbatti and Lischka 2008 and references therein), theoretical analyses have converged to two main pathways, one involving puckering of the C_2 atom (envelope E_2) and another one involving puckering of the C_1 and C_6 atoms along with out-of-plane displacement of the amino group (screwboat 1S_6). Surface hopping dynamics simulations based on MR-CIS potential energy surfaces showed that adenine is trapped in envelope E_2 geometries much before it can reach other reaction pathways (Barbatti and Lischka 2008). Nevertheless, a different result has been achieved by surface hopping simulations based on semiempirical methods, with most of trajectories following the 1S_6 pathway (Alexandrova et al. 2010; Fabiano and Thiel 2008). Dynamics simulations based on TD-DFTB surfaces indicated a intermediary situation with selective activation of the 1S_6 and E_2 pathways depending on the excitation energy (Lei et al. 2008) and very long lifetime (Mitric et al. 2009b).

Guanine: A series of investigations on guanine tautomers and derivatives in gas phase and water employing surface hopping based on ROKS potential energy surfaces have shown that out-of-plane distortions increase the non-adiabatic transition probability and induce the transition to the ground state in about 1 ps (see Santoro et al. 2007a and references therein). Surface hopping dynamics based on OM2 surfaces predicted a major role for processes involving C_2 puckering, with a lifetime of about 0.2 fs. Similar process has been observed at the ab initio level (Barbatti et al. 2011b). In the particular case of the enol form of guanine, simulations predicted that OH dissociation channel is a major deactivation pathway (Langer et al. 2005).

Cytosine: Multiple spawning dynamics simulations based on CASSCF(2,2) surfaces predicted that deactivation of this nucleobase should occur in approximately 0.5 ps mainly through conical intersections involving out-of-plane distortions of the amino group (65 %) and with minor contribution of the $n_O\pi^*$ (15 %) and of the ethylene (5 %) pathways (Hudock and Martinez 2008). OM2-based surface hopping dynamics, on the other hand, predicts that the ethylene pathway should dominate the deactivation, which occurs in about 0.4 ps (Lan et al. 2009). Dynamics for this molecule computed with surface hopping at CASSCF surfaces predict dominance of the $n_O\pi^*$ pathway (Barbatti et al. 2011a; González-Vázquez and González 2010).

Thymine: This nucleobase has the longest lifetime among the five naturally occurring bases (Canuel et al. 2005; Kang et al. 2002). Multiple spawning dynamics simulations based on CASSCF surfaces have revealed that the source of this feature may be a trapping of thymine in a minimum of the S_2 state for a few picoseconds (Hudock et al. 2007). Surface hopping dynamics based on the OM2 method, however, predict a very short stay in the S_2 state (0.02 ps) and an overall excited-state lifetime of 0.4 ps (Lan et al. 2009). In this latter case, the decay to S_0 occurred mainly through the out-of-plane O conical intersection (82 %) and pathways involving puckering at the C_5 atom were observed in the other 12 % of trajectories. Surface hopping dynamics based on CASSCF surfaces agree with the multiple spawning results and predicts that the $n\pi^* S_1$ state will be populated after 2.6 ps (Gustavsson et al. 2006). Surface hopping simulations also based on CASSCF surfaces, but starting at the transition state in the S_2 state, predict that both the $\pi\pi^* S_1$ state (65 %) and the $n\pi^* S_1$ state should be populated (Asturiol et al. 2009). While according to Gustavsson et al. (2006) trajectories should remain in the $n\pi^* S_1$ state for additional few picoseconds, results of Asturiol et al. (2009) indicated that, independently of the diabatic character, the S_1 state should be depopulated in about 0.1 ps through conical intersections involving puckering at the C_5 atom.

Uracil: Similarly to thymine, multiple spawning CASSCF dynamics predict a long stay of uracil in the S_2 state (Hudock et al. 2007). Surface hopping OM2 dynamics, in contrast, predict a very fast decay to S_1 and an overall lifetime of about 0.6 ps with trajectories following out-of-plane O pathways in 85 % of the cases and puckering at C_5 atom in the remaining 15 % of the cases (Lan et al. 2009). Wave packet propagation in a 3-D model surface also predicts a very fast $\pi\pi^* \rightarrow n\pi^*$ transition taking place in less than 0.1 ps (Improta et al. 2009). Surface hopping dynamics simulations on ROKS surfaces in gas phase and in water favor deactivation through the C_5 puckering pathway (Nieber and Doltsinis 2008).

As we have seen, dynamics simulations carried out with different methods have often presented conflicting results. This may be mainly attributed to differences between quantum-chemical electronic structure methods used to compute the potential energy surface, rather than to the different non-adiabatic dynamics methods. More about the limitation of the methods will be discussed in the conclusions for this chapter. For now, we may note that even though dynamics simulations have not achieved a degree of agreement that would allow assigning specific relaxation

mechanisms for each nucleobase, they have been able to reduce the possibilities to a few mechanisms in each case.

Formamide and Peptide Bonds

Probing Photoexcitation of Proteins

The peptide linkage plays an important role in many biological processes and the knowledge of how it interacts with ultraviolet light is of fundamental importance for understanding biological activities in living organism. This holds in particular for interaction of peptides with 193 nm excimer ArF laser due to its proximity to the absorption maximum of peptide bonds (Gingell et al. 1997). This is an effect of considerable importance because ArF lasers are widely used clinically for laser surgery (Mulcahy et al. 2000). It is shown, for instance, that the peptide bond is the main target of irradiation in collagen and its scission is claimed to be the most important photochemical reaction taking place during photorefractive keratectomy (Torikai and Shibata 1999). Furthermore, the peptide group appears as the repeat unit in some industrially important polymeric materials (Greenberg et al. 2002) and knowledge about the mechanism of its interaction with ultraviolet light has implications for understanding crucial factors influencing their photostability. The smallest molecule of the peptide prototype is formamide (Fig. 11) and as such it has been frequently used as the simplest model for studying fundamental aspects of photochemical and photophysical processes taking place in peptides and proteins (Antol et al. 2007, 2008a, 2008a, b; Back and Boden 1971; Kang and Kim 2006; Liu et al. 2000; Lundell et al. 1998). An overview of the recent studies related to the second topic will be presented in section “Model I: Formamide” of this chapter.

Photoinduced fragmentation of charged small polypeptides has also been studied extensively in recent years due to their importance in “soft” ionization methods such as MALDI and ESI (Cui et al. 2005; Grégoire et al. 2007; Jeong et al. 2005, 2007; Kang et al. 2004; Nolting et al. 2006; Thompson et al. 2004, 2007). Photo-dissociation dynamics of small protonated peptides containing aromatic amino acids

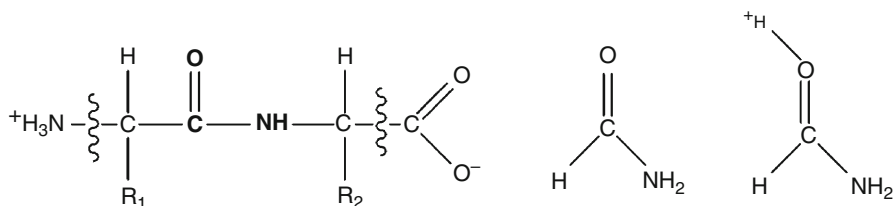


Fig. 11 Schematic presentation of a polypeptide chain with the peptide bond showed in bold (*left*), of formamide (*center*) and of oxygen-protonated formamide (*right*). R_1 and R_2 stand for the amino acid chains

excited by 266 nm laser beams have been also explored by pump/probe femtosecond measurements (Grégoire et al. 2006, 2007; Nolting et al. 2006). However, the deactivation mechanism of possible pathways could not be derived, neither experimentally nor theoretically due to the size and complexity of the explored systems. Concomitantly, several experimental and computational photoexcitation studies have been undertaken on gas-phase protonated amino acids (Cui et al. 2005; Thompson et al. 2007). This led to the proposition of a deactivation mechanism (Kang et al. 2005, 2004) compatible with experimental results obtained for small polypeptides containing aromatic amino acids (Grégoire et al. 2006, 2007; Nolting et al. 2006). A recent review on this topic is given in Grégoire et al. (2009). With the aim of exploring possible alternative pathways of photodissociation of polypeptides we recently examined impact of protonation on fragmentation of the peptide bond using protonated formamide as the model molecule (Antol et al. 2008b). The most salient results from this work will be described in section “Model 2: Protonated Formamide.”

Model 1: Formamide

Photodissociation of formamide has been studied in the gas phase (Back and Boden 1971), solution (Petersen et al. 2008) and in noble-gas matrices of argon and xenon (Lundell et al. 1998). In gas phase photolysis, CH, H₂, and NH₃ fragments with quantum yields of 0.8, 0.6, and 0.2, respectively, were detected (Boden and Back 1970). Their formation was rationalized by three major primary processes forming the products NH₂ + CO + H, HNCHO + H, and NH₃ + CO. Subsequently, Lundell et al. (1998) reported that photolysis of formamide in the argon matrix by 193 nm laser leads to formation of a weakly bound complex between NH₃ and CO. Its formation was ascribed to the breaking of the C–N bond in formamide followed by hydrogen atom transfer from the formyl radical HCO to the NH₂ radical. On the other hand, when the formamide was trapped into solid Xe matrix major photoproducts were H₂ and HNCO. The change in composition of the products was explained by intersystem crossing to the triplet state induced by a strong external heavy-atom effect in the Xe matrix.

The mechanism of photodissociation in the gas phase has been also studied computationally at various levels of theory (Back and Boden 1971; Liu et al. 2000; Lundell et al. 1998). For instance, Liu and co-workers (2000) considered the mechanism of formamide photodissociation in the S₁ and the T₁ excited states using the CASPT2/CASSCF method. They located transition structures for each of the paths presented by Eqs. 1, 2, 3, 4, and 5 and found that the transition structure for C–N dissociation (Eq. 1) has the lowest energy on the S₁ as well on the T₁ potential energy surfaces, indicating that this reaction path is the most probable one.





However, the full understanding of the photodeactivation process in addition to characterization of the energy minima of the reactants and products and the transition state structures, requires also knowledge about minima on the crossing seams between involved excited states and the ground state, as well as information about nuclear motion of atomic nuclei during the process. Therefore, we recently reexamined the reaction paths considered by Liu and coworkers (2000) for the S_1 state taking into account both of these requirements and at higher level of theory. In addition, photodissociation of formamide in the second excited singlet state was considered also. The details of computational methods are given in Antol et al. (2007). Following location of all critical points on the S_1 and S_2 surfaces, we turned to dynamic simulations, which were carried by the direct trajectory surface hopping method based on the state average SA-3-CASSCF(10,8)/6-31G(d) calculations. For each of the initial states, 100 trajectories were propagated for 1 ps. The average occupations of the S_1 and S_2 states are shown in Fig. 12. It appears that the average occupation of the S_1 state remains practically unchanged for about 43 fs. After this initial stage, its occupation starts to decay exponentially. The lifetime of the state was calculated to be 441 ± 4 fs. On the other hand depopulation of the S_2 state starts after only 12 fs and the overall estimated lifetime is 66 ± 1 fs. After passing through the S_2/S_1 conical intersection, formamide resides on the S_1 surface only shortly and decays to the ground state.

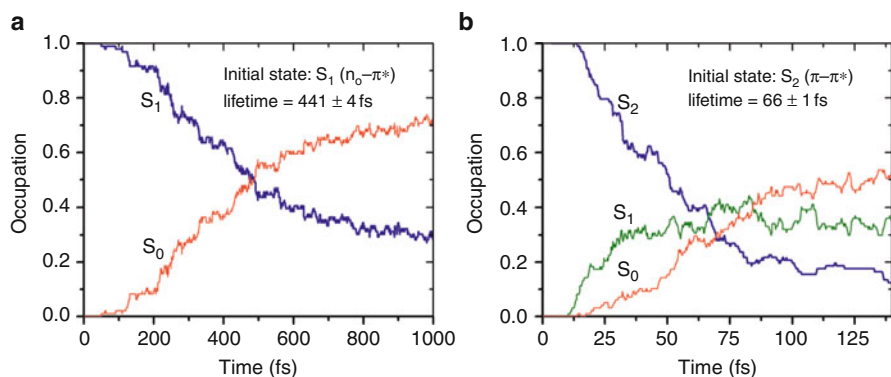


Fig. 12 Fraction of trajectories for each state as a function of time after photoexcitation into (a) S_1 initial state; and (b) S_2 initial state (Data from Antol et al. 2007)

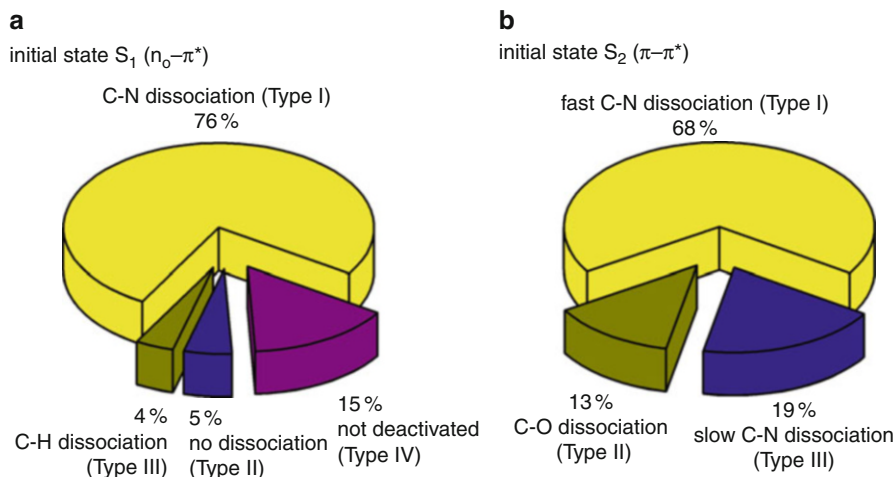


Fig. 13 Photodeactivation paths observed during dynamics simulations of formamide starting at the (a) S_1 ($n_O\pi^*$) and at the; (b) S_2 ($\pi\pi^*$) states (Data from Antol et al. 2007)

Analysis of the trajectories (Fig. 13) revealed that after excitation into the S_1 state about 75 % of the examined trajectories followed the C–N dissociation channel. After reaching the ground state, the energy of the system is transferred to translational motion of the HCO and the NH_2 radicals and to C–H dissociation from vibrationally hot CHO fragment, which finally results in formation of NH_3 and CO.

In addition to this channel, three additional types of trajectories were observed, but their fraction was found to be much lower. In trajectories of type II, the molecules decay to the ground state either via the same C–N conical intersection as in the previous case, followed by back formation of formamide on the S_0 PES or return to the ground state via a nondissociative MXS (Antol et al. 2007). In either case, no dissociation is observed within the simulation time. In trajectories of type III, hoppings to the ground state cause activation of the C–H vibrations and the molecule undergoes C–H bond stretching. Full C–H dissociation, however, was observed in only one trajectory. Finally, in 15 % trajectories, formamide did not decay to the ground state and did not dissociate within the simulation time.

Following excitation to the second excited state, three types of trajectories were observed. The dominant type, corresponding to 68 % of the trajectories, is characterized by a very fast C–N bond cleavage starting immediately after photoexcitation and still taking place on the S_2 surface. The deactivation to the S_1 occurs in an average time of 31 fs. The system remains on the S_1 surface only for a short time and then switches to the S_0 state in an average time of 65 fs. Such fast $S_1 \rightarrow S_0$ decay is opposite to the slow $S_1 \rightarrow S_0$ decay observed when the initial state was S_1 . This is ascribed to the cooperative action of the energy transferred from the S_2 state, and the momentum acquired by the system along the C–N stretching coordinate during its motion from the Franck-Condon region to the S_2/S_1 conical

intersection. A similar process has been observed in the dynamics of protonated formamide (Antol et al. 2008b) and the CH_2NH_2^+ cation (Barbatti et al. 2007a).

In addition to the C–N dissociation channel, two alternative types of trajectories, both of them involving the C=O stretching mode, were observed, but their fraction was found to be much lower. The first among them (trajectories of type II) is characterized by dissociation of the oxygen atom and concomitant merging of all three potential energy surfaces. It corresponds to 13 % of all observed events. Finally, trajectories of type III (19 %) are best described as combination of the trajectories of types I and II. In this case, initial motion involves the C–O stretching mode, but the C–O bond cleavage does not occur due to recombination of the molecule in the S_2 followed by C–N bond dissociation.

Model 2: Protonated Formamide

The most basic site of formamide in the ground, as well as in the excited state is the oxygen atom (Antol et al. 2008a). Therefore, in this section we shall restrict our discussion to the photofragmentation of the O-protonated species (Fig. 11). The calculations were carried out by means of the same computational approach used for the parent molecule, the SA-3-CASSCF(10,8)/6-31G(d) method. A summary of dynamics calculations results are presented in Fig. 14. More detailed discussion, including dynamics simulations of photodissociation of N-protonated formamide, can be found in Antol et al. (2008b).

Examination of the results shows that dissociation of the O-protonated formamide from the S_1 state strongly resembles the mechanism of dissociation observed from the second excited state in the neutral molecule, as expected, due to the fact that both states are of $(\pi\pi^*)$ character. This is also reflected in considerable shortening of the calculated lifetime relative to that calculated for the S_1 state in the neutral molecule (33 ± 1 fs vs. 441 ± 4 fs).

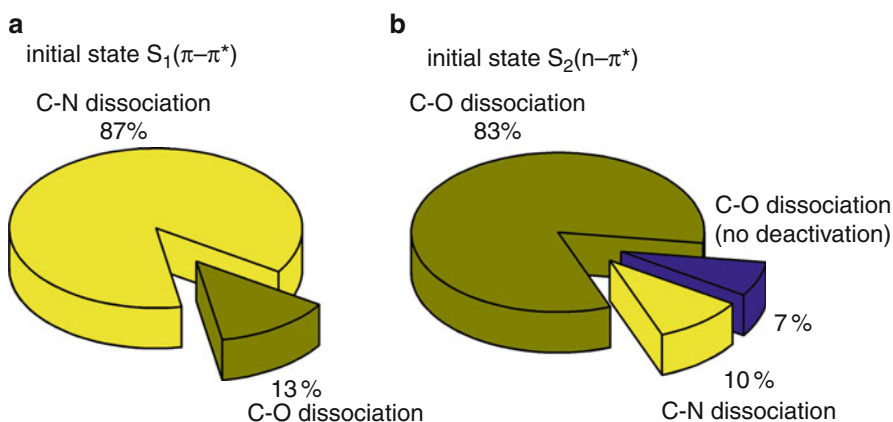


Fig. 14 Photodeactivation paths observed during dynamics simulations of O-protonated formamide starting at the (a) S_1 ($\pi\pi^*$) and at the; (b) S_2 ($n\pi^*$) states (Data from Antol et al. 2008b)

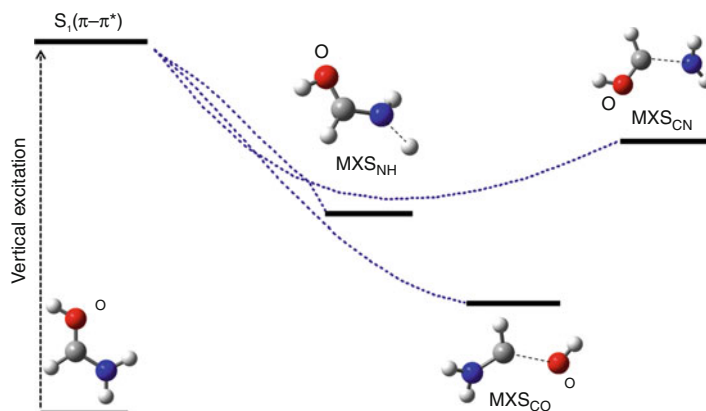


Fig. 15 Schematic presentation of reaction paths in photodissociation of oxygen protonated formamide. MXS_{NH} , MXS_{CO} , and MXS_{CN} denote minima on the crossing seam for dissociation of N-H, C=O, and C-N bond, respectively. Geometries are optimized at the CASSCF(10,8)/6-31G(d) level (Data from Antol et al. 2008b)

In searching for possible deactivation paths, three minima on the crossing seam between S_1 and S_0 were found to be accessible without energy barrier indicating very fast dynamics. The other interesting feature is that, in contrast to the neutral molecule, the MXS for the C-N dissociation has the highest energy (Fig. 15). In spite of that, analysis of the calculated trajectories revealed that this is the major dissociation channel. The reason for that lies in the fact that this reaction path has the steepest “slope” in the Franck-Condon region (Fig. 15). In other words, as the molecule in the Franck-Condon region does not have information about energies of conical intersections, it selects the energetically most favorable reaction path and not one leading to the most stable minimum on the crossing seam. The predominance of non-adiabatic transitions near highest energy MXS has also been observed in photodissociation of aminopyrimidine (Barbatti and Lischka 2007). In protonated formamide this deactivation path is activated in 87 % of trajectories. After reaching the ground state, the energy of the system is transferred to translational motion of the HCOH and NH_2 fragments in 33 % of trajectories, while in 47 % of trajectories HCOH and NH_2 recombine into the vibrationally excited starting structure. Furthermore, in 7 % of trajectories the proton transfer from the HCOH radical cation to the NH_2 radical occurs resulting in formation of the NH_3 radical cation and HCO radical.

In addition to the C-N dissociation path, in 13 % of trajectories the dissociation channel involving stretching of the C-O bond becomes activated. This dissociation path strongly resembles the C-O dissociation channel from the second excited state of the neutral molecule (trajectories of type II).

Dissociation of the O-protonated formamide following excitation to the second excited state is also interesting. It triggers a very fast switch (10 fs) from the S_2 to the S_1 state via a highly pyramidalized non-dissociated MXS (Antol et al. 2008b), followed by activation of two dissociation processes (C-O and C-N) on the S_1

surface, with the C–O dissociation path being observed in 83 % of trajectories. It is interesting to note that the resulting branching ratio between C–O and C–N dissociation pathway differs considerably from that observed upon photodissociation from the $n\pi^*$ state in the neutral molecules in which the C–N dissociation was the major photodissociation channel. The major reason for this behavior could be connected to the direction of the momentum along the C–O stretching coordinate inherited from the S_2/S_1 conical intersection in the O-protonated formamide. Another point worth of noting concerns formation of the OH radical, indicating that protonated formamide might be a precursor in the photogeneration of potentially deleterious OH radical.

Current Status of the Field

One point of particular relevance in the photodynamics studies of formamide and peptides that has not been considered in this section is the effect of the environment. Environmental effects are expected to change the relative position of the electronic states, the position, and the topology of the conical intersections (Burghardt and Hynes 2006). QM/MM dynamics simulations of photodissociation of formamide embedded in argon matrix showed strong impact on the outcome of the process. This was illustrated by formation of the 1:1 complex between ammonia and CO and prevention of full separation of the NH_2^+ and HCO^+ subunits in the $\text{NH}_2^+ + \text{HCO}^+$ radical pair. Also, the argon matrix influenced the lifetime of the S_1 state, which increased by 211 fs relative to the gas phase (Eckert-Maksić et al. 2010b). Another important issue that deserves attention concerns the effect of substituents on the mechanism of photodissociation of formamide. Some experimental, as well as computational studies, on the effect of methyl (Chen et al. 2003; Eckert-Maksić and Antol 2009; Liu 2002) and benzyl (Chen and Fang 2004; Pei et al. 2008) substituents related to this issue have been recently reviewed (Eckert-Maksić et al. 2010a). Finally, the possibility of intersystem crossing to the first triplet state has not been fully explored. This holds, in particular, for photodissociation of formamide in matrices, but also for photodissociation of higher analogues of formamide. This possibility, however, cannot be considered by dynamic simulations because it demands propagation times that are too long.

Conclusions and Outlook

We have reviewed a series of investigations of π -conjugated molecules using mixed quantum-classical dynamics simulations. All these molecules have been chosen as models for biologically relevant systems. Dynamics simulations are able to provide information on time-dependent phenomena, which can only be obtained in a very indirect way by conventional static quantum-chemistry simulations. The main pieces of information brought by dynamics in excited states are the time constants for the relaxation processes following the photoexcitation and the relative importance of each available pathway.

In an effort to reduce computational demands, several levels of approximation are employed, starting with the dynamics methodology itself, which propagates the nuclear motion classically for all degrees of freedom, while partially recovering the quantum information of the non-adiabatic transitions between states. Although methodological progress has been achieved by the development of the multiple spawning (Ben-Nun et al. 2000; Hack et al. 2001; Levine et al. 2008) and MCTDH (Worth et al. 2003) approaches, the everyday method for tackling dynamics simulations in excited state is still the surface hopping approach, mainly due to its easy implementation and intuitive conceptual background.

At another level of approximation, simplified models have often been employed in substitution of the full biochemical system. In this chapter, we have discussed protonated Schiff base chains as models for the chromophore of rhodopsin proteins, heteroaromatic rings as models for UV-excited nucleic acids, and formamide as a model for photoexcited peptide bonds. With development of quantum chemical methods and computational resources, effective progress has been achieved by taking larger models at higher theoretical levels. This process, however, evolves very slowly and there are no perspectives of extending the quantum-mechanical treatment much beyond the chromophore region of the system. Instead, the development of hybrid methods, especially the quantum-mechanics/molecular-mechanics based approach, have been demonstrated to be of great utility in accounting for the effects of the molecular environment into the excited-state region of the molecule (Ciminelli et al. 2008; Hayashi et al. 2009; Levine et al. 2008; Lin and Truhlar 2007). Recently, methods for treating the excited-state evolution under the influence of an external electric field have been proposed (Jones et al. 2008; Mitric et al. 2009a; Yagi and Takatsuka 2005). They open the perspective of achieving a more complete treatment of the photo-excitation phenomena, by, for example, taking into account the effect of laser fields.

In a third level of approximations, the quantum chemical methods employed in dynamics simulations are naturally much lower than those employed at conventional quantum chemical calculations. As an example, dynamics simulations for molecules of biological interest are quite restricted to the use of double- ζ basis sets. Moreover, it is common to have a poor treatment of ionic states and non-adiabatic events, to neglect Rydberg states, and to neglect electron dynamic correlation. Although all these approximations are usually carefully employed by extended comparisons to high-level benchmark data, dynamics results obtained under such approximations should be regarded only as semi-quantitative.

Even in view of all these limitations, excited-state non-adiabatic dynamics simulations based on mixed quantum-classical methods constitute a fascinating and active research field providing essential information on the nature of molecular phenomena.

Acknowledgments The authors gratefully acknowledge computer time at the Vienna Scientific Cluster (project nos. 70019 and 70151). This work was supported by the Austrian Science Fund within the framework of the Special Research Program F41 Vienna Computational Materials Laboratory (ViCoM). The work in Zagreb (M.E.-M., I.A., and M.V.) is supported by the Ministry

of Science, Education and Sport of Croatia through the project no. 098-0982933-2920. The support by the COST D37 action, WG0001-06 and the WTZ treaty between Austria and Croatia (Project No. HR17/2008) is also acknowledged. This work was also performed as part of research supported by the National Science Foundation Partnership in International Research and Education (PIRE) Grant No. OISE-0730114; support was also provided by the Robert A. Welch Foundation under Grant No. D-0005.

Bibliography

- Abouaf, R., Pommier, J., & Dunet, H. (2003). Electronic and vibrational excitation in gas phase thymine and 5-bromouracil by electron impact. *Chemical Physics Letters*, 381(3–4), 486–494. doi:10.1016/j.cplett.2003.09.121.
- Alexandrova, A. N., Tully, J. C., & Granucci, G. (2010). Photochemistry of DNA fragments via semiclassical nonadiabatic dynamics. *The Journal of Physical Chemistry B*, 114(37), 12116–12128. doi:10.1021/jp103322c.
- Andruniow, T., Ferre, N., & Olivucci, M. (2004). Structure, initial excited-state relaxation, and energy storage of rhodopsin resolved at the multiconfigurational perturbation theory level. *Proceedings of the National Academy of Sciences of the United States of America*, 101(52), 17908–17913. doi:10.1073/pnas.0407997101.
- Antol, I., Eckert-Maksić, M., Barbatti, M., & Lischka, H. (2007). Simulation of the photodeactivation of formamide in the $n_{\sigma}\pi^{*}$ and $\pi\pi^{*}$ states: An ab initio on-the-fly surface-hopping dynamics study. *Journal of Chemical Physics*, 127(23), 234303–234308. doi:10.1063/1.2804862.
- Antol, I., Barbatti, M., Eckert-Maksić, M., & Lischka, H. (2008a). Quantum chemical calculations of electronically excited states: Formamide, its protonated form and alkali cation complexes as case studies. *Monatshefte für Chemie/Chemical Monthly*, 139(4), 319–328. doi:10.1007/s00706-007-0803-2.
- Antol, I., Vazdar, M., Barbatti, M., & Eckert-Maksić, M. (2008b). The effect of protonation on the photodissociation processes in formamide – An ab initio surface hopping dynamics study. *Chemical Physics*, 349(1–3), 308–318. doi:10.1016/j.chemphys.2008.01.026.
- Araujo, M., Lasorne, B., Bearpark, M. J., & Robb, M. A. (2008). The photochemistry of formaldehyde: Internal conversion and molecular dissociation in a single step? *The Journal of Physical Chemistry A*, 112(33), 7489–7491.
- Asturiol, D., Lasorne, B., Robb, M. A., & Blancafort, L. (2009). Photophysics of the $\pi\pi^{*}$ and n,π^{*} states of thymine: MS-CASPT2 minimum energy paths and CASSCF on-the-fly dynamics. *The Journal of Physical Chemistry A*, 113(38), 10211–10218. doi:10.1021/jp905303g.
- Back, R. A., & Boden, J. C. (1971). High-temperature photolysis and pyrolysis of formamide vapour, and thermal decomposition of carbamyl radical. *Transactions of the Faraday Society*, 67(577), 88–96.
- Barbatti, M., & Lischka, H. (2007). Can the nonadiabatic photodynamics of aminopyrimidine be a model for the ultrafast deactivation of adenine? *The Journal of Physical Chemistry A*, 111(15), 2852–2858. doi:10.1021/jp070089w.
- Barbatti, M., & Lischka, H. (2008). Nonadiabatic deactivation of 9H-adenine: A comprehensive picture based on mixed quantum-classical dynamics. *Journal of the American Chemical Society*, 130(21), 6831–6839. doi:10.1021/ja800589p.
- Barbatti, M., Vazdar, M., Aquino, A. J. A., Eckert-Maksić, M., & Lischka, H. (2006). The nonadiabatic deactivation paths of pyrrole. *Journal of Chemical Physics*, 125(16), 164323. doi:10.1063/1.2363376.
- Barbatti, M., Granucci, G., Persico, M., Ruckebauer, M., Vazdar, M., Eckert-Maksić, M., & Lischka, H. (2007a). The on-the-fly surface-hopping program system Newton-X: Application to ab initio simulation of the nonadiabatic photodynamics of benchmark systems. *Journal of Photochemistry and Photobiology A*, 190(2–3), 228–240. doi:10.1016/j.jphotochem.2006.12.008.

- Barbatti, M., Granucci, G., Ruckebauer, M., Pittner, J., Persico, M., & Lischka, H. (2007b). NEWTON-X: A package for Newtonian dynamics close to the crossing seam. www.newtonx.org
- Barbatti, M., Ruckebauer, M., Szymczak, J. J., Aquino, A. J. A., & Lischka, H. (2008). Nonadiabatic excited-state dynamics of polar p-systems and related model compounds of biological relevance. *Physical Chemistry Chemical Physics*, *10*, 482. doi:10.1039/b709315m.
- Barbatti, M., Aquino, A. J. A., Szymczak, J. J., Nachtigallová, D., Hobza, P., & Lischka, H. (2010a). Relaxation mechanisms of UV-photoexcited DNA and RNA nucleobases. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21453–21458. doi:10.1073/pnas.1014982107.
- Barbatti, M., Pittner, J., Pederzoli, M., Werner, U., Mitric, R., Bonacic-Koutecký, V., & Lischka, H. (2010b). Non-adiabatic dynamics of pyrrole: Dependence of deactivation mechanisms on the excitation energy. *Chemical Physics*, *375*(1), 26–34.
- Barbatti, M., Aquino, A. J. A., Szymczak, J. J., Nachtigallová, D., & Lischka, H. (2011a). Photodynamics simulations of cytosine: Characterization of the ultra fast bi-exponential UV deactivation. *Physical Chemistry Chemical Physics*, *13*, 6145–6155. doi:10.1039/C1030CP02142C.
- Barbatti, M., Szymczak, J. J., Aquino, A. J. A., Nachtigallová, D., & Lischka, H. (2011b). The decay mechanism of photoexcited guanine – A nonadiabatic dynamics study. *Journal of Chemical Physics*, *134*(1), 014304.
- Ben-Nun, M., Molnar, F., Lu, H., Phillips, J. C., Martinez, T. J., & Schulten, K. (1998). Quantum dynamics of the femtosecond photoisomerization of retinal in bacteriorhodopsin. *Faraday Discussions*, *110*, 447–462.
- Ben-Nun, M., Quenneville, J., & Martínez, T. J. (2000). Ab initio multiple spawning: Photochemistry from first principles quantum molecular dynamics. *The Journal of Physical Chemistry. A*, *104*(22), 5161–5175.
- Birge, R. R. (1981). Photophysics of light transduction in rhodopsin and bacteriorhodopsin. *Annual Review of Biophysics & Bioengineering*, *10*, 315–354.
- Birge, R. R. (1990). Nature of the primary photochemical events in rhodopsin and bacteriorhodopsin. *Biochimica et Biophysica Acta*, *1016*(3), 293–327.
- Blank, D. A., North, S. W., & Lee, Y. T. (1994). The ultraviolet photodissociation dynamics of pyrrole. *Chemical Physics*, *187*(1–2), 35–47.
- Boden, J. C., & Back, R. A. (1970). Photochemistry and free-radical reactions in formamide vapour. *Transactions of the Faraday Society*, *66*, 175–182.
- Boeyens, J. C. A. (1978). The conformation of six-membered rings. *Journal of Chemical Crystallography*, *8*, 317–320.
- Burghardt, I., & Hynes, J. T. (2006). Excited-state charge transfer at a conical intersection: Effects of an environment. *The Journal of Physical Chemistry. A*, *110*(40), 11411–11423.
- Cadet, J., & Berger, M. (1985). Radiation-induced decomposition of the purine-bases within DNA and related model compounds. *International Journal of Radiation Biology*, *47*(2), 127–143.
- Canuel, C., Mons, M., Piuze, F., Tardivel, B., Dimicoli, I., & Elhanine, M. (2005). Excited states dynamics of DNA and RNA bases: Characterization of a stepwise deactivation pathway in the gas phase. *Journal of Chemical Physics*, *122*(7), 074316.
- Chen, X.-B., & Fang, W.-H. (2004). Insights into photodissociation dynamics of benzamide and formamide from ab initio calculations. *Journal of the American Chemical Society*, *126*(29), 8976–8980.
- Chen, X.-B., Fang, W.-H., & Fang, D. C. (2003). An ab initio study toward understanding the mechanistic photochemistry of acetamide. *Journal of the American Chemical Society*, *125*(32), 9689–9698.
- Chin, C. H., Mebel, A. M., Kim, G. S., Baek, K. Y., Hayashi, M., Liang, K. K., & Lin, S. H. (2007). Theoretical investigations of spectroscopy and excited state dynamics of adenine. *Chemical Physics Letters*, *445*(4–6), 361–369.
- Ciminelli, C., Granucci, G., & Persico, M. (2008). The photoisomerization of a peptidic derivative of azobenzene: A nonadiabatic dynamics simulation of a supramolecular system. *Chemical Physics*, *349*(1–3), 325–333. doi:10.1016/j.chemphys.2008.01.030.

- Clark, L. B., & Tinoco, I. (1965). Correlations in ultraviolet spectra of purine and pyrimidine bases. *Journal of the American Chemical Society*, *87*(1), 11.
- Clark, L. B., Peschel, G. G., & Tinoco, I. (1965). Vapor spectra and heats of vaporization of some purine and pyrimidine bases. *The Journal of Physical Chemistry*, *69*(10), 3615–3618.
- Cremer, D., & Pople, J. A. (1975). General definition of ring puckering coordinates. *Journal of the American Chemical Society*, *97*(6), 1354–1358.
- Crespo-Hernández, C. E., Cohen, B., Hare, P. M., & Kohler, B. (2004). Ultrafast excited-state dynamics in nucleic acids. *Chemical Reviews*, *104*(4), 1977–2019.
- Crespo-Hernandez, C. E., Cohen, B., & Kohler, B. (2005). Base stacking controls excited-state dynamics in A-T DNA. *Nature*, *436*(7054), 1141–1144. doi:[10.1038/Nature03933](https://doi.org/10.1038/Nature03933).
- Cui, W., Thompson, M. S., & Reilly, J. P. (2005). Pathways of peptide ion fragmentation induced by vacuum ultraviolet light. *Journal of The American Society for Mass Spectrometry*, *16*(8), 1384–1398.
- Dallos, M., Lischka, H., Shepard, R., Yarkony, D. R., & Szalay, P. G. (2004). Analytic evaluation of nonadiabatic coupling terms at the MR-CI level. II. Minima on the crossing seam: Formaldehyde and the photodimerization of ethylene. *Journal of Chemical Physics*, *120*(16), 7330–7339.
- Daura, X., Antes, I., van Gunsteren, W. F., Thiel, W., & Mark, A. E. (1999). The effect of motional averaging on the calculation of NMR-derived structural properties. *Proteins-Structure Function and Genetics*, *36*(4), 542–555.
- Devine, A. L., Cronin, B., Nix, M. G. D., & Ashfold, M. N. R. (2006). High resolution photofragment translational spectroscopy studies of the near ultraviolet photolysis of imidazole. *Journal of Chemical Physics*, *125*(18), 184302, doi:[10.1063/1.2364504](https://doi.org/10.1063/1.2364504), Artn184302.
- Duggan, D. E., Bowman, R. L., Brodie, B. B., & Udenfriend, S. (1957). A spectrophotofluorometric study of compounds of biological interest. *Archives of Biochemistry and Biophysics*, *68*(1), 1–14.
- Eckert-Maksić, M., & Antol, I. (2009). Study of the mechanism of the N-CO photodissociation in N, N-Dimethylformamide by direct trajectory surface hopping simulations. *The Journal of Physical Chemistry: A*, *113*(45), 12582–12590. doi:[10.1021/jp9046177](https://doi.org/10.1021/jp9046177).
- Eckert-Maksić, M., Antol, I., Vazdar, M., Barbatti, M., & Lischka, H. (2010a). Formamide as the model compound for photodissociation studies of the peptide bond. In P. Paneth & A. Dybala-Defratyka (Eds.), *Kinetics and dynamics: Challenges and advances in computational chemistry and physics* (pp. 77–106). Netherlands: Springer.
- Eckert-Maksić, M., Vazdar, M., Ruckebauer, M., Barbatti, M., Muller, T., & Lischka, H. (2010b). Matrix-controlled photofragmentation of formamide: Dynamics simulation in argon by nonadiabatic QM/MM method. *Physical Chemistry Chemical Physics*, *12*(39), 12719–12726.
- Fabiano, E., & Thiel, W. (2008). Nonradiative deexcitation dynamics of 9H-adenine: An OM2 surface hopping study. *The Journal of Physical Chemistry: A*, *112*(30), 6859–6863. doi:[10.1021/jp8033402](https://doi.org/10.1021/jp8033402).
- Fabiano, E., Groenhof, G., & Thiel, W. (2008a). Approximate switching algorithms for trajectory surface hopping. *Chemical Physics*, *351*(1–3), 111–116.
- Fabiano, E., Keal, T. W., & Thiel, W. (2008b). Implementation of surface hopping molecular dynamics using semiempirical methods. *Chemical Physics*, *349*(1–3), 334–347. doi:[10.1016/j.chemphys.2008.01.044](https://doi.org/10.1016/j.chemphys.2008.01.044).
- Ferretti, A., Granucci, G., Lami, A., Persico, M., & Villani, G. (1996). Quantum mechanical and semiclassical dynamics at a conical intersection. *Journal of Chemical Physics*, *104*(14), 5517–5527.
- Frutos, L. M., Andruniow, T., Santoro, F., Ferre, N., & Olivucci, M. (2007). Tracking the excited-state time evolution of the visual pigment with multiconfigurational quantum chemistry. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(19), 7764–7769.
- Garavelli, M., Celani, P., Bernardi, F., Robb, M. A., & Olivucci, M. (1997). The C5H6NH2+ protonated Schiff base: An ab initio minimal model for retinal photoisomerization. *Journal of the American Chemical Society*, *119*(29), 6891–6901.

- Garavelli, M., Bernardi, F., Robb, M. A., & Olivucci, M. (1999a). The short-chain acroleini-
minium and pentadieniminium cations: Towards a model for retinal photoisomerization. A
CASSCF/PT2 study. *Journal of Molecular Structure: THEOCHEM*, 463(1–2), 59–64.
- Garavelli, M., Negri, F., & Olivucci, M. (1999b). Initial excited-state relaxation of the isolated
11-cis protonated Schiff base of retinal: Evidence for in-plane motion from ab initio quantum
chemical simulation of the resonance Raman spectrum. *Journal of the American Chemical
Society*, 121(5), 1023–1029.
- Gascon, J. A., & Batista, V. S. (2004). QM/MM study of energy storage and molecular
rearrangements due to the primary event in vision. *Biophysical Journal*, 87(5), 2931–2941.
doi:10.1529/biophysj.104.048264.
- Gingell, J. M., Mason, N. J., Zhao, H., Walker, I. C., & Siggel, M. R. F. (1997). VUV optical-
absorption and electron-energy-loss spectroscopy of formamide. *Chemical Physics*, 220(1–2),
191–205.
- González-Luque, R., Garavelli, M., Bernardi, F., Merchán, M., Robb, M. A., & Olivucci, M.
(2000). Computational evidence in favor of a two-state, two-mode model of the retinal
chromophore photoisomerization. *Proceedings of the National Academy of Sciences of the
United States of America*, 97(17), 9379–9384.
- González-Vázquez, J., & González, L. (2010). A time-dependent picture of the ultrafast deac-
tivation of keto-cytosine including three-state conical intersections. *A European Journal of
Chemical Physics and Physical Chemistry*, 11(17), 3617–3624.
- Granucci, G., & Persico, M. (2007). Critical appraisal of the fewest switches algorithm for surface
hopping. *Journal of Chemical Physics*, 126(13), 134114–134111.
- Granucci, G., Persico, M., & Toniolo, A. (2001). Direct semiclassical simulation of photochemical
processes with semiempirical wave functions. *Journal of Chemical Physics*, 114(24), 10608–
10615.
- Grégoire, G., Lucas, B., Barat, M., Fayeton, J. A., Dedonder-Lardeux, C., & Jouvét, C. (2009).
UV photoinduced dynamics in protonated aromatic amino acid. *European Physical Journal D*,
51(1), 109–116.
- Grégoire, G., Dedonder-Lardeux, C., Jouvét, C., Desfrancois, C., & Fayeton, J. A. (2007). Ultrafast
excited state dynamics in protonated GWG and GYG tripeptides. *Physical Chemistry Chemical
Physics*, 9(1), 78–82.
- Grégoire, G., Kang, H., Dedonder-Lardeux, C., Jouvét, C., Desfrancois, C., Onidas, D., Lepere,
V., & Fayeton, J. A. (2006). Statistical vs. non-statistical deactivation pathways in the UV
photo-fragmentation of protonated tryptophan-leucine dipeptide. *Physical Chemistry Chemical
Physics*, 8(1), 122–128.
- Greenberg, A., Breneman, C. M., & Liebman, J. F. (2002). *The amide linkage: Structural
significance in chemistry, biochemistry, and materials science. The amide linkage: Selected
structural aspects in chemistry, biochemistry, and materials science*. New York: Wiley.
- Gustavsson, T., Sarkar, N., Lazzarotto, E., Markovitsi, D., & Improtà, R. (2006). Singlet excited
state dynamics of uracil and thymine derivatives: A femtosecond fluorescence upconversion
study in acetonitrile. *Chemical Physics Letters*, 429(4–6), 551–557. doi:10.1016/j.cplett.2006.
08.058.
- Hack, M. D., Wensmann, A. M., Truhlar, D. G., Ben-Nun, M., & Martínez, T. J. (2001).
Comparison of full multiple spawning, trajectory surface hopping, and converged quantum
mechanics for electronically nonadiabatic dynamics. *Journal of Chemical Physics*, 115(3),
1172–1186.
- Hammes-Schiffer, S., & Tully, J. C. (1994). Proton-transfer in solution—Molecular-dynamics with
quantum transitions. *Journal of Chemical Physics*, 101(6), 4657–4667.
- Hayashi, S., Taikhorshid, E., & Schulten, K. (2009). Photochemical reaction dynamics of the
primary event of vision studied by means of a hybrid molecular simulation. *Biophysical Journal*,
96(2), 403–416. doi:10.1016/j.bpj.2008.09.049.
- He, Y., Wu, C., & Kong, W. (2003). Decay pathways of thymine and methyl-substituted uracil
and thymine in the gas phase. *The Journal of Physical Chemistry. A*, 107(26), 5145–5148.
doi:10.1021/jp034733s.

- Hudock, H. R., & Martinez, T. J. (2008). Excited-state dynamics of cytosine reveal multiple intrinsic subpicosecond pathways. *A European Journal of Chemical Physics and Physical Chemistry*, 9(17), 2486–2490. doi:10.1002/cphc.200800649.
- Hudock, H. R., Levine, B. G., Thompson, A. L., Satzger, H., Townsend, D., Gador, N., Ullrich, S., Stolow, A., & Martinez, T. J. (2007). Ab initio molecular dynamics and time-resolved photoelectron spectroscopy of electronically excited uracil and thymine. *The Journal of Physical Chemistry. A*, 111(34), 8500–8508.
- Improta, R., Barone, V., Lami, A., & Santoro, F. (2009). Quantum dynamics of the ultrafast $\pi\pi^*/n\pi^*$ population transfer in uracil and 5-fluoro-uracil in water and acetonitrile. *The Journal of Physical Chemistry B*, 113(43), 14491–14503.
- Ishida, T., Nanbu, S., & Nakamura, H. (2009). Nonadiabatic ab initio dynamics of two models of schiff base retinal. *The Journal of Physical Chemistry. A*, 113(16), 4356–4366. doi:10.1021/Jp8110315.
- Jasper, A. W., Stechmann, S. N., & Truhlar, D. G. (2002). Fewest-switches with time uncertainty: A modified trajectory surface-hopping algorithm with better accuracy for classically forbidden electronic transitions. *Journal of Chemical Physics*, 116(13), 5424–5431.
- Jeong, H. M., Young, S. S., Hyun, J. C., & Myung, S. K. (2007). Photodissociation at 193 nm of some singly protonated peptides and proteins with m/z 2000–9000 using a tandem time-of-flight mass spectrometer equipped with a second source for delayed extraction/post-acceleration of product ions. *Rapid Communications in Mass Spectrometry*, 21(3), 359–368.
- Jeong, H. M., So, H. Y., & Myung, S. K. (2005). Photodissociation of singly protonated peptides at 193 nm investigated with tandem time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 19(22), 3248–3252.
- Jones, G. A., Acocella, A., & Zerbetto, F. (2008). On-the-fly, electric-field-driven, coupled electron-nuclear dynamics. *The Journal of Physical Chemistry. A*, 112(40), 9650–9656. doi:10.1021/Jp805360v.
- Jorgensen, W. L., & McDonald, N. A. (1998). Development of an all-atom force field for heterocycles. Properties of liquid pyridine and diazenes. *Journal of Molecular Structure: Theochem*, 424(1–2), 145–155.
- Kang, T. Y., & Kim, H. L. (2006). Photodissociation of formamide at 205 nm: The H atom channels. *Chemical Physics Letters*, 431(1–3), 24–27.
- Kang, H., Jouvret, C., Dedonder-Lardeux, C., Martrenchard, S., Grégoire, G., Desfrancois, C., Schermann, J. P., Barat, M., & Fayetteon, J. A. (2005). Ultrafast deactivation mechanisms of protonated aromatic amino acids following UV excitation. *Physical Chemistry Chemical Physics*, 7(2), 394–398.
- Kang, H., Dedonder-Lardeux, C., Jouvret, C., Martrenchard, S., Grégoire, G., Desfrancois, C., Schermann, J. P., Barat, M., & Fayetteon, J. A. (2004). Photo-induced dissociation of protonated tryptophan TrpH⁺: A direct states dissociation channel in the excited controls the hydrogen atom loss. *Physical Chemistry Chemical Physics*, 6(10), 2628–2632.
- Kang, H., Lee, K. T., Jung, B., Ko, Y. J., & Kim, S. K. (2002). Intrinsic lifetimes of the excited state of DNA and RNA bases. *Journal of the American Chemical Society*, 124(44), 12958–12959.
- Keal, T., Wanko, M., & Thiel, W. (2009). Assessment of semiempirical methods for the photoisomerisation of a protonated Schiff base. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 123(1), 145–156.
- Kochendoerfer, G. G., & Mathies, R. A. (1996). Spontaneous emission study of the femtosecond isomerization dynamics of rhodopsin. *The Journal of Physical Chemistry*, 100(34), 14526–14532.
- Köppel, H., Gromov, E. V., & Trofimov, A. B. (2004). Multi-mode-multi-state quantum dynamics of key five-membered heterocycles: Spectroscopy and ultrafast internal conversion. *Chemical Physics*, 304(1–2), 35–49.
- Kukura, P., McCamant, D. W., Yoon, S., Wandschneider, D. B., & Mathies, R. A. (2005). Structural observation of the primary isomerization in vision with femtosecond-stimulated Raman. *Science*, 310(5750), 1006–1009.

- Lan, Z., & Domcke, W. (2008). Role of vibrational energy relaxation in the photoinduced nonadiabatic dynamics of pyrrole at the $^1\pi\sigma^* - S_0$ conical intersection. *Chemical Physics*, 350(1–3), 125–138.
- Lan, Z., Fabiano, E., & Thiel, W. (2009). Photoinduced nonadiabatic dynamics of pyrimidine nucleobases: On-the-fly surface-hopping study with semiempirical methods. *The Journal of Physical Chemistry B*, 113(11), 3548–3555. doi:10.1021/jp809085h.
- Langer, H., Doltsinis, N. L., & Marx, D. (2005). Excited-state dynamics and coupled proton–electron transfer of guanine. *A European Journal of Chemical Physics and Physical Chemistry*, 6(9), 1734–1737.
- Lasser, C., & Swart, T. (2008). Single switch surface hopping for a model of pyrazine. *Journal of Chemical Physics*, 129(3), 034302–034308.
- Lei, Y., Yuan, S., Dou, Y., Wang, Y., & Wen, Z. (2008). Detailed dynamics of the nonradiative deactivation of adenine: A semiclassical dynamics study. *The Journal of Physical Chemistry. A*, 112(37), 8497–8504. doi:10.1021/jp802483b.
- Levine, B. G., Ko, C., Quenneville, J., & Martínez, T. J. (2006). Conical intersections and double excitations in time-dependent density functional theory. *Molecular Physics*, 104(5–7), 1039–1051.
- Levine, B. G., Coe, J. D., Virshup, A. M., & Martinez, T. J. (2008). Implementation of ab initio multiple spawning in the MOLPRO quantum chemistry package. *Chemical Physics*, 347(1–3), 3–16. doi:10.1016/j.chemphys.2008.01.014.
- Li, X. S., Tully, J. C., Schlegel, H. B., & Frisch, M. J. (2005). Ab initio Ehrenfest dynamics. *Journal of Chemical Physics*, 123(8), 084106.
- Lin, H., & Truhlar, D. G. (2007). QM/MM: What have we learned, where are we, and where do we go from here? *Theoretical Chemistry Accounts*, 117(2), 185–199.
- Lippert, H., Ritze, H. H., Hertel, I. V., & Radloff, W. (2004). Femtosecond time-resolved hydrogen-atom elimination from photoexcited pyrrole molecules. *A European Journal of Chemical Physics and Physical Chemistry*, 5(9), 1423–1427.
- Lischka, H., Shepard, R., Brown, F. B., & Shavitt, I. (1981). New implementation of the graphical unitary-group approach for multi-reference direct configuration-interaction calculations. *International Journal of Quantum Chemistry*, S.15, 91–100.
- Lischka, H., Shepard, R., Pitzer, R. M., Shavitt, I., Dallos, M., Müller, T., Szalay, P. G., Seth, M., Kedziora, G. S., Yabushita, S., & Zhang, Z. Y. (2001). High-level multireference methods in the quantum-chemistry program system COLUMBUS: Analytic MR-CISD and MR-AQCC gradients and MR-AQCC-LRT for excited states, GUGA spin-orbit CI and parallel CI density. *Physical Chemistry Chemical Physics*, 3(5), 664–673.
- Lischka, H., Dallos, M., & Shepard, R. (2002). Analytic MRCI gradient for excited states: Formalism and application to the $n-\pi^*$ valence- and $n-(3s,3p)$ Rydberg states of formaldehyde. *Molecular Physics*, 100(11), 1647–1658.
- Lischka, H., Dallos, M., Szalay, P. G., Yarkony, D. R., & Shepard, R. (2004). Analytic evaluation of nonadiabatic coupling terms at the MR-CI level. I. Formalism. *Journal of Chemical Physics*, 120(16), 7322–7329.
- Lischka, H., Shepard, R., Shavitt, I., Pitzer, R. M., Dallos, M., Mueller, T., Szalay, P. G., Brown, F. B., Ahlrichs, R., Boehm, H. J., Chang, A., Comeau, D. C., Gdanitz, R., Dachsels, H., Ehrhardt, C., Ernzerhof, M., Hocht, P., Irl, S., Kedziora, G., Kovar, T., Parasuk, V., Pepper, M. J. M., Scharf, P., Schiffer, H., Schindler, M., Schueler, M., Seth, M., Stahlberg, E. A., Zhao, J.-G., Yabushita, S., Zhang, Z., Barbatti, M., Matsika, S., Schuurmann, M., Yarkony, D. R., Brozell, S. R., Beck, E. V., & Blaudeau, J.-P. (2006). COLUMBUS, an ab initio electronic structure program, release 5.9.1. www.univie.ac.at/columbus
- Liu, R. S. H., & Asato, A. E. (1985). Photochemistry of polyenes. 22. The primary process of vision and the structure of bathorhodopsin – A mechanism for photoisomerization of polyenes. *Proceedings of the National Academy of Sciences of the United States of America*, 82(2), 259–263.
- Liu, R. S. H. (2002). Introduction to the symposium-in-print: Photoisomerization pathways, torsional relaxation and the Hula Twist & para. *Photochemistry and Photobiology*, 76(6), 580–583.

- Liu, D., Fang, W. H., & Fu, X. Y. (2000). Ab initio molecular orbital study of the mechanism of photodissociation of formamide. *Chemical Physics Letters*, *318*(4–5), 291–297.
- Logunov, I., & Schulten, K. (1996). Quantum chemistry: Molecular dynamics study of the dark-adaptation process in bacteriorhodopsin. *Journal of the American Chemical Society*, *118*(40), 9727–9735.
- Longworth, J. W., Rahn, R. O., & Shulman, R. G. (1966). Luminescence of pyrimidines purines nucleosides and nucleotides at 77 degrees K. Effect of ionization and tautomerization. *Journal of Chemical Physics*, *45*(8), 2930.
- Lundell, J., Krajewska, M., & Räsänen, M. (1998). Matrix isolation Fourier transform infrared and ab initio studies of the 193-nm-induced photodecomposition of formamide. *The Journal of Physical Chemistry. A*, *102*(33), 6643–6650.
- Malone, R. J., Miller, A. M., & Kohler, B. (2003). Singlet excited-state lifetimes of cytosine derivatives measured by femtosecond transient absorption. *Photochemistry and Photobiology*, *77*(2), 158–164.
- Marian, C. M. (2005). A new pathway for the rapid decay of electronically excited adenine. *Journal of Chemical Physics*, *122*(10), 104314.
- Matsika, S. (2004). Radiationless decay of excited states of uracil through conical intersections. *The Journal of Physical Chemistry. A*, *108*(37), 7584–7590.
- Merchan, M., Gonzalez-Luque, R., Climent, T., Serrano-Andres, L., Rodriiguez, E., Reguero, M., & Pelaez, D. (2006). Unified model for the ultrafast decay of pyrimidine nucleobases. *The Journal of Physical Chemistry B*, *110*(51), 26471–26476. doi:10.1021/Jp066874a.
- Middleton, C. T., de La Harpe, K., Su, C., Law, Y. K., Crespo-Hernandez, C. E., & Kohler, B. (2009). DNA excited-state dynamics: From single bases to the double helix. *Annual Review of Physical Chemistry*, *60*(1), 217–239. doi:10.1146/annurev.physchem.59.032607.093719.
- Migani, A., Robb, M. A., & Olivucci, M. (2003). Relationship between photoisomerization path and intersection space in a retinal chromophore model. *Journal of the American Chemical Society*, *125*(9), 2804–2808.
- Mitric, R., Petersen, J., & Bonacic-Koutecky, V. (2009a). Laser-field-induced surface-hopping method for the simulation and control of ultrafast photodynamics. *Physical Review A*, *79*(5), 053416.
- Mitric, R., Werner, U., Wohlgemuth, M., Seifert, G., & Bonacic-Koutecky, V. (2009b). Nonadiabatic dynamics within time-dependent density functional tight binding method. *The Journal of Physical Chemistry. A*, *113*(45), 12700–12705. doi:10.1021/Jp905600w.
- Mulcahy, M., McInerney, J. G., Nikogosyan, D. N., & Görner, H. (2000). 193 Nm photolysis of aromatic and aliphatic dipeptides in aqueous solution: Dependence of decomposition quantum yield on the amino acid sequence. *Biological Chemistry*, *381*(12), 1259–1262.
- Muller, U., & Stock, G. (1997). Surface-hopping modeling of photoinduced relaxation dynamics on coupled potential-energy surfaces. *Journal of Chemical Physics*, *107*(16), 6230–6245.
- Nachtigallova, D., Zeleny, T., Ruckebauer, M., Muller, T., Barbatti, M., Hobza, P., & Lischka, H. (2010). Does stacking restrain the photodynamics of individual nucleobases? *Journal of the American Chemical Society*, *132*(24), 8261–8263. doi:10.1021/ja1029705.
- Nieber, H., & Doltsinis, N. L. (2008). Elucidating ultrafast nonradiative decay of photoexcited uracil in aqueous solution by ab initio molecular dynamics. *Chemical Physics*, *347*(1–3), 405–412. doi:10.1016/j.chemphys.2007.09.056.
- Nolting, D., Schultz, T., Hertel, I. V., & Weinkauff, R. (2006). Excited state dynamics and fragmentation channels of the protonated dipeptide H2N-Leu-Trp-COOH. *Physical Chemistry Chemical Physics*, *8*(44), 5247–5254. doi:10.1039/B609726.
- Palings, I., Pardoën, J. A., Vandenberg, E., Winkel, C., Lugtenburg, J., & Mathies, R. A. (1987). Assignment of fingerprint vibrations in the resonance Raman-spectra of rhodopsin, isorhodopsin, and bathorhodopsin – Implications for chromophore structure and environment. *Biochemistry (Mosc)*, *26*(9), 2544–2556.
- Pei, K.-M., Ma, Y., & Zheng, X. (2008). Resonance Raman and theoretical investigation of the photodissociation dynamics of benzamide in S_3 state. *The Journal of Chemical Physics*, *128*(22), 224310–224310.

- Perun, S., Sobolewski, A. L., & Domcke, W. (2005a). Ab initio studies on the radiationless decay mechanisms of the lowest excited singlet states of 9H-adenine. *Journal of the American Chemical Society*, *127*(17), 6257–6265.
- Perun, S., Sobolewski, A. L., & Domcke, W. (2005b). Photostability of 9H-adenine: Mechanisms of the radiationless deactivation of the lowest excited singlet states. *Chemical Physics*, *313*(1–3), 107–112.
- Peteanu, L. A., Schoenlein, R. W., Wang, Q., Mathies, R. A., & Shank, C. V. (1993). The 1st Step in vision occurs in femtoseconds – Complete blue and red spectral studies. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(24), 11762–11766.
- Petersen, C., Dahl, N. H., Jensen, S. K., Poulsen, J. A., Thøgersen, J., & Keiding, S. R. (2008). Femtosecond Photolysis of Aqueous Formamide. *The Journal of Physical Chemistry. A*, *112*(15), 3339–3344.
- Pittner, J., Lischka, H., & Barbatti, M. (2009). Optimization of mixed quantum-classical dynamics: Time-derivative coupling terms and selected couplings. *Chemical Physics*, *356*(1–3), 147–152. doi:[10.1016/j.chemphys.2008.10.013](https://doi.org/10.1016/j.chemphys.2008.10.013).
- Polli, D., Altoe, P., Weingart, O., Spillane, K. M., Manzoni, C., Brida, D., Tomasello, G., Orlandi, G., Kukura, P., Mathies, R. A., Garavelli, M., & Cerullo, G. (2010). Conical intersection dynamics of the primary photoisomerization event in vision. *Nature*, *467*(7314), 440–443. doi:[10.1038/nature09346](https://doi.org/10.1038/nature09346).
- Poterya, V., Profant, V., Farnik, M., Slavicek, P., & Buck, U. (2007). Experimental and theoretical study of the pyrrole cluster photochemistry: Closing the $\pi\sigma^*$ dissociation pathway by complexation. *Journal of Chemical Physics*, *127*(6), 064307. doi:[10.1063/1.2754687](https://doi.org/10.1063/1.2754687).
- Powner, M. W., Gerland, B., & Sutherland, J. D. (2009). Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature*, *459*(7244), 239–242. doi:[10.1038/Nature08013](https://doi.org/10.1038/Nature08013).
- Rohrig, U. F., Guidoni, L., Laio, A., Frank, I., & Rothlisberger, U. (2004). A molecular spring for vision. *Journal of the American Chemical Society*, *126*(47), 15328–15329. doi:[10.1021/ja048265r](https://doi.org/10.1021/ja048265r).
- Rohrig, U. F., Guidoni, L., & Rothlisberger, U. (2005). Solvent and protein effects on the structure and dynamics of the rhodopsin chromophore. *A European Journal of Chemical Physics and Physical Chemistry*, *6*(9), 1836–1847. doi:[10.1002/cphc.200500066](https://doi.org/10.1002/cphc.200500066).
- Ruckenbauer, M., Barbatti, M., Muller, T., & Lischka, H. (2010). Nonadiabatic excited-state dynamics with hybrid ab initio quantum-mechanical/molecular-mechanical methods: Solvation of the pentadieniminium cation in Apolar media. *The Journal of Physical Chemistry. A*, *114*(25), 6757–6765. doi:[10.1021/jp103101t](https://doi.org/10.1021/jp103101t).
- Saam, J., Tajkhorshid, E., Hayashi, S., & Schulten, K. (2002). Molecular dynamics investigation of primary photoinduced events in the activation of rhodopsin. *Biophysical Journal*, *83*(6), 3097–3112.
- Santoro, F., Barone, V., & Improta, R. (2007a). Influence of base stacking on excited-state behavior of polyadenine in water, based on time-dependent density functional calculations. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(24), 9931–9936. doi:[10.1073/pnas.0703298104](https://doi.org/10.1073/pnas.0703298104).
- Santoro, F., Lami, A., & Olivucci, M. (2007b). Complex excited dynamics around a plateau on a retinal-like potential surface: Chaos, multi-exponential decays and quantum/classical differences. *Theoretical Chemistry Accounts*, *117*(5–6), 1061–1072. doi:[10.1007/s00214-006-0220-3](https://doi.org/10.1007/s00214-006-0220-3).
- Shapiro, I., Weingart, O., & Buss, V. (2009). Bicycle-pedal isomerization in a rhodopsin chromophore model. *Journal of the American Chemical Society*, *131*(1), 16. doi:[10.1021/Ja805586z](https://doi.org/10.1021/Ja805586z).
- Schoenlein, R. W., Peteanu, L. A., Mathies, R. A., & Shank, C. V. (1991). The 1st step in vision – Femtosecond isomerization of rhodopsin. *Science*, *254*(5030), 412–415.
- Schultz, T., Samoylova, E., Radloff, W., Hertel, I. V., Sobolewski, A. L., & Domcke, W. (2004). Efficient deactivation of a model base pair via excited-state hydrogen transfer. *Science*, *306*(5702), 1765–1768.

- Sellner, B., Barbatti, M., & Lischka, H. (2009). Dynamics starting at a conical intersection: Application to the photochemistry of pyrrole. *Journal of Chemical Physics*, *131*(2), 024312. doi:10.1063/1.3175799.
- Send, R., & Sundholm, D. (2007). Stairway to the conical intersection: A computational study of the retinal isomerization. *The Journal of Physical Chemistry. A*, *111*(36), 8766–8773.
- Serrano-Andres, L., & Merchán, M. (2009). Are the five natural DNA/RNA base monomers a good choice from natural selection? A photochemical perspective. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, *10*(1), 21–32. doi:10.1016/j.jphotochemrev.2008.12.001.
- Shepard, R. (1995). The analytic gradient method for configuration interaction wave functions. In D. R. Yarkony (Ed.), *Modern electronic structure theory. Advanced series in physical chemistry* (Vol. 1, p. 345). Singapore: World Scientific.
- Shepard, R., Lischka, H., Szalay, P. G., Kovar, T., & Ernzerhof, M. (1992). A general multireference configuration-interaction gradient program. *Journal of Chemical Physics*, *96*(3), 2085–2098.
- Shin, E. J. (2004). Photochemistry of anthrylene derivatives containing heteroaromatic ring: Pyrrole and indole derivatives. *Bulletin of the Korean Chemical Society*, *25*(6), 907–909.
- Sobolewski, A. L., & Domcke, W. (2000). Conical intersections induced by repulsive $^1\pi\sigma^*$ states in planar organic molecules: Malonaldehyde, pyrrole and chlorobenzene as photochemical model systems. *Chemical Physics*, *259*(2–3), 181–191.
- Sobolewski, A. L., Domcke, W., & Hättig, C. (2005). Tautomeric selectivity of the excited-state lifetime of guanine/cytosine base pairs: The role of electron-driven proton-transfer processes. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(50), 17903–17906.
- Szymczak, J. J., Barbatti, M., & Lischka, H. (2008). Mechanism of ultrafast photodecay in restricted motions in protonated Schiff bases: The pentadieniminium cation. *Journal of Chemical Theory and Computation*, *4*(8), 1189–1199. doi:10.1021/Ct800148n.
- Szymczak, J. J., Barbatti, M., & Lischka, H. (2009). Is the photoinduced isomerization in retinal protonated schiff bases a single- or double-torsional process? *The Journal of Physical Chemistry. A*, *113*(43), 11907–11918. doi:10.1021/jp903329j.
- Szymczak, J. J., Barbatti, M., & Lischka, H. (2011). Influence of the active space on CASSCF nonadiabatic dynamics simulations. *International Journal of Quantum Chemistry*, *111*(13), 3307–3315. doi:10.1002/qua.22978.
- Tapavicza, E., Tavernelli, I., & Rothlisberger, U. (2007). Trajectory surface hopping within linear response time-dependent density-functional theory. *Physical Review Letters*, *98*(2), 023001.
- Thompson, M. S., Cui, W., & Reilly, J. P. (2007). Factors that impact the vacuum ultraviolet photofragmentation of peptide ions. *Journal of The American Society for Mass Spectrometry*, *18*(8), 1439–1452.
- Thompson, M. S., Cui, W., & Reilly, J. P. (2004). Fragmentation of singly charged peptide ions by photodissociation at $\hat{\nu} = 157$ nm. *Angewandte Chemie-International Edition*, *43*(36), 4791–4794.
- Torikai, A., & Shibata, H. (1999). Effect of ultraviolet radiation on photodegradation of collagen. *Journal of Applied Polymer Science*, *73*(7), 1259–1265.
- Tully, J. C. (1990). Molecular-dynamics with electronic-transitions. *Journal of Chemical Physics*, *93*(2), 1061–1071.
- Tully, J. C. (1998). Mixed quantum-classical dynamics. *Faraday Discussions*, *110*, 407–419.
- Ullrich, S., Schultz, T., Zgierski, M. Z., & Stolow, A. (2004). Electronic relaxation dynamics in DNA and RNA bases studied by time-resolved photoelectron spectroscopy. *Physical Chemistry Chemical Physics*, *6*(10), 2796–2801.
- Vallet, V., Lan, Z. G., Mahapatra, S., Sobolewski, A. L., & Domcke, W. (2004). Time-dependent quantum wave-packet description of the $^1\pi\sigma^*$ photochemistry of pyrrole. *Faraday Discussions*, *127*, 283–293.

- Vallet, V., Lan, Z. G., Mahapatra, S., Sobolewski, A. L., & Domcke, W. (2005). Photochemistry of pyrrole: Time-dependent quantum wave-packet description of the dynamics at the ${}^1\text{GREEK}(\text{ps})^* - S_0$ conical intersections. *Journal of Chemical Physics*, *123*(14), 144307.
- van den Brom, A. J., Kapelios, M., Kitsopoulos, T. N., Nahler, N. H., Cronin, B., & Ashfold, M. N. R. (2005). Photodissociation and photoionization of pyrrole following the multiphoton excitation at 243 and 364.7 nm. *Physical Chemistry Chemical Physics*, *7*(5), 892–899.
- Vazdar, M., Eckert-Maksić, M., Barbatti, M., & Lischka, H. (2009). Excited-state non-adiabatic dynamics simulations of pyrrole. *Molecular Physics*, *107*(8), 845–854. doi:[10.1080/00268970802665639](https://doi.org/10.1080/00268970802665639).
- Venkatesan, T. S., Mahapatra, S., Meyer, H. D., Koppel, H., & Cederbaum, L. S. (2007). Multimode Jahn-Teller and pseudo-Jahn-Teller interactions in the cyclopropane radical cation: Complex vibronic spectra and nonradiative decay dynamics. *The Journal of Physical Chemistry. A*, *111*(10), 1746–1761.
- Virshup, A. M., Punwong, C., Pogorelov, T. V., Lindquist, B. A., Ko, C., & Martinez, T. J. (2009). Photodynamics in complex environments: Ab initio multiple spawning quantum mechanical/molecular mechanical dynamics. *The Journal of Physical Chemistry B*, *113*(11), 3280–3291. doi:[10.1021/Jp8073464](https://doi.org/10.1021/Jp8073464).
- Vreven, T., Bernardi, F., Garavelli, M., Olivucci, M., Robb, M. A., & Schlegel, H. B. (1997). Ab initio photoisomerization dynamics of a simple retinal chromophore model. *Journal of the American Chemical Society*, *119*(51), 12687–12688.
- Wald, G. (1968). Molecular basis of visual excitation. *Science*, *162*(3850), 230.
- Wang, Q., Schoenlein, R. W., Peteanu, L. A., Mathies, R. A., & Shank, C. V. (1994). Vibrationally coherent photochemistry in the femtosecond primary event of vision. *Science*, *266*(5184), 422–424.
- Wanko, M., Hoffmann, M., Strodet, P., Koslowski, A., Thiel, W., Neese, F., Frauenheim, T., & Elstner, M. (2005). Calculating absorption shifts for retinal proteins: Computational challenges. *The Journal of Physical Chemistry B*, *109*(8), 3606–3615.
- Warshel, A. (1976). Bicycle-pedal model for 1st step in vision process. *Nature*, *260*(5553), 679–683.
- Warshel, A., & Barboy, N. (1982). Energy-storage and reaction pathways in the 1st step of the vision process. *Journal of the American Chemical Society*, *104*(6), 1469–1476.
- Warshel, A., & Chu, Z. T. (2001). Nature of the surface crossing process in bacteriorhodopsin: Computer simulations of the quantum dynamics of the primary photochemical event. *The Journal of Physical Chemistry B*, *105*(40), 9857–9871.
- Wei, J., Riedel, J., Kuczmann, A., Renth, F., & Temps, F. (2004). Photodissociation dynamics of pyrrole: Evidence for mode specific dynamics from conical intersections. *Faraday Discussions*, *127*, 267–282.
- Weingart, O., Buss, V., & Robb, M. A. (2005). Excited state molecular dynamics of retinal model chromophores. *Phase Transitions*, *78*(1–3), 17–24.
- Weingart, O., Migani, A., Olivucci, M., Robb, M. A., Buss, V., & Hunt, P. (2004). Probing the photochemical funnel of a retinal chromophore model via zero-point energy sampling semiclassical dynamics. *The Journal of Physical Chemistry A*, *108*(21), 4685–4693.
- Weingart, O., Schapiro, I., & Buss, V. (2006). Bond torsion affects the product distribution in the photoreaction of retinal model chromophores. *Journal of Molecular Modeling*, *12*(5), 713–721.
- Weingart, O., Schapiro, I., & Buss, V. (2007). Photochemistry of visual pigment chromophore models by ab initio molecular dynamics. *The Journal of Physical Chemistry B*, *111*(14), 3782–3788.
- Werner, U., Mitric, R., Suzuki, T., & Bonacic-Koutecký, V. (2008). Nonadiabatic dynamics within the time dependent density functional theory: Ultrafast photodynamics in pyrazine. *Chemical Physics*, *349*(1–3), 319–324.
- Worth, G. A., & Cederbaum, L. S. (2004). Beyond Born-Oppenheimer: Molecular dynamics through a conical intersection. *Annual Review of Physical Chemistry*, *55*, 127–158.

- Worth, G. A., Hunt, P., & Robb, M. A. (2003). Nonadiabatic dynamics: A comparison of surface hopping direct dynamics with quantum wave packet calculations. *The Journal of Physical Chemistry. A*, *107*(5), 621–631.
- Yagi, K., & Takatsuka, K. (2005). Nonadiabatic chemical dynamics in an intense laser field: Electronic wave packet coupled with classical nuclear motions. *Journal of Chemical Physics*, *123*(22), 224103. doi:[10.1063/1.2130335](https://doi.org/10.1063/1.2130335), Artn224103.
- Zechmann, G., & Barbatti, M. (2008). Photophysics and deactivation pathways of thymine. *The Journal of Physical Chemistry. A*, *112*(36), 8273–8279. doi:[10.1021/jp804309x](https://doi.org/10.1021/jp804309x).
- Zewail, A. H. (2000). Femtochemistry: Atomic-scale dynamics of the chemical bond. *The Journal of Physical Chemistry. A*, *104*(24), 5660–5694. doi:[10.1021/Jp001460h](https://doi.org/10.1021/Jp001460h).
- Zhu, C. Y., Nangia, S., Jasper, A. W., & Truhlar, D. G. (2004). Coherent switching with decay of mixing: An improved treatment of electronic coherence for non-Born-Oppenheimer trajectories. *Journal of Chemical Physics*, *121*(16), 7658–7670.

Low-Energy Electron (LEE)-Induced DNA Damage: Theoretical Approaches to Modeling Experiment

40

Anil Kumar and Michael D. Sevilla

Contents

Introduction	1742
Proposed Mechanism of LEE-Induced DNA Strand Breaks	1745
Electron–Molecule Interaction Events	1747
Resonance (TNI) Formation: A Molecular Orbital Approach	1748
Shape Resonances of DNA Bases	1752
Electron Attachment to DNA/RNA Bases in Gas Phase	1756
Choice of the Basis Set	1759
Effect of Solvation (Condensed Media)	1760
Physical Observables of Solvated Electron (e_{aq}^-)	1762
Proposed Theoretical Models of DNA Damage	1766
Excited States of TNI (Resonance Formation)	1771
Profile of DNA Damage Produced by LEE and Ionizing Radiation	1779
LEE-Induced Base Release	1780
Effect of Solvation on Strand Break Formation	1781
LEE Interaction with 5-Bromouracil (5-BrU) and Radiosensitization Activity	1784
Hydration of the Uracil-5-yl Radical	1785
Modeling 5-Substituted Uracil as an Efficient DNA Radiosensitizer	1788
Conclusion	1789
Bibliography	1790

Abstract

Low-energy electrons (LEE) have been experimentally found to structurally alter DNA by induction of base damage, base release, and strand breaks. This has engendered a considerable number of theoretical studies of the mechanisms involved in this DNA damage. In this chapter, we discuss the various pathways for LEE interaction with DNA and the theoretical treatments most suited to

A. Kumar (✉) • M.D. Sevilla

Department of Chemistry, Oakland University, Rochester, MI, USA

e-mail: kumar@oakland.edu; sevilla@oakland.edu

unravel these pathways. For example, inelastic electron scattering produces excitation, ionization, and transient negative ions (TNI) via shape, core-excited, and vibrational Feshbach resonances, which can all lead to DNA damage. Each of these pathways is distinguished, and their pertinence to the experimental results observed is described. Theoretical approaches used to explain these pathways are also described. Shape resonances can be understood as interactions with the electron with unoccupied molecular orbitals of neutral molecule, while core-excited states involve excitation of inner shell electrons and can be treated with theoretical methods such as time-dependent density functional theory (TD-DFT) or CASSCF. In treating the electron–molecule interaction, special care is needed to distinguish between diffuse and valence states of the TNI. The role of the vertical and adiabatic states of the radical anion is important as the electron adds to the neutral molecular framework, and reactions induced likely occur before equilibration to the adiabatic state. The effect of solvation is critical to both energetics of the interaction and the nature of the TNI formed. For example, gas-phase calculations show diffuse dipole-bound character for adenine, guanine, and cytosine anion radicals, but each of these is found to be in a valence state in aqueous solution by experiment. DNA base anion radicals often show ground states that are diffuse in character and that collapse to valence states on solvation. Such processes are shown to be accounted for inclusion of the polarized continuum model (PCM) for solvation. TD-DFT excited-state calculations including solvation show that the diffuse states rise in energy on solvation as expected. For LEE in the aqueous phase, new energy states become available such as conduction band or presolvated electrons, which may have sufficient energy to cause DNA damage. DNA radiosensitizers such as halogen-substituted uracils are being used to produce highly reactive uracil-5-yl radical, a precursor of DNA strand breaks, through its reaction with LEE.

Introduction

The DNA sequence contains the information necessary for the functioning of all cellular processes. Exposure of DNA to high-energy radiation leads to disruption of this sequence or actual strand cleavage that results in potential mutagenic or lethal damage (Becker and Sevilla 1993, 2008; Swarts et al. 1992; Becker et al. 2007, 2010; Li and Sevilla 2007; Swiderek 2006; Kumar and Sevilla 2008b, 2010a, b; Sevilla and Bernhard 2008; Sanche 2008; Yokoya et al. 2008; Denifl et al. 2007; von Sonntag 1991). In the early stage of the radiation-induced DNA damage, high-energy radiation ionizes and excites each component of DNA (i.e., base, sugar, and phosphate backbone) and the surrounding water molecules randomly and produces “holes,” secondary electrons, and excited states (Kumar and Sevilla 2008b, 2010a, b; Sevilla and Bernhard 2008; Sevilla et al. 1991). The hole (positive charge or radical cation), produced during ionization event in DNA, usually transfers to the site in DNA with the lowest ionization energy. Of the four DNA bases (adenine [A], guanine [G], cytosine [C], and thymine [T]), guanine is found to have the lowest ionization potential (Kumar and Sevilla 2008b; Hush and Cheung 1975;

Orlov et al. 1976; Yang et al. 2004; Steenken and Jovanovic 1997) and acts as the locus for hole transfer in DNA (Becker and Sevilla 1993, 2008; Swarts et al. 1992; Becker et al. 2010; Li and Sevilla 2007; Kumar and Sevilla 2008b, 2010a, b; Sevilla and Bernhard 2008; Sevilla et al. 1991). The guanine radical cation ($G^{\bullet+}$), formed on ionization, reacts further to produce oxidative DNA damage in the form of 8-oxoG (Burrows and Muller 1998; Cadet et al. 2008). A hole created on the sugar phosphate backbone undergoes two competitive reactions: (a) formation of neutral sugar radical after deprotonation ($-H^+$) from the specific sites of the sugar ring (Becker et al. 2003; Shukla et al. 2004, 2005; Adhikary et al. 2005, 2006a, b, 2007, 2008; Khanduri et al. 2008) and (b) hole transfer to the nearby base in DNA followed by *base-to-base* transfer to guanine (Adhikary et al. 2006b). Secondary electrons formed by the ionizing event lose their kinetic energy, thermalize, add to DNA, and transfer to the bases of highest electron affinities, that is, the pyrimidines, T and C. This produces the anion radicals ($T^{\bullet-}$ and $C^{\bullet-}$) (Sevilla et al. 1991). Secondary electrons before thermalization possess substantial kinetic energy, which can lead to DNA damage. Such electrons with kinetic energies below 20 eV are known as *low-energy electrons* (LEE) (Sanche 2008; Pimblott et al. 1996; Pimblott and LaVerne 2007) and are produced in great numbers ($\sim 4 \times 10^4$ electrons per MeV energy deposited) (International Commission on Radiation Units and Measurements 1979) along the tracks of the ionizing radiation. LEEs have been recognized as the potential significant contributor to DNA damage. Collisions of LEEs with the aqueous medium result in most being thermalized within 10^{-12} s to become solvated electrons. Those LEEs produced in or near DNA are able to damage DNA via resonant absorption processes (Kumar and Sevilla 2010a). In their pioneering work, Sanche and coworkers (Boudaïffa et al. 2000; Huels et al. 2003; Sanche 2005; Zheng et al. 2005, 2006, 2008; Ptasinska and Sanche 2007a, b) established that LEEs are able to create single- and double-strand breaks (SSB and DSB) in plasmid DNA through a dissociative electron attachment (DEA) mechanism. This finding initiated several studies by a number of workers on the interaction of LEEs with DNA and its components (Baccarelli et al. 2007; Sulzer et al. 2006; Bald et al. 2006, 2008; König et al. 2006; Ptasinska et al. 2005a, 2006). A number of other recent reviews have focused on this important area of research (Li and Sevilla 2007; Swiderek 2006; Kumar and Sevilla 2008b, 2010a, b; Sanche 2008, 2009, 2010; Rak et al. 2008). The initial events induced in DNA by high-energy radiation discussed above are summarized in Fig. 1. The events shown in Fig. 1 follow the oxidative and reductive processes that are initiated by radiation to the dominant end products formed. When a number of the damage sites, shown in Fig. 1, are formed in proximity, a multiple damage site (MDS) is formed. MDSs are often irreparable owing to the loss of DNA information and the inability of enzymatic repair processes to recognize individual damages. They therefore represent the most lethal type of damage.

Ionization and, to a lesser extent, excitation as outline above were in the past considered the primary causes of DNA radiation damage. However, results have strongly implicated the role of low-energy electrons in this damage, and their study has been of intense interest in recent times. In this chapter, we will be focusing on the interaction of LEE with DNA, which leads to a variety of physical and chemical changes in DNA. A number of efforts to elucidate the mechanisms of

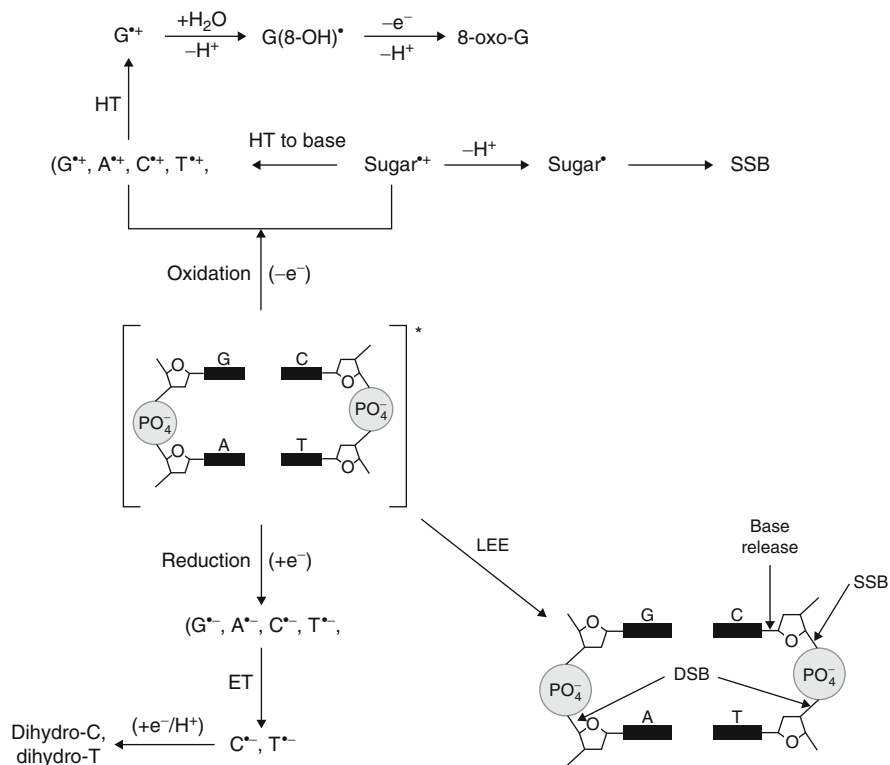


Fig. 1 Schematic diagram showing the overall processes (oxidation, reduction, and LEE-induced DNA damage) that occur during high-energy radiation interaction with DNA. *HT* hole transfer, *ET* electron transfer

LEE-induced DNA damage have been made in recent years employing theory (Li and Sevilla 2007; Kumar and Sevilla 2008b, 2010a; Rak et al. 2008; Simons 2006) and experiments (Sanche 2005, 2008, 2009; Denifl et al. 2007; Boudaïffa et al. 2000; Huels et al. 2003; Zheng et al. 2005, 2006, 2008; Ptasńska and Sanche 2007a, b; Baccarelli et al. 2007; Sulzer et al. 2006; Bald et al. 2006, 2008; König et al. 2006; Ptasńska et al. 2005a, 2006; Rak et al. 2008). Recent efforts have implicated the role of vibrational and electronic excited states in LEE interactions with DNA. The advent of substantial computing power at low cost and the accessibility of computational resources at national centers (Baker et al. 2009) allow the use of rigorous ab initio quantum chemical methods such as density functional theory (DFT), Møller–Plesset perturbation theory (MP2), CCSD(T), and CASPT2 for DNA components. Such theoretical studies shed light on the underlying mechanisms of LEE-induced DNA damage. The applicability of MP2, CCSD(T), and CASPT2 methods are still limited to small-to-modest size molecules such as DNA base pairs. However, DFT can be applied to larger molecular systems having as many as a 1,000 atoms (Neese et al. 2009) and has been found suitable to handle

electron–molecule interaction processes (Schwabe and Grimme 2008; Zhao and Truhlar 2008; Shao et al. 2006). In this chapter, we include a brief summary of the experimental findings and provide a detailed description of the recent applications of theory toward the understanding of the mechanisms of action of LEE-induced DNA damage. The chapter includes discussions of (1) shape and core-excited resonances, (2) valence and dipole-bound states of DNA bases and their associated electron affinities, (3) the best choice of the basis sets to tackle the valence and dipole-bound radical anions, (4) effects of condensed phase on dipole-bound states and thermodynamic properties of the solvated electron, (5) single-strand breaks formation in the ground state of the radical anion, (6) importance of the excited states of radical anions (resonances) for LEE-induced DNA single-strand breaks formation, (7) LEE-induced base release, (8) the effect of solvation on the formation of LEE-induced strand breaks, and (9) LEE interaction with DNA radiosensitizers.

Proposed Mechanism of LEE-Induced DNA Strand Breaks

As briefly described in the introduction, low-energy electrons (LEE) are produced abundantly during high-radiation events and are increasingly being recognized as potent DNA-damaging agents. In fact, LEEs are found to be several times more damaging than photons that have similar energy. The pioneering series of experiments by Sanche and coworkers (Boudaïffa et al. 2000; Huels et al. 2003; Martin et al. 2004; Panajotovic et al. 2006; Brun et al. 2009; Li et al. 2008; Zheng et al. 2004; Abdoul-Carime et al. 2001), showed that LEEs in the 0–10 eV range produce single- and double-strand breaks in DNA. In their first report (Boudaïffa et al. 2000), the plasmid DNA (pGEM 3Zf(–)) was extracted from *E. coli*, hydrated to 2.5 water/base pair, and then irradiated with LEE beam with kinetic energies 3–20 eV. From these experiments, it was concluded that electrons having energy below the ionization limit of DNA (ca. 7.5–10 eV (Hush and Cheung 1975; Orlov et al. 1976; Yang et al. 2004; Colson et al. 1992)) were able to cause SSB and DSB. The corresponding yields of SSB and DSB were found to be 8.2×10^{-4} and 2×10^{-4} strand breaks per incident 10 eV electron. The yields of SSB and DSB in DNA depended on the energy of the interacting electron and have been suggested to be caused by the rapid fragmentation reactions of transient molecular resonances localized on DNA components, that is, base, sugar, and phosphate backbone (Boudaïffa et al. 2000). In later experiments, Sanche and coworkers (Martin et al. 2004; Panajotovic et al. 2006) found that 0–4.7 eV electrons can effectively produce SSB in plasmid DNA with similar yields to those produced at far higher energies. (See Fig. 2, where it is evident that the cross section for SSB formation at 1 eV is actually greater than that found at 10 eV (Panajotovic et al. 2006).) They also found that 0.1–4.5 eV electrons induce only SSB (Martin et al. 2004) (see Fig. 1 in Martin et al. (2004)). Apparent resonances for inducing SSB are found at 0.8 and 2.2 eV, characterizing the formation of transient negative ion (TNI) at these energies (Martin et al. 2004). Due to low energies, these resonances were interpreted as shape resonances (Martin et al. 2004). The SSB yields at 0.8

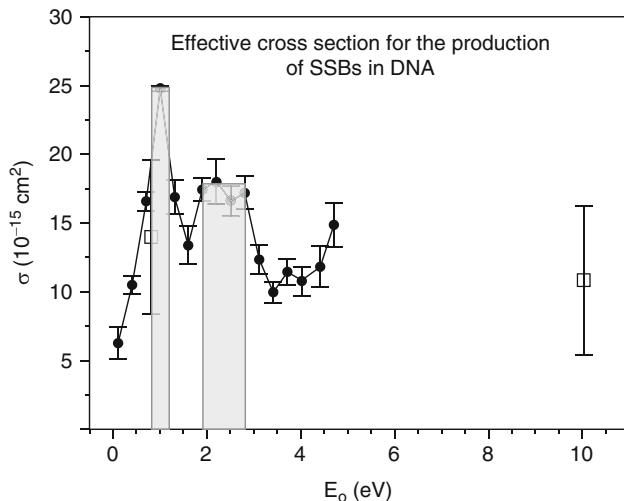


Fig. 2 Effective cross sections (σ) for the formation of SSB in plasmid DNA by 0.1–4.7 eV electrons. The peaks at 1 eV and around 2.5 eV (see *shaded portion*) correspond to shape resonances (Reprinted with permission from Panajotovic et al. (2006). © (2006) Radiation Research Society)

and 2.2 eV, found in Sanche's work, were compared with electron capture cross sections of the DNA bases which were assigned to π^* -shape resonances. The vertical attachment energies of DNA bases were taken from electron transmission spectroscopy (ETS) (Aflatooni et al. 1998), and their peak magnitudes were scaled to reflect the inverse energy dependence of the electron capture cross sections. The lowest peak present at 0.39 eV in the modeled cross section was shifted by 0.41 eV to match the SSB yield at 0.8 eV (Martin et al. 2004). A good agreement between the experimental and simulated SSB yield between 0 and 4 eV certainly suggests the involvement of shape resonances in creating LEE-induced single-strand breaks. Indeed, from the studies of strand breaks in plasmid DNA induced by 3–100 eV electrons (Huels et al. 2003; Martin et al. 2004; Panajotovic et al. 2006), it was concluded that below 5 eV LEE-induced SSB occurs through dissociative electron attachment (DEA) via shape resonances (Boudaïffa et al. 2000; Huels et al. 2003; Martin et al. 2004; Panajotovic et al. 2006; Brun et al. 2009), while between 5 and 15 eV, the core-excited resonances induce SSB and DSB similar to those produced at far higher energies such as 100 eV. Illenberger and coworkers (Baccarelli et al. 2007; Sulzer et al. 2006; Bald et al. 2006, 2008; König et al. 2006; Ptasińska et al. 2005a, b, 2006; Abdoul-Carime et al. 2004a, b; Hanel et al. 2003) have contributed to the understanding of these processes by showing that electron attachment directly to the phosphate moiety can lead to fragmentation (Bald et al. 2006). The involvement of shape resonances to create molecular fragmentation in DNA model compounds was also proposed in their work (Baccarelli et al. 2007; Sulzer et al. 2006; Bald et al. 2006, 2008; König et al. 2006; Ptasińska et al.

2005a, c, 2006). Several groups used scattering theory to help elucidate the nature of LEE interactions resulting in DNA damage (Caron and Sanche 2003, 2004, 2005, 2006; Caron et al. 2008; Tonzani and Greene 2006b; Dora et al. 2009; Gianturco et al. 2008; Winstead and McKoy 2006) through resonances. The modeling of shape resonances, created during electron–molecule interaction, using theory is discussed in section “[Resonance \(TNI\) Formation: A Molecular Orbital Approach.](#)”

Electron–Molecule Interaction Events

Electron–molecule interactions give rise to energetic species that result in physical and, subsequently, chemical changes in the parent molecule. The initial energy of the interacting electron is critical to the formation of such species. Understanding the physical and chemical changes induced requires an understanding of the underlying fundamental processes that govern the electron–molecule interaction. One of the most important of these interactions is dissociative electron attachment (DEA), which may result in DNA strand breaks discussed in section “[Proposed Mechanism of LEE-Induced DNA Strand Breaks.](#)” A scheme that depicts the various paths available during electron–molecule interaction is shown in Fig. 3. The collision of an electron with molecule involves two types of interactions: (a) direct scattering and (b) resonant interactions. In nonresonance electron–molecule scattering (direct scattering), the electron interacts with the molecule for a very short time (typically less than ca. 10^{-15} s), that is, approximately the transit time for an electron through

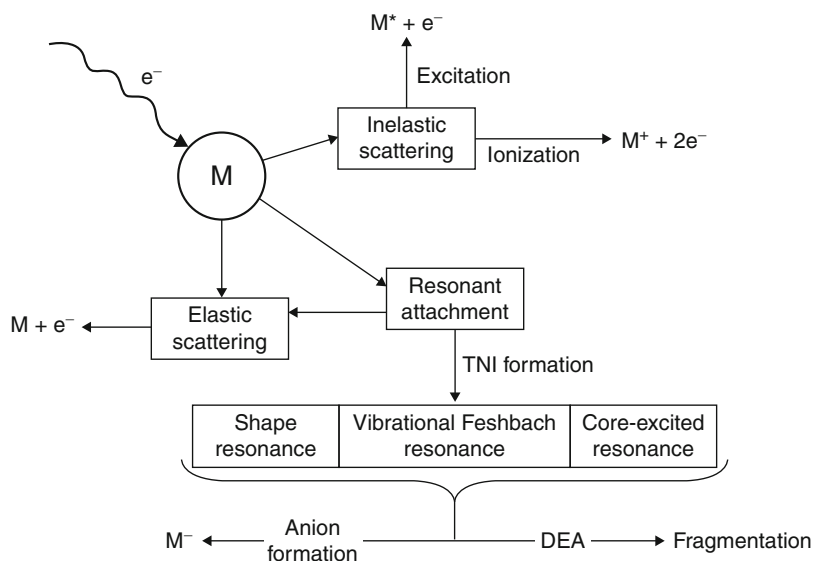


Fig. 3 Schematic diagram showing the scattering and resonant attachment events taking place during electron–molecule interactions

the dimensions of a molecular framework. A state with angular momentum such as a p-state can extend the time of the resonance. Depending on the energy transferred from electron to the molecule during collision, the direct scattering is known as (a) elastic, if no kinetic energy is transferred to the internal degrees of freedom (electronic, rotational, and vibrational) of the parent molecule, and (b) inelastic, if some part of the electron energy is transferred to the molecule and excites these internal degrees of freedom. In an elastic collision, only kinetic energy can be transferred through the conservation of momentum. However, no significant kinetic energy is transferred in an elastic collision owing to the small mass of the electron, which is negligible compared to the mass of the molecule. Inelastic scattering can result in ionization as well as vibrational and/or electronic excitation, as shown in Fig. 3.

In resonant interactions, the electron–molecule interaction lasts for a longer time ($\sim 10^{-12}$ to 10^{-15} s) and results in the formation of transient negative ions (TNI) of several types discussed in section “[Resonance \(TNI\) Formation: A Molecular Orbital Approach](#).” Once these resonances are formed, they may decay via (a) autodetachment, in which the electron is emitted from the negative ion $(M^-)^*$ resulting M^* and e^- ; (b) dissociative electron attachment (DEA), in which the negative ion $(M^-)^*$ stabilizes by dissociating one of its chemical bonds to form anion and neutral fragments; and (c) radiative cooling, which means the negative ion $(M^-)^*$ is stabilized by emission of one or more photons. This last process does not usually occur in the gas phase as the lifetimes are short owing to processes (a) and (b).

For the DEA process (b), the time needed to stretch a specific bond in its neutral potential energy surface (PES) so that it crosses the repulsive anion surface is usually less than the vibrational period of that bond. On the basis of the timescale, DEA and autodetachment are clearly in competition with each other. Such electron resonances are well described in the literature (Schulz 1973a, b; Lane 1980; Sanche 1995; Jordan and Burrow 1987; Simons and Jordan 1987; Simons 2008; Modelli and Martin 2002; Wetzel and Brauman 1987; Chen and Gallup 1990; McConkey et al. 2008; Hotop et al. 2003). Since the TNI formation can occur in the continuum, special care is needed to handle these problems in any theoretical treatment. Quantum chemical methods employing compact basis sets (Chen and Gallup 1990) have been used to study the nature and properties of TNIs/resonances by several groups such as Burrow (Jordan and Burrow 1987), Jordon (Jordan and Burrow 1987; Simons and Jordan 1987), Simons (Simons and Jordan 1987; Simons 2008), and Modelli (Modelli and Martin 2002). This method has been widely used and can give reasonable estimates of the energetics involved. However, the scattering phenomenon can only be approximated by such approaches. The present chapter elucidates the mechanism of DNA damage caused by resonant interaction events only.

Resonance (TNI) Formation: A Molecular Orbital Approach

From the above discussion, we see that electron–molecule interactions can lead to transient negative ion (TNI) formation and subsequently to a variety of chemical changes. On TNI formation, an extra electron

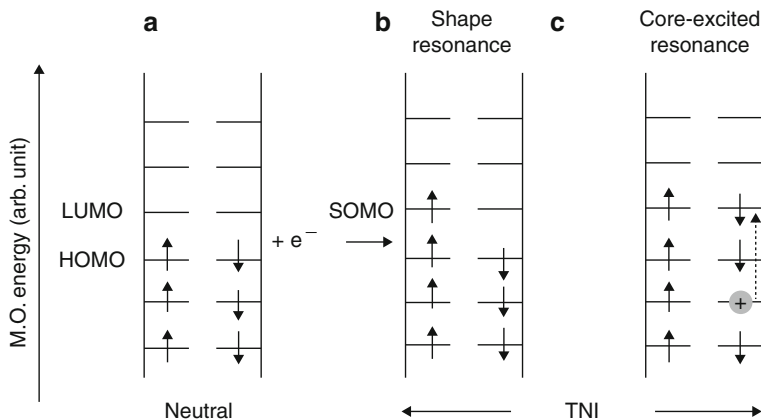


Fig. 4 Schematic diagram showing the electronic configuration of a neutral (a) and transient negative ion (TNI) (b, c). The interacting electron initially captures into the unoccupied MOs of the neutral molecule resulting in TNI formation via (a) shape resonance or (b) core-excited resonance. For a shape resonance, the electron can interact with any unoccupied MO. The SOMO was the empty LUMO before the LEE interaction. In core-excited resonance, on electron interaction an electronic transition takes place from an inner shell to the vacant MOs creating a “hole” (+ charge) in the inner shell, shown by an arrow (c). The up and down arrows show the occupancy of the molecular orbitals (MOs) with electrons of α and β spins. HOMO highest occupied molecular orbital, LUMO lowest unoccupied molecular orbital, SOMO singly occupied molecular orbital

is captured into the unoccupied molecular orbital (UMO) of the neutral molecule, and a shape or core-excited resonance results (see Fig. 4).

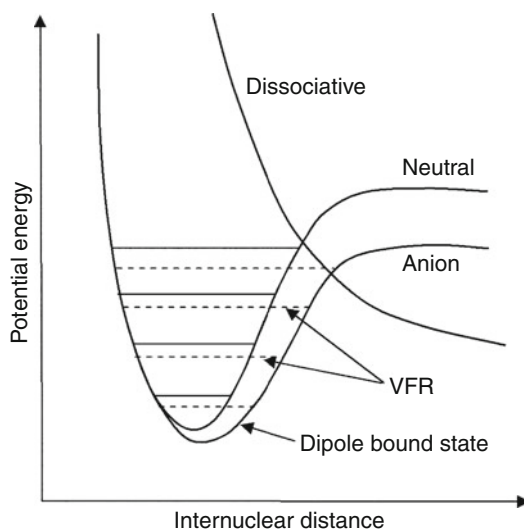
Shape resonances: In a shape resonance, an extra electron occupies a normally unoccupied (vacant) MO without perturbing the inner electronic configuration of the parent molecule. This results in the formation of a transient negative ion (TNI) (Schulz 1973a, b; Lane 1980; Sanche 1995) (see Fig. 4b). In a time-dependent picture, such TNIs (Schulz 1973a, b) are formed by an electron transition from the continuum ($e + M$) to a quasi-discrete state of the transient molecular anion (M^-)*. Because electron attachment is rapid with respect to nuclear motion, the negative ion (M^-)* is formed at the geometry of the neutral molecule. Shape resonances are also known as “single-” or “one-particle, no-hole” resonances. Generally, they occur at low energies (0–4 eV) and have lifetimes in the range 10^{-15} to 10^{-10} s. Longer lifetimes may occur if the attachment is followed by vibrational motion into portions of the nuclear potential surface lying below that of the neutral, where the electron cannot autodetach. Shape resonances may result in a dissociative electron attachment (DEA) process forming an anion fragment and a neutral fragment.

Core-excited resonances: A core-excited resonance is formed when an interacting electron excites one of the core electrons of the parent molecule, resulting in a hole in the inner filled MOs and two electrons, an extra electron and one from the core, in unoccupied MOs (see Fig. 4c). Such resonances are also known as “two-particle, one-hole” resonances. Core-excited resonances occur typically above 4 eV;

those lying below the “parent” molecular excited state are known as “Feshbach resonances” and may have lifetimes comparable to vibrational periods. If they lie above the parent, they are called “core-excited shape resonances” (Schulz 1973a) and generally have much shorter lifetimes because of decay into the parent state. Thus, core-excited shape resonances exhibit a negative electron affinity, relative to the excited parent, whereas Feshbach resonances have a small positive electron affinity (Schulz 1973a). Core-excited resonances are highly energetic and have been suggested to play a role in double-strand breaks in DNA (Huels et al. 2003; Martin et al. 2004; Panajotovic et al. 2006; Brun et al. 2009).

Vibrational Feshbach resonances (VFRs): In a vibrational Feshbach resonance, the interaction of a slow electron takes the form of a virtual excitation of a vibrational level of the neutral molecule with capture of the electron (Hotop et al. 2003; Dessent et al. 2000). For the zero point vibration, the maximum probability of interaction of the electron with parent molecule occurs at zero energy. If the dipole moment of the neutral molecule exceeds the critical value of approximately 2 Debye, the impinging electron may be trapped into the diffuse bound state, which provides a much longer timescale for the electron to stay near the molecule (Hotop et al. 2003; Dessent et al. 2000; Illenberger 1992), and VFRs may appear as shown in Fig. 5. Here, the PES of the anion lies close to, but below, the neutral surface, and the lowest vibrational levels of the dipole-bound anion are located just below the corresponding vibrational levels of the neutral. These long-lived VFRs appear as narrow features below the vibrational thresholds in elastic or vibrationally inelastic electron scattering cross sections (Hotop et al. 2003). VFRs may also decay through the dissociative electron attachment mechanism. For an excellent, detailed overview, see reviews by Schulz (1973a, b) and Hotop et al. (2003).

Fig. 5 Diabatic potential energy surfaces of neutral and dipole-bound anion showing the origin of the vibrational Feshbach resonances (VFRs). An extra electron attaching into the diffuse dipole-bound MO of the parent can excite the nearby vibrational levels of the parent molecule resulting in VFRs



An excellent example of the synergy of vibrational Feshbach resonances with dipole-bound anion states of uracil and thymine in DEA of N_1 -H bond has been reported by Burrow et al. (Burrow et al. 2006; Scheer et al. 2004). Sharp peaks (<3 eV) in the dissociative electron attachment cross sections of uracil and thymine were assigned as vibrational Feshbach resonances (Burrow et al. 2006; Scheer et al. 2004). Methylation at N_1 or N_3 site of these bases selectively eliminates the hydrogen atom (H) loss at the site of methyl substitution and thus allows for site assignments of DEA yield with energy of the LEE. At 1 eV, the H loss occurred from the N_1 site, while at 1.8 eV, the H loss occurred from the N_3 site (Ptasińska et al. 2005a), indicating the involvement of two distinct DEA pathways (Ptasińska et al. 2005a; Burrow et al. 2006) for dissociation. Using quantum chemical calculations and very diffuse basis sets 6-311G(d) augmented with four Gaussian sets of (s , p_x , p_y , p_z) orbitals with scale factors of 0.04, 0.004, 0.0004, and 0.00004, Burrow et al. (2006) showed that an excess electron is bound in the electric dipole potential field of the uracil molecule, resulting a dipole-bound anion state (DBS), which lies below the neutral state of the uracil (see Fig. 6). From Fig. 6 (left), we see that the DBS and the σ^* (N_1 -H) anion state can mix by symmetry, and LEE excitation of the N_1 -H bond vibrations can then lead to dissociation.

This results in H loss in uracil ($(U-H)^- + H^*$) (see Fig. 4 of Burrow et al. 2006). LEE resonances at 0.69 and 1.01 eV were identified with the second and third vibrational levels for the N_1 -H bond stretch (see Fig. 6).

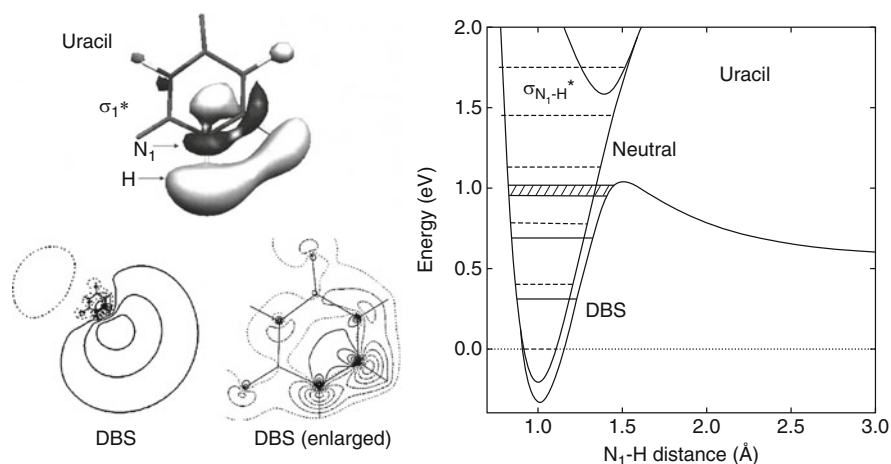


Fig. 6 (Left) The valence σ_1^* orbitals and expanded view of the dipole-bound anion state (DBS) of uracil on the molecular frame. (Right) Potential energies of the neutral, dipole-bound state, and valence σ^* anion state as a function of N_1 -H separation (Reprinted with permission from Burrow et al. (2006). © (2006) American Institute of Physics)

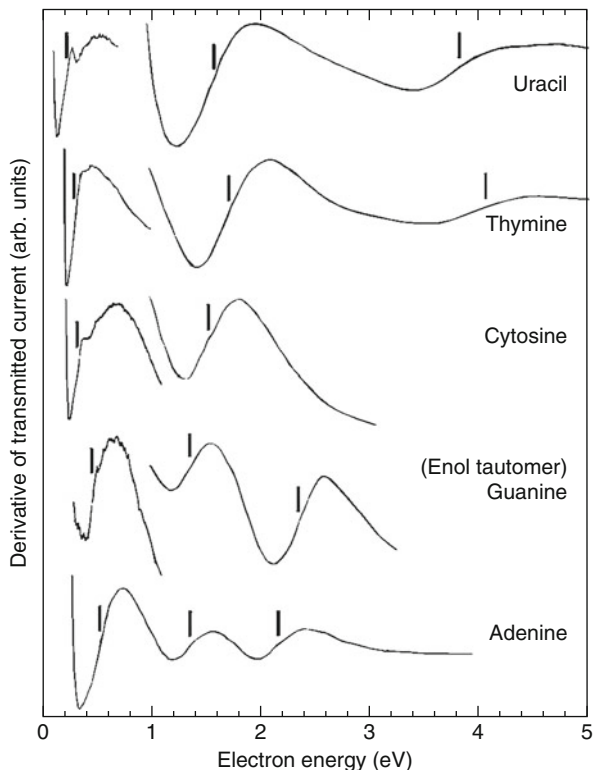
Shape Resonances of DNA Bases

As pointed out above, when LEEs are captured in the virtual (unfilled) orbitals of the parent (neutral) molecule, a transient negative ion formation results, known as a “shape resonance.” These resonances are generally unstable with respect to the autodetachment and therefore possess negative electron affinity, that is, the anion state lies above the neutral state of the molecule. In the gas phase, they are commonly studied by electron transmission spectroscopy (ETS). In ETS, a signal consisting of the derivative with respect to energy of the electron beam current transmitted through a gas cell is plotted as a function of electron kinetic energy. The energies of the resonances are termed as electron vertical attachment energies (VAEs) and are always found positive by ETS experiment (Aflatooni et al. 1998; Schulz 1973a, b; Jordan and Burrow 1987; Simons and Jordan 1987; Allan 1989). In this chapter, we will refer to such states as having negative vertical electron affinities. To our knowledge, there are no experimental available LEE scattering data for a complete DNA fragment (composing base, sugar, and phosphate). However, there are ETS results for the DNA/RNA bases, that is, adenine, cytosine, thymine, uracil, and guanine, reported by Aflatooni et al. (1998).

From ETS experiments (Aflatooni et al. 1998), Aflatooni et al. showed that energies required to attach an electron into the lowest empty valence molecular orbitals of all the bases are positive. This means that all the bases have negative valence vertical electron affinity (<0). The vertically formed anions of pyrimidines bases (C, T, and U) were found to be more stable than the purines (G and A) by ca. 0.2 eV. The ETS spectra of U, T, C, G, and A are shown in Fig. 7. The spectra in Fig. 7 arise due to the occupation of the lowest empty π^* MOs of the parent molecules by the electron. The vertical lines on the spectra (Fig. 7) show the position of the vertical attachment energies. For guanine anion, the ETS (shown in Fig. 7) was observed for its enol tautomer (for details see Aflatooni et al. (1998)). For each of these molecules, three VAEs, associated with the three lowest vacant π_1^* , π_2^* , and π_3^* MOs, were determined. The VAEs 0.22, 1.58, and 3.83 eV for uracil; 0.29, 1.71, and 4.05 eV for thymine; 0.32, 1.53, and 4.50 eV for cytosine; 0.54, 1.36, and 2.17 eV for adenine; and 0.46, 1.37, and 2.36 eV for guanine(enol) tautomer were determined by ETS.

The VAEs of bases were also predicted using theoretical methods such as semiempirical Pariser–Parr–Pople (PPP) (Compton et al. 1980; Younkin et al. 1976) and ab initio methods (Sevilla et al. 1995). Since these anion states are metastable with respect to the autodetachment of the electron, special care was taken in the aforementioned theoretical approaches to handle these anion states by employing confined basis set, Koopmans’ theorem approximation, and empirical determination of parameters to scale the theoretical results to the experimentally measured anion state energies for other compounds. Sevilla et al. (1995) used calculated values for VAEs of benzene, naphthalene, pyridine, pyrimidines, and uracil, which have experimentally known values to scale theoretically calculated VAEs for the DNA bases. A comparison of the experimental VAEs of the bases to those predicted

Fig. 7 ETS spectra of uracil, thymine, cytosine, guanine, and adenine. The experimentally determined π^* anion state energies are indicated by vertical lines (Reprinted with permission from Aflatooni et al. (1998). © (1998) American Chemical Society)



theoretically is presented in Table 1. From Table 1, we see that the calculated VAE values are in excellent agreement with experimental VAEs and the difference between theory and experiment lies in the range 0.03–0.2 eV. The PPP method only shows a qualitative agreement with the experiment. The resonance energies of DNA/RNA bases, including deoxynucleosides and deoxynucleotides, were studied using scattering theory (Winstead and McKoy 2008; Gianturco and Lucchese 2004; Tonzani and Greene 2006a; Yalunin and Leble 2007). The calculated values (Gianturco and Lucchese 2004; Tonzani and Greene 2006a; Yalunin and Leble 2007) were found to be several electronvolts more than the experimental ETS values (Aflatooni et al. 1998); however, the first π^* resonance energies reported by Winstead and McKoy (2008) were in reasonable agreement with experiment (Aflatooni et al. 1998) for uracil, thymine, and cytosine.

Recently, we studied the vertical attachment energies of 5'-thymidine monophosphate (5'-dTMP) using the density function theory (DFT) B3LYP/6-31G* method (Kumar and Sevilla 2007). A nucleotide such as 5'-dTMP represents the smallest fragment of DNA containing all three components: base, sugar, and phosphate. This study was carried out with a view to draw some fruitful insights about the nature and location of the “shape resonances” during transient negative ion (TNI) formation.

Table 1 Experimental (ETS) and theoretical vertical attachment energies (VAEs) in uracil and the DNA bases^{a,b}

Compound	Vertical attachment energies (eV)				
	Assignment	Theory			
		Exp. (ETS)	PPP	Sevilla et al. (1995)	6-31G*/3-21G
Uracil	π_1^*	0.22	0.48	0.19	0.216
Thymine	π_2^*	1.58	1.50	–	1.609
	π_3^*	3.83	3.35	–	5.009
	π_1^*	0.29	–	0.32	0.364
Cytosine	π_2^*	1.71	–	–	1.771
	π_3^*	4.05	–	–	4.057
	π_1^*	0.32	–	0.4	0.524
Adenine	π_2^*	1.53	–	–	1.918
	π_3^*	4.50	–	–	5.235
	π_1^*	0.54	0.82	0.74	0.794
Guanine (keto)	π_2^*	1.36	1.43	–	1.379
	π_3^*	2.17	2.39	–	2.405
	π_1^*	–	0.94	1.23	1.191
Guanine (enol)	π_2^*	–	1.64	–	1.441
	π_3^*	–	2.55	–	2.575
	π_1^*	0.46	–	–	0.908
	π_2^*	1.37	–	–	1.662
	π_3^*	2.36	–	–	2.597

^aSee Aflatooni et al. (1998)^bValues correspond to negative electron affinities

In this study, the five lowest unoccupied molecular orbitals (LUMOs) along with the highest occupied molecular orbital (HOMO) of the neutral 5'-dTMPH were considered. The B3LYP/6-31G* method predicted HOMO and the lowest two UMOs as π -type localizing on the thymine base, while the third, fourth, and fifth UMOs as σ -type localized on the phosphate (PO_4) group, the sugar moiety, and the sugar phosphate group, respectively (see Fig. 8). The B3LYP/6-31G* calculated orbital energies in eV of HOMO, and the five LUMOs are $-6.24(\pi)$, $-0.84(\pi_1^*)$, $0.43(\pi_2^*)$, $0.73(\sigma_1^*)$, $1.27(\sigma_2^*)$, and $1.78(\sigma_3^*)$, respectively (see Fig. 8). From ETS and the theoretical calculations discussed above, it is evident that within Koopmans' theorem approximation, the vertical attachment energies (VAEs) are equivalent to the virtual orbital energies (VOEs). But these VOEs calculated at the Hartree–Fock or DFT (B3LYP) levels are appreciably overestimated by several electronvolts and thus need scaling to appropriately represent the measured VAEs (Aflatooni et al. 1998; Sevilla et al. 1995). Using this procedure, Modelli (2003) proposed an equation for appropriate scaling of B3LYP results for π^* resonances. Using Modelli's equation (Modelli 2003), the scaled VOEs of corresponding B3LYP/6-31G* computed LUMOs of 5'-dTMPH are found to be $0.53(\pi_1^*)$, $1.56(\pi_2^*)$, $1.80(\sigma_1^*)$, $2.23(\sigma_2^*)$, and $2.64(\sigma_3^*)$ eV, respectively (Kumar and Sevilla 2007). The

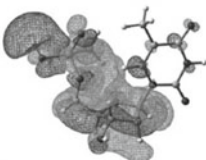
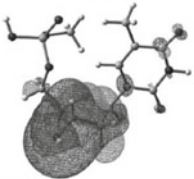
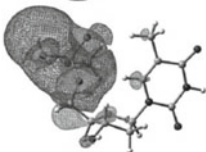
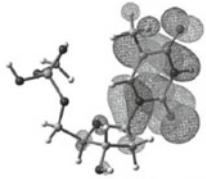
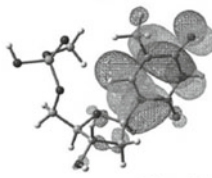
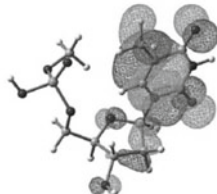
MOs		Orbital Energy (eV)	Scaled VOE (eV)
LUMO + 4 (σ_3^*)		1.78	2.64
LUMO + 3 (σ_2^*)		1.27	2.23
LUMO + 2 (σ_1^*)		0.73	1.80
LUMO + 1 (π_2^*)		0.43	1.56 (1.71)
LUMO (π_1^*)		-0.84	0.53 (0.29)
HOMO (π)		-6.24	

Fig. 8 Molecular orbital plots of neutral 5'-dTMPH, calculated using the B3LYP/6-31G* method. B3LYP/6-31G* calculated orbital energies along with scaled values are given in electronvolts. The experimental VOEs of thymine (Aflatooni et al. 1998) are given in electronvolts (in parentheses) (Reprinted with permission from Kumar and Sevilla (2007). © (2007) American Chemical Society)

calculated lowest two VAEs 0.53(π_1^*) and 1.56(π_2^*) are in reasonable agreement with those reported by Aflatooni et al. (1998) using EsTS of thymine, as shown in Table 1. This theoretical calculation predicts that for energies of less than 2 eV, there are UMOs on base, sugar, and phosphate available to capture the LEE. The theoretical procedure of scaling of orbital energies of neutral LUMOs gives good

Table 2 Semantics of electron molecular resonances^a

Resonance ^a		“Parent” ^a	Energy <i>vis-a-vis</i> parent ^a	Characteristics	
				Gas phase ^a	Condensed phase
Shape (1 particle, 0 hole)		Ground electronic state (GS)	Above parent (0–4 eV)	Yes	Yes
Core-excited (2 particles, 1 hole)	Feshbach	Mostly Rydberg ES	Below parent (ca. 0.5 eV)	Yes	No
	Shape	Valence ES	Above parent (0–5 eV)	Yes	Yes
Vibrational Feshbach ^b (VFR)		Vibrational level of GS	Below	Yes	No

^aTerminology adapted from Schulz (1973a). GS ground electronic state, ES excited electronic state

^bHotop et al. (2003)

estimates of the resonance energies; however, it has several drawbacks, for example, (a) during TNI formation, the UMOs of the neutral molecule can be perturbed and can reorder the UMOs energies so they are not as in the neutral molecule and (b) this procedure gives no information about core-excited resonances. In this context, the excited-state calculations for the TNI provide the only convenient means to rectify these drawbacks. This will be discussed in section “[Excited States of TNI \(Resonance Formation\)](#).”

A summary of the semantics of electron molecular resonances occurring in the gas phase has been described by Schulz (1973a) (see Table 2) with relevance to the condense phase.

Electron Attachment to DNA/RNA Bases in Gas Phase

The energy change on electron attachment to a molecule is described as the electron affinity of that molecule. Upon electron attachment to the neutral molecule (M), the geometry of the neutral molecule undergoes nuclear relaxation with adiabatic anion (M⁻) formation. Electron attachment to the neutral molecule without a nuclear relaxation gives vertical electron affinity (VEA), and on nuclear relaxation the overall energy change gives the adiabatic electron affinity (AEA). The energy needed to detach an electron from the adiabatic anion without relaxation of the resulting neutral molecule is termed as vertical detachment energy (VDE). Depending on the stability of the anion formation with respect to the neutral molecule, a molecule has positive or negative electron affinity (VEA and AEA) values (see Fig. 9). The various processes taking place during electron attachment to a molecule (M) are shown in Fig. 9.

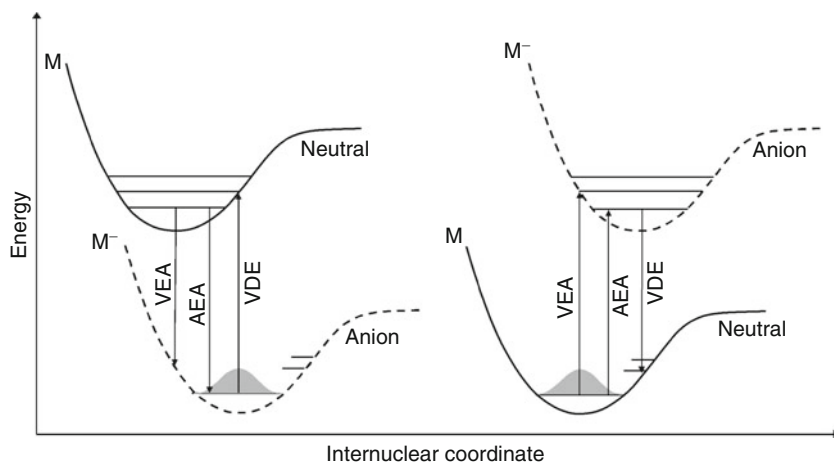


Fig. 9 Potential energy surfaces (PES) of an anion formation (M^-) after electron (e^-) attachment to the neutral molecule (M). When the energy of an anion (M^-) lies below the neutral molecule, positive electron affinity is observed (*left*); otherwise the molecule is said to have negative electron affinity (*right*). Energy changes shown for an electron attachment or detachment event are (1) the vertical electron affinity (VEA), (2) the adiabatic electron affinity (AEA), and (3) the vertical detachment energy (VDE). The VDE and VEA impose the upper and lower bounds to AEA

From theory (Compton et al. 1980; Younkin et al. 1976; Sevilla et al. 1995; Winstead and McKoy 2008; Gianturco and Lucchese 2004; Tonzani and Greene 2006a; Yalunin and Leble 2007; Kumar and Sevilla 2007) and experiment (Aflatooni et al. 1998), it is well established that all the DNA/RNA bases have negative vertical electron affinity (VEA) values in gas phase. The adiabatic electron affinities (AEA) of these bases were also studied using anion photoelectron spectroscopy by Bowen (Hendricks et al. 1996, 1998) and Weinkauff (Schiedt et al. 1998), and Rydberg electron transfer (RET) experiments by Schermann (Desfancois et al. 1996, 1998) and their coworkers. These experiments showed the formation of stable dipole-bound anion of uracil, thymine, and cytosine having AEA in the range of 0.06–0.1 eV (see Table 3). The experimentally estimated dipole-bound AEA of uracil ~ 93 meV and thymine ~ 68 meV were found to be in excellent agreement with the theoretically predicted dipole-bound AEA values 86 and 88 meV of the corresponding bases by Oyler and Adamowicz (1993, 1994) using SCF and MP2 methods in combination with a very diffuse basis set. In recent years, density functional theory (DFT) and CCSD(T) methods were also applied to evaluate the electron affinities (EAs) of the bases (Li et al. 2002a; Vera and Pierini 2004; Puiatti et al. 2009; Wesolowski et al. 2001; Bachorz et al. 2007). The DFT-calculated positive electron affinity values for the pyrimidines are overestimated by about 0.15 eV (see Table 3). From theory and experiment, it is now well established that VEAs of the bases are negative and the AEAs of U, T, and C are near zero and still negative for A and G. The best estimated theoretical electron affinities of bases along with their experimental values are presented in Table 3.

Table 3 Selected theoretical and experimental gas-phase electron affinities of DNA/RNA bases (eV)

Bases	VEA		AEA				
	Experiment	Theory	DB (meV) ^{a,b,c}		Valence		
			Experiment	Theory	Experiment	DFT	Ab initio
U	−0.22 ^d	−0.27 ^e	93 ± 7 ^b	86 ^f	ca. 0 ^g	0.20 ^h	0.002 ⁱ
	−0.30 ^j	–	86 ± 8 ^c				0.040 ^k
	–	–	–	–	–	–	0.024 ± 0.013 ^p
T	−0.29 ^d	−0.30 ^e	69 ± 7 ^b	88 ^l	ca. 0 ^g	0.15 ^h	0.018 ^m
	–	–	62 ± 8 ^c			–	–
C	−0.32 ^d	−0.55 ^e	85 ± 8 ^c	–		−0.05 ^h	−0.13 ⁱ
	−0.55 ^j	–				–	
A	−0.54 ^d	−0.74 ^e	12 ± 5 ^a	–	–	−0.35 ^h	–
	−0.45 ^j	–				–	–
G	(−0.74) ^d	−1.25 ⁿ	–	–	–	−0.75 ^h	−0.52 ^o

^aDipole-bound (DB) (Desfancois et al. 1996)

^bDipole-bound (DB) (Hendricks et al. 1996; Schiedt et al. 1998)

^cDipole-bound (DB) (Schiedt et al. 1998)

^dElectron transmission spectroscopy (ETS) results reported by Aflatooni et al. (1998). For G, VEA was estimated for keto tautomer from enol tautomer experimental value (−0.46 eV) plus the calculated difference in total energies between the two tautomers (0.28 eV)

^eB3LYP/6-311G(2df,p) values due to Vera and Pierini (2004) and Wetmore et al. (2000)

^fOyler and Adamowicz (1993)

^gEstimated from stable valence anion complexes, e.g., U(Ar)[−] (Hendricks et al. 1998)

^hBest estimates from DFT basis set dependence study by Li et al. (2002a). The DFT values are generally overestimated by 0.15 eV

ⁱCalculated using CBS-Q (Li et al. 2004)

^jRydberg electron transfer spectroscopy (RET) results due to Periquet et al. (2000)

^kBachorz et al. (2007)

^lOyler and Adamowicz (1994)

^mCCSD(T) correction to MP2 complete basis set (CBS) limit including ZPE correction. See Svozil et al. (2005)

ⁿBest estimates from B3LYP/D95V+(D) trends for other bases (Li et al. 2002a)

^oHaranczyk and Gutowski (2005)

^pGu et al. (2014b)

More recently, Schaefer and coworkers (Gu et al. 2014b) used the Weizmann Bruecker doubles composite method (W1BD) to benchmark the valence AEA of uracil. The W1BD method predicted a positive electron affinity of uracil and estimated its valence AEA to be 0.024 ± 0.013 eV (see Table 3). From their benchmark study (Gu et al. 2014b), these authors proposed that Hartree-Fock and MP2 methods are not suitable and severely underestimate the AEA of uracil, while the Gaussian-4 (G4) theory (Curtiss et al. 2007), CASPT2 and M06-2X (Zhao and Truhlar 2008) methods were found to be quite accurate, see Table 1 in Gu et al. (2014) and Table 2 in Gu et al. (2012).

Choice of the Basis Set

From the discussions in the preceding sections, it is clear from theory and experiment that vertical electron affinities of all the nucleic acid bases are negative, while the adiabatic electron affinities of pyrimidines are near zero. This means that an electron is unbound with the bases in the vertical state (TNI formation) and very loosely bound into the diffuse orbital of the bases in the adiabatic state, and in both the cases the choice of the basis set is crucial in determining the electron attachment process. The extent of the electron density localization with respect to the molecular frame for “dipole-bound” and “valence-bound” states can only be treated by an appropriate selection of the basis set. Generally, the basis set needed in a calculation consists of valence and extended (double- and triple- ζ) basis, which is further augmented with polarization and diffuse functions. While the polarization functions are important to study molecules with strained chemical bonds (Simons 2008), diffuse functions are necessary for dealing with anions or Rydberg type of species. In calculations for “dipole-bound” anions, the valence basis set must be augmented with very diffuse functions to successfully describe the system. Such basis sets were employed by Burrow et al. (2006) in the study of VFRs of thymine and uracil and by Oyler and Adamowicz (1993, 1994) for estimating the dipole-bound adiabatic binding energy of excess electrons to uracil and thymine. At times, the diffuseness of the basis functions is inadequate and the dipole-bound state mixes with the valence-bound state of the molecule (see Fig. 1 in Oyler and Adamowicz (1993)). If the energy of the dipole-bound and valence-bound states is equal, such a mixture could represent a proper solution for the molecular anion radical.

Diffuse functions are not a good choice for dealing with virtual “valence” states in TNIs or the available energies for shape resonances in the continuum. Any ab initio calculation employing the variational method with diffuse functions for a molecule with a negative electron affinity will fail and result in an unbound electron in the continuum at zero energy as the lowest energy state. Thus, the calculation of the negative electron affinity is not properly treated by a variational method. Nevertheless, virtual TNI states can be estimated with the use of compact basis sets such as 6-31G* as these compact basis sets avoid mixing with dipole-bound states and the continuum. Calculations using compact basis sets, when scaled, can provide good estimates of the virtual “valence” states in reasonable agreement with experiment. Further, these calculations can provide insights into the nature of such intermediate states and mechanisms of action of LEE in the induction of chemical change. The use of the compact basis set in the study of resonances (TNIs) is well documented in the literature (Jordan and Burrow 1987; Simons and Jordan 1987; Simons 2008; Modelli and Martin 2002; Staiey and Strnad 1994). Using compact 6-31G(d) to very diffuse 6-311++G(2d,p) basis sets, Li et al. (2002a) calculated the negative valence electron affinities of the bases. Their calculated SOMO surfaces of the DNA bases demonstrated that while the compact basis set represents the valence type π^* anions (the SOMO surfaces reside on the molecular frame) and gives good estimates of negative electron affinities, the use of extended basis sets results in

a collapse of the electron-binding energy to near zero, which is expected for the continuum and dipole-bound states. SOMO surfaces for extended basis sets show the characteristics of the dipole-bound state with the electron density residing away from the molecular frame. In the extreme of extension of the basis set, the electron would reside in the continuum, or if the molecule has a dipole, a weakly bound electron in a dipole-bound state.

Effect of Solvation (Condensed Media)

The discussion above describes the events of electron–molecule interaction in the gas phase. However, in condensed aqueous media, the nature of these processes is significantly altered. For example, dipole-bound states are not likely to be present as they are suppressed by the surrounding medium, and this is confirmed from the photoelectron spectroscopy of the hydrated uracil, thymine, cytosine, and adenine (Schiedt et al. 1998; Eustis et al. 2007) as the photoelectron spectra are blue-shifted with the increase of the number of hydrated water molecules and show valence-bound anion formation. For a visual inspection of this phenomenon, we plotted the LUMO and SOMO surfaces of guanine in neutral and anion radical states, using the B3LYP/6-311++G(2d,p) method in gas phase and in aqueous media using polarized continuum model (PCM). The LUMO and SOMO surfaces in the gas phase and in the solvated phase of guanine in neutral and in anion radical states are shown in Fig. 10. From Fig. 10, it is inferred that LUMO and SOMO represent a dipole-bound state (Li et al. 2002a) in the gas phase (see Fig. 10a, c), which is destabilized under the influence of the full solvation and becomes the valence-bound state (see Fig. 10b, d).

The solvent has a number of other effects on TNIs. First, relaxation of an excited TNI in the condensed media is rapid owing to the coupling with the high density of states in the matrix such as vibrational and phonon modes. Coupling to matrix energy levels quickly deactivates the excited vibrational states (within 1–10 ps). Further, the dielectric of the medium changes the energetics of the TNI first by the fast dielectric response, and as nuclear relaxation and reorientation takes place in the matrix, the full dielectric response ensues. The fast electronic response accounts for ca. 1 eV of stabilization and the slow dielectric response can account for several electronvolts. This stabilization can move the TNI (resonances) from the continuum into a valence-bound state. As a result, the autodissociation pathway is diminished or stopped altogether. In this context, it is worth considering the effect of secondary electrons generated by radiation damage to liquid water surrounding the DNA. It is found that on thermalization, the spatial extent of radiation-produced electrons in liquid water is in the range of 20–50 Å in what is considered the conduction band (Galduel et al. 1989; Crowell and Bartels 1996). In solution, the electron has several energy states (see Table 4), all lower than in the gas phase. The ground or the lowest energy state for the electron is the fully solvated electron (e_{aq}^-), which is stabilized

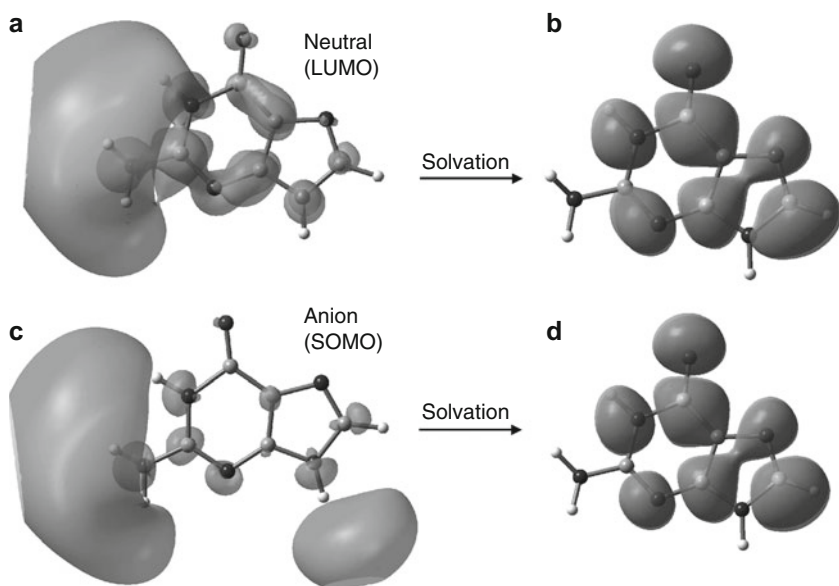


Fig. 10 B3LYP/6-311++G(2d,p)-calculated lowest unoccupied molecular orbital (LUMO) of neutral guanine in gas phase (**a**) and in solution (**b**). The singly occupied molecular orbital (SOMO) of guanine anion radical in gas phase (**c**) and in solution (**d**). The dipole-bound anion in (**a**, **c**) becomes a π^* -valence anion in (**b**, **d**). The effect of full solvation was considered through the use of polarized continuum model (PCM) considering the optimized gas-phase geometries

Table 4 Electron energies in gas phase and in liquid water

Electron type	Medium	Energy (eV)
LEE	Gas phase	0–20
Thermalized electron	Gas phase	0.026 ^a
LEE	Water	–1 to 20
Thermalized electrons	Water	
(–) VDE of conduction band electron (presolvated electron [e_{pre}^-])		–0.5 to –1
(–VDE) of p-state (excited state of solvated electron)		–1.5
(–) VDE of solvated electron		–3.3
Adiabatic energy of conduction band electron		Ca. 0.0
Adiabatic solvation energy (free energy of solvation)		–1.6

^aEnergy at 300 K

by ca. 1.6 electronvolts of solvation energy. The electron can be photoexcited to a p-state ($h\nu = 1.8$ eV) or to the conduction band by a ca. 2.5 eV photon or photoejected from water by a 3.3 eV photon. To properly account for the reaction of the solvated electron (e_{aq}^-) with DNA, knowledge about its thermodynamic properties is of utmost importance. These properties are discussed below.

Physical Observables of Solvated Electron (e_{aq}^-)

In 1962, Hart and Boag (1962) reported the absorption spectrum of e_{aq}^- using pulse radiolysis experiments. Owing to its fundamental importance to radiation chemistry (Young and Neumark 2012; Casey et al. 2013; Abel et al. 2012; Feng and Kevan 1980; Coe et al. 2008; Jordan 2004), the hydrated electron (e_{aq}^-) has been extensively studied using theory (Jordan 2004; Marsalek et al. 2012; Uhlig et al. 2012; Herbert and Head-Gordon 2005; Larsen et al. 2010; Jacobson and Herbert 2011; Turi and Madarász et al. 2011; Hammer et al. 2004; Madarász et al. 2009; Vysotskiy et al. 2012; Shkrob 2007; Bartels 2001; Bartels et al. 2005) and experiment (Coe et al. 1990; Coe 2001; Verlet et al. 2005; Siefermann et al. 2010; Siefermann and Abel 2011; Shreve et al. 2010; Tang et al. 2010; Kevan 1981) to elucidate its structure, dynamics, and binding energy. Considerable controversy exists in the literature regarding the energetics and structural aspects of its binding with bulk water, at water interface and in water clusters (Young and Neumark 2012; Casey et al. 2013; Abel et al. 2012; Feng and Kevan 1980; Coe et al. 2008; Jordan 2004; Marsalek et al. 2012). From electron spin resonance (ESR) data at 77 K in glasses, Kevan proposed a “cavity model” for binding the electron in water (Feng and Kevan 1980; Kevan 1981). In this model, the electron was surrounded by six water molecules arranged in an octahedral fashion and OH bond of each water points towards the electron at the center with *electron-to-hydrogen* ($e^- \rightarrow \text{HO}$) distance ca. 2.1 Å (Kevan 1981). Cavity size representing the radius of gyration (R_g) of the e_{aq}^- in the bulk water was computed as 2.5 Å by Bartels (Bartels 2001; Bartels et al. 2005) using moment analysis of the anionic water cluster spectra.

Coe et al. (1990, 2001, 2008) used photoelectron spectroscopy to measure the vertical detachment energy (VDE) of water anion clusters ranging from 2 to 69 in the gas phase. From the photoelectron spectra, they found that for clusters ranging from 11 to 69, the VDE varies linearly with $n^{-1/3}$ where n is the number of cluster. The extrapolation of this plot for $n = \infty$ (bulk water) had provided VDE = 3.3 eV, which was suggested to be the threshold for the solvated electron in bulk water.

Very recently, Siefermann et al. (2010) used ultrafast liquid micro-jet photoelectron spectroscopy and reported the VDE of e_{aq}^- in the water solution as = 3.3 eV. Using the time-resolved photoelectron spectroscopy, Tang et al. (2010) reported the VDE = 3.17 ± 0.1 eV in H_2O and 3.20 ± 0.1 eV in D_2O of e_{aq}^- . Neumark and coworkers (Shreve et al. 2010) measured the VDE = 3.6 ± 0.1 eV at 280 K using liquid micro-jet photoelectron spectroscopy, while Schultz and coworkers (Lübcke et al. 2010) reported the VDE = 3.4 eV of e_{aq}^- .

The thermodynamic properties of e_{aq}^- , such as the solvation free energy (ΔG_{hyd}^0), the enthalpy of hydration (ΔH_{hyd}^0), and entropy of hydration (ΔS_{hyd}^0), were first estimated by Jortner and Noyes (1966) assuming a cavity model for e_{aq}^- . Their estimated $\Delta G_{\text{hyd}}^0 = -37.5$ kcal/mol is in close agreement with experimental values (Han and Bartels 1990; Schwarz 1991; Hickel and Sehested 1985; Buxton et al. 1988); however, ΔH_{hyd}^0 and ΔS_{hyd}^0 were poorly estimated. Using Arrhenius parameters for the reaction $\text{H}^+ + \text{OH}^- \rightarrow e_{\text{aq}}^- + \text{H}_2\text{O}$, Bartels (Han and

Bartels 1990) reported $\Delta G^{\circ}_{\text{hyd}} = -38.5$ kcal/mol, $\Delta S^{\circ}_{\text{hyd}} = 28.2$ cal/mol deg. From pulse radiolysis, Schwarz (1991) estimated the $\Delta H^{\circ}_{\text{hyd}}$ and $\Delta S^{\circ}_{\text{hyd}}$ of e_{aq}^{-} as -34.0 kcal/mol and 16 cal/mol deg, respectively, giving $\Delta G^{\circ}_{\text{hyd}} = -38.8$ kcal/mol. Using pulse radiolysis, Hickel and Sehested (1985) and Buxton et al. (1988) estimated $\Delta G^{\circ}_{\text{hyd}} = -38.9$ kcal/mol, $\Delta H^{\circ}_{\text{hyd}} = -33.9$ kcal/mol, and $\Delta S^{\circ}_{\text{hyd}} = 11.7$ cal/mol deg. From the standard thermal constants of ions in aqueous solution, Ryabukhin (1977) estimated the $\Delta S^{\circ}_{\text{hyd}}$ of e_{aq}^{-} as 15.6 cal/mol deg. Zhan and Dixon (2003) used first-principles calculations including full solvent through dielectric continuum and calculated $\Delta G^{\circ}_{\text{hyd}} = -35.5$ kcal/mol for e_{aq}^{-} for anionic $(\text{H}_2\text{O})_8$. In addition to the thermodynamical quantities, the experimental redox potential (E° vs NHE) = -2.87 V of e_{aq}^{-} (see Table 5) is a quantity of fundamental interest to the understanding of the reaction of e_{aq}^{-} with DNA components. Very recently Kumar et al. (2015) model the solvated electron residing in a cavity formed

Table 5 Thermodynamic properties and physical observables of solvated electron (e_{aq}^{-})

ΔG° (kcal/mol)	ΔH° (kcal/mol)	ΔS° (cal/mol deg)	VDE (eV)	R_g (Å)	E° vs NHE (V)
-37.8^a	-29.9^a	26.6^a	-3.45^a	$2.4\text{--}2.6^a$	
-35.5^b	–	–	$3.3^k; 3.3^l;$ $3.17 \pm 0.1^m;$ $3.6 \pm 0.1^n; 3.4^o$	2.5^p	$-2.87^{q,r}$
-37.5^c					
-38.8^d	-34^d	16^d			
-38.5^e	-30.1^e	28.2^e			
$-38.9^{f,g}$	$-33.9^{f,g}$	$11.7^{f,g}$			
–	–	$15.6^{f,h}$			
-36.2^i					
$(-37.6)^j$	$(-32.0)^j$	$(19.6)^j$			

^aKumar et al. (2015)

^bZhan and Dixon (2003)

^cJortner and Noyes (1966)

^dSchwarz (1991)

^eHan and Bartels (1990)

^fBuxton et al. (1988)

^gHickel and Sehested (1985)

^hRyabukhin (1977)

ⁱThe E° vs NHE = -2.87 V for e_{aq}^{-} plus 4.44 V (NHE) gives $E^{\circ} = 1.57$ V or $\Delta G^{\circ} = -36.2$ kcal/mol

^jAverage of values in table

^kCoe et al. (1990)

^lSiefermann et al. (2010)

^mTang et al. (2010)

ⁿShreve et al. (2010)

^oLübcke et al. (2010)

^pBartels (2001)

^qSchwarz (1981)

^rWardman (1989)

by four water molecules in tetrahedral structure employing density functional theory with PCM. This simple model successfully described many of the properties of the solvated electron including the thermodynamic properties, see Table 5. In Table 5, we summarized the physical observables of the solvated electron.

The values in Table 5 suggest that the $\Delta G^{\circ}_{\text{hyd}}$ of e_{aq}^- lies in the range -35 kcal/mol to -38 kcal/mol, $\Delta H^{\circ}_{\text{hyd}}$ in the range -29 kcal/mol to -34 kcal/mol, and $\Delta S^{\circ}_{\text{hyd}}$ in the range between 12 and 28 cal/mol deg. The positive entropy indicates that solvated electron is a structure breaker. The VDE lies in a small range of 3.3 eV to 3.6 eV. This experimental VDE value is also supported by the recent molecular dynamics simulations of large anionic water clusters ($32\text{H}_2\text{O}$ treated by quantum mechanics (QM) and $992\text{H}_2\text{O}$ treated by molecular mechanics (MM)) from the group of Jungwirth (Uhlig et al. 2012).

Recently, Siefermann and Abel (Siefermann et al. 2010; Siefermann and Abel 2011) misinterpreted the redox properties of e_{aq}^- based on its VDE = 3.3 eV. Arguing from this value of the VDE (3.3 eV), they suggested that the hydrated electron (e_{aq}^-) should not react with DNA, see Fig. 5 in Siefermann et al. 2010 and Fig. 4 in Siefermann and Abel 2011. Their interpretation is incorrect because the VDE is clearly not the energy of the thermally equilibrated system. The redox potential of a aqueous electron depends on the AEA ($-\text{AEA} = \Delta G^{\circ} = -1.52$ eV to -1.65 eV, see Table 5), which is derived from the thermodynamically equilibrated system in question. The known E° vs NHE is -2.87 V for e_{aq}^- (Schwarz 1981; Wardman 1989) when added to the absolute potential for the NHE of 4.44 V gives absolute $E^{\circ} = 1.57$ V or $\Delta G^{\circ} = -36.2$ kcal/mol, which agrees well with the values in Table 5. The redox potential (E° vs NHE) for one electron reduction of DNA/RNA bases (G (<-2.76 V), C (-2.35), A (-2.52), T (-2.18), and U (-2.07 V)) was measured in DMF using pulse polarography and cyclic voltammetry by Seidel et al. (1996). From the known redox potential (E° vs NHE) values of the e_{aq}^- (-2.87 V) and of the of DNA/RNA bases, it is clear that e_{aq}^- can efficiently reduce all the DNA/RNA bases and the reaction will be exergonic in nature as shown in Fig. 11.

The energetics of the electron in the gas phase and in solution is shown in Fig. 11. Figure 11 also shows the TNI of DNA in the gas phase or in solution. Once formed in solution, the TNI will undergo solvent and molecular reorganization, which liberates additional solvation energy. This occurs within 100 ps and results in stabilization by several electronvolts. For a stable (adiabatic) anion radical to be formed, the energy must be below the energy of the solvated electron (e_{aq}^-), as shown in Fig. 11. The adiabatically solvated DNA radical anion ($\text{DNA}^{\bullet-}$) can subsequently undergo protonation reactions on pyrimidines for further stabilization, but will not produce a strand break. The reaction of the TNI to form a DNA strand break must take place before solvation, which forms the stable anion radical. This allows only a short time window for the reaction to take place. In Fig. 12, the energetic cycle for the adiabatic electron solvation (AEA), photo ejection of the solvated electron (e_{aq}^-) into the gas phase (VDE), and the solvent reorganization to equilibrated water (SRE), based on values in Table 5, are shown.

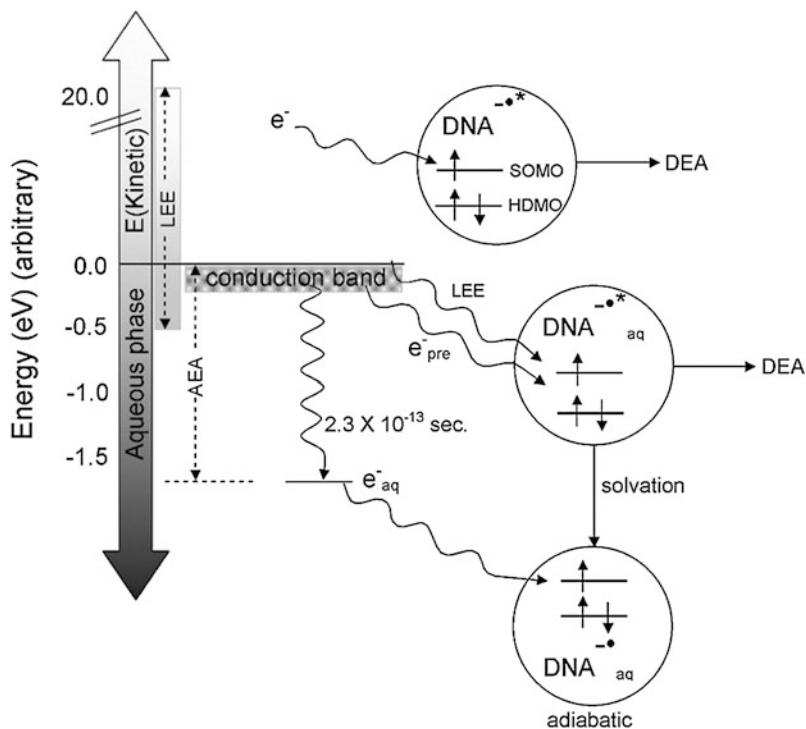


Fig. 11 Schematic diagram showing the addition of LEE to DNA in gas phase (e_{gas}^-) and in liquid water via prehydrated electrons (e_{pre}^-) from the conduction band and by LEE addition. In the gas phase, LEE (e_{gas}^-) can be captured into one of the UMOS of DNA (shown as SOMO after the (e_{gas}^-) capture) creating the TNI ($\text{DNA}_{\text{gas}}^{-.*}$). A TNI formed in the liquid water ($\text{DNA}_{\text{aq}}^{-.*}$) is quickly solvated, resulting in the adiabatic DNA radical anion ($\text{DNA}_{\text{aq}}^{-.*}$). The energy of the adiabatic radical anion must be below the energy of the solvated electron (e_{aq}^-) to be stable to electron loss to water. For clarity, the MOs of DNA below the HDMO (highest doubly occupied MOs) and above the SOMO are not shown in the figure (Figure based in part from Wang et al. (2009))

Because of the importance of solvation to the reactivity of electrons, studies of electrons in water at various energy states are needed (Alizadeh and Sanche 2012; Kheir et al. 2013; Petrovici et al. 2014; Adhikary et al. 2010). A few recent studies (Wang et al. 2009; Orlando et al. 2008) have appeared. The first (Wang et al. 2009) reported the plasmid DNA damage by the prehydrated electron (e_{pre}^-) (energy of the conduction band in Fig. 11). The electrons were generated by water excitation using two UV photon laser pulses. The TNI formation was reported to result in fast DEA processes, resulting in strand breaks. This work is the first to

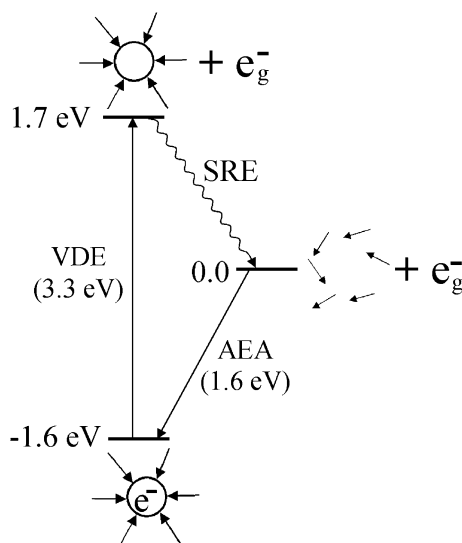


Fig. 12 Schematic diagram showing the energetic cycle for the adiabatic electron solvation (AEA), the photoejection of the solvated electron (e^-_{aq}) into the gas phase (VDE), and the solvent reorganization to equilibrated water (SRE). The various energy changes are designated as VDE (vertical detachment energy), AEA (adiabatic electron attachment), and SRE (Solvent reorganization energy). The VDE has the same geometry for the upper state as in the ground state. The solvent after photoejection of the electron is unstable by ca. 1.7 eV over a stable water-phase geometry. The difference between the AEA and VDE gives the value of the SRE

suggest conduction band electrons (e^-_{pre}) would be reactive in this manner. The results appear problematic because the concentrations of the solutes may not have been high enough to scavenge sufficient extremely short lived (0.3 ps) prehydrated electrons. Confirmation experiments are needed to verify this very interesting result. LEE-induced damage in hydrated A- and B-DNA has also been studied by Orlando et al. (2008) using elastic scattering of 5–30 eV electrons. In this study, a featureless amplitude buildup of elastic scattered electrons on the sugar and phosphate groups was observed for all energy ranges from 5 to 30 eV.

Proposed Theoretical Models of DNA Damage

After the discovery that low-energy (3–20 eV) electrons induced single- and double-strand breaks in plasmid DNA (Boudaïffa et al. 2000), a number of theoretical studies have attempted to elucidate the mechanism of LEE induction of DNA strand breaks. As the calculation of the entire DNA molecule at the ab initio or DFT levels is currently prohibitive, fragments of DNA structure usually including the base, sugar, and phosphate attached at 3'- and 5'-ends of the sugar ring have been treated.

The first mechanism for LEE-induced SSBs was proposed by Simons and coworkers employing a Hartree–Fock (UHF) level of theory and 6-31+G* basis set

(Barrios et al. 2002). In their study, they considered 3'-cytidine mono phosphate (3'-dCMPH) as a model of DNA and studied the ground-state potential energy surfaces (PESs) of the $C_{3'}-O_{3'}$ bond dissociation in the neutral and in radical anion states, and at each $C_{3'}-O_{3'}$ bond increment, the geometries were fully optimized. The calculations suggest a shape resonance, that is, the initial attachment of the excess electron on the π^* virtual molecular orbital localized on the cytosine base. During the $C_{3'}-O_{3'}$ bond elongation process on the PES of the anion radical, the excess electron, captured on the base, was found to be transferred to the $C_{3'}-O_{3'}$ bond region (joining the sugar and phosphate groups) at the transition state. From the calculated energy profile of $C_{3'}-O_{3'}$ bond dissociation, the energy barrier was estimated to be ca. 13 kcal/mol and the fragmentation process was found to be exothermic in nature. Based on this study, the authors proposed an *indirect mechanism* of SSB in which an electron transfers from base (π^*) to break the remote (σ^*) C–O bond. Here, we note that the statement “indirect mechanism” is not related to the terms direct and indirect effect of radiation-induced DNA damage. This first theoretical calculation clearly predicted that SSBs can occur below 3 eV as later found in LEE experiments (Martin et al. 2004; Panajotovic et al. 2006). While groundbreaking in nature, this study had several drawbacks: (a) it did not consider the higher energy shape resonances; (b) it considered the adiabatic state of the radical anion for the bond fragmentation, not the TNI, which is the initial state of the shape resonance formation as discussed in the section “[Shape Resonances of DNA Bases](#)”; and (c) it was not clear whether the proposed mechanism was pertinent to the actual DNA environment, where the bases are stacked over each other and the electron transfer process between the bases is likely preferred over the electron transfer process to the sugar phosphate region (Wagenknecht 2005).

Just after Simons's work (Barrios et al. 2002), Li, Sevilla, and Sanche (Li et al. 2003a) proposed that the direct attachment of the excess electron to the sugar phosphate backbone of the DNA may create the SSB. To investigate this process, they modeled the sugar–phosphate–sugar (S-P-S) backbone of the DNA without bases attached to the sugar ring (see Fig. 13). Because of the large size of the model, they used the ONIOM approach to optimize the geometry of the S-P-S model shown in Fig. 13. In the ONIOM treatment, B3LYP/6-31+G(d) level of theory was used for the critical bond cleavage atoms (higher layer), and the semiempirical AM1 method was used for the rest of the atoms (lower layer). In their treatment, they calculated the PESs of the $C_{3'}-O_{3'}$ and $C_{5'}-O_{5'}$ bonds dissociation of the S-P-S model in their neutral and anionic radical states. The study showed that in the radical anion state, the 3'- and 5'-C–O bonds cleavage occurred with a low activation barrier (ca. 10 kcal/mol) in comparison with the neutral state. The bond dissociation process for radical anion was found to be highly favorable thermodynamically (see Fig. 14). This work suggested that a vibrational excitation of the C–O bond could capture an electron and lead to a strand break. An important aspect about the nature of an excess electron association into the sugar phosphate region showed the formation of the dipole-bound anion state rather than the valence anion state, which is expected as only electronic excited states are expected to be available for electron capture in the sugar phosphate structure.

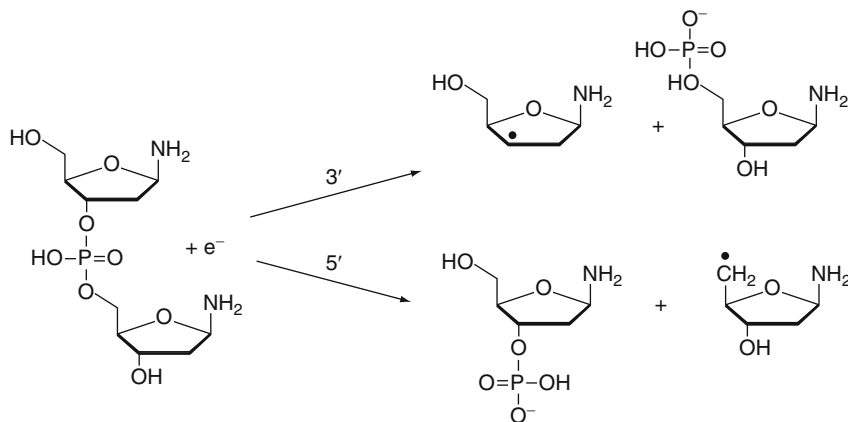


Fig. 13 Sugar–phosphate–sugar (S-P-S) model representing a section of DNA backbone. Electron-induced bond dissociations at the 3'- and 5'-ends of the model are investigated. Figure modified according to the spin density distribution shown in Li and Sevilla (2007), which showed no valence state but a dipole-bound precursor (Reprinted with permission from Li et al. (2003a). © (2003) American Chemical Society)

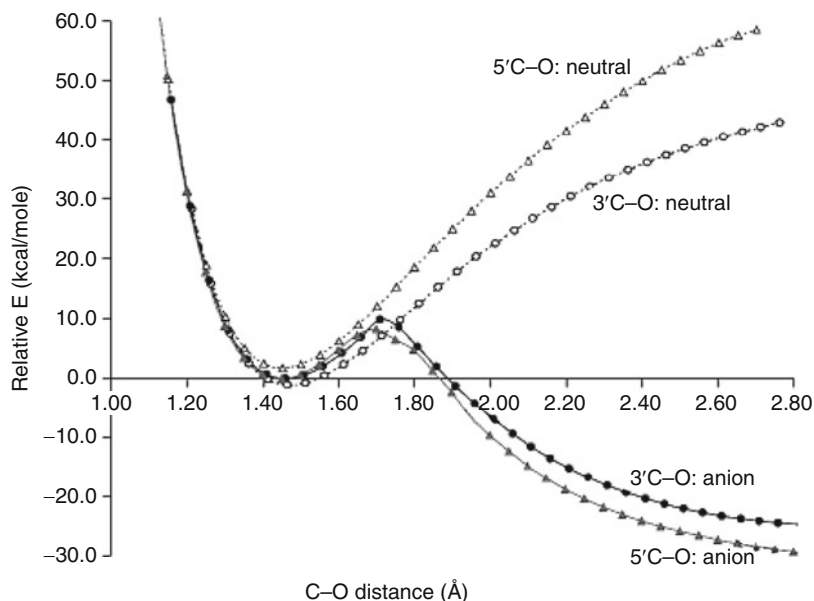
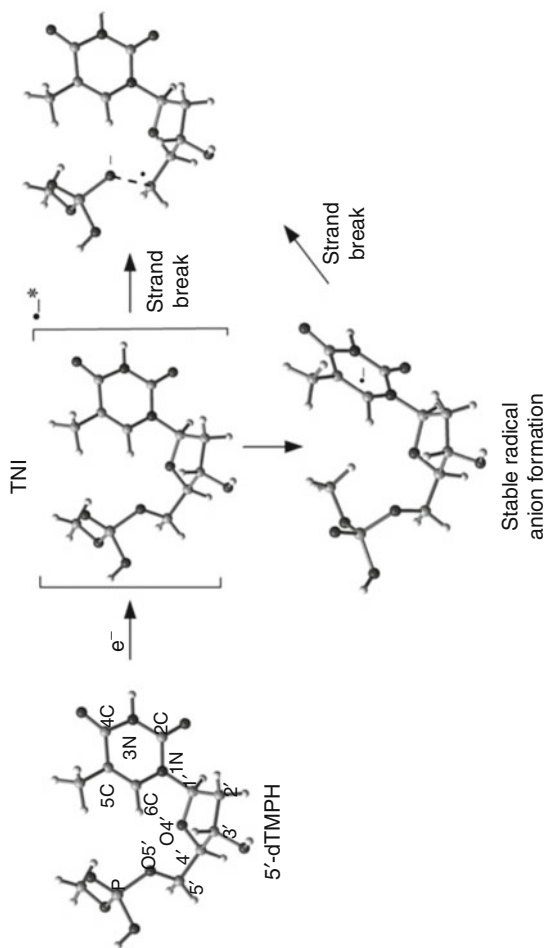


Fig. 14 Adiabatic potential energy surfaces for 3'C-O and 5'C-O bond rupture before and after electron addition to the S-P-S model. All energies are relative to the energies of the anions at equilibrium (Reprinted with permission from Li et al. (2003a). © (2003) American Chemical Society)

In a recent work (Kumar and Sevilla 2007), the interaction of LEE with a model for DNA (5'-dTMPH) was studied at the B3LYP level of theory considering both the vertical (TNI) and the adiabatic states of the radical anion. Since the TNI (shape resonance) formation has only a short life time of 10^{-15} to 10^{-10} s, the study was performed with the hypothesis that SSBs may occur before as well as after molecular relaxation, as shown in Scheme 1.

In this calculation, the potential energy surface (PES) for $C_{5'}-O_{5'}$ bond dissociation of 5'-dTMPH in vertical and adiabatic states was calculated using both compact (6-31G*) and diffuse (6-31++G**) basis sets. At each step on the PES, the singly occupied molecular orbital (SOMO) was plotted to observe the pattern of the excess electron localization on the molecular frame. To elucidate the mechanism of SSB, the PES was scanned by stretching the $C_{5'}-O_{5'}$ bond from the equilibrium bond length of the neutral and anion radical of 5'-dTMPH to 2Å in the increments of 0.1Å. On the vertical radical anion (TNI) PES, the $C_{5'}-O_{5'}$ bond was scanned by maintaining the remainder of the geometry in the neutral state; however, on the adiabatic state, the geometry was optimized at each step of the bond elongation on the PES. The B3LYP/6-31G* and B3LYP/6-31++G** calculated PESs of $C_{5'}-O_{5'}$ bond dissociation are shown in Figs. 15 and 16, respectively. On the adiabatic surface, both methods gave similar barrier heights, that is, 14.8 and 13.5 kcal/mol for the $C_{5'}-O_{5'}$ bond dissociation. The SOMO, which is initially localized on the thymine as expected (Becker et al. 2010; Li and Sevilla 2007; Kumar and Sevilla 2008b), transfers to the $C_{5'}-O_{5'}$ bond region only after the transition state for the $C_{5'}-O_{5'}$ bond rupture. For the vertical PES, B3LYP/6-31G* theory predicted the barrier height for $C_{5'}-O_{5'}$ bond dissociation as 9 kcal/mol, while the corresponding barrier was predicted to be 17 kcal/mol at the B3LYP/6-31++G** level of theory. An important feature of this study is that in the vertical state excess electron begins transferring smoothly into the $C_{5'}-O_{5'}$ bond region before the TS during bond elongation process; however, in the adiabatic state electron transfer takes place from base to $C_{5'}-O_{5'}$ bond region after the TS (see SOMO plots in 15 and 16). The facile electron transfer from base to sugar in the vertical state supports the hypothesis that the transiently bound electron in the TNI (shape resonance) results in SSBs in DNA without the need for significant molecular relaxation. From this work, it was also inferred that TNI formation may excite some specific vibrational mode which has the lower barrier for the bond fragmentation, as already proposed by Burrow et al. (Afatooni et al. 1998; Jordan and Burrow 1987) and Sanche et al. (Levesque et al. 2003). It is also noted that on timescale appropriate for the transition state ($<10^{-12}$ s) (Zewail 2000), specific vibrational motions may dominate. Therefore, on TNI formation a vibrational excitation of the $C_{5'}-O_{5'}$ bond would need only a small or negligible barrier for the bond dissociation. In this model, the incoming LEE must vibrationally excite the $C_{5'}-O_{5'}$ bond during the TNI (shape resonance) at thymine base. Since this mechanism seems improbable, alternatively, the LEE may directly attach into the phosphate group in an UMO and create such a vibrational excitation through resonance formation on phosphate. From the computed vertical attachment energies of 5'-dTMPH, shown in Fig. 8, it is evident that <2 eV UMOs on phosphate group and base are available to capture



Scheme. 1 Proposed mechanism of single-strand break (SSB) due to attachment of LEE with 5'-dTMPH molecule in vertical and adiabatic states (Reprinted with permission from Kumar and Sevilla (2007), © (2007) American Chemical Society)

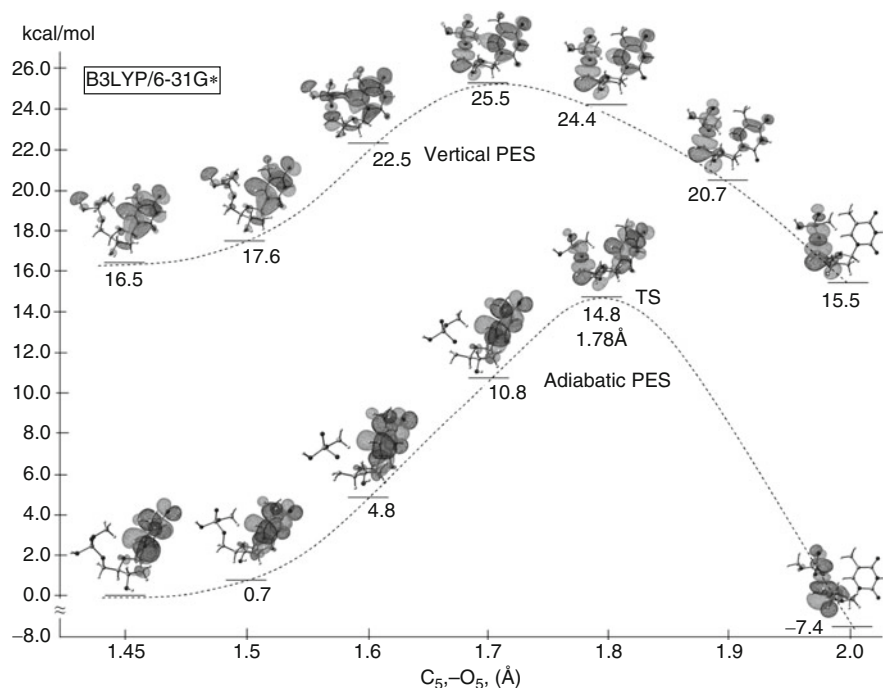


Fig. 15 B3LYP/6-31G* calculated adiabatic and vertical potential energy surfaces (PESs) of $C_{5'}$ - $O_{5'}$ bond dissociation of 5'-dTMPH radical anion. Energies and distances are given in kcal/mol and angstroms ($^{\circ}$ Å), respectively. The singly occupied molecular orbital (SOMO) is also shown (Reprinted with permission from Kumar and Sevilla (2007). © (2007) American Chemical Society)

LEEs (Kumar and Sevilla 2007). Thus, it is quite possible that on TNI formation these UMOs will excite the vibrational mode and result in SSBs.

Excited States of TNI (Resonance Formation)

The theoretical model (Kumar and Sevilla 2007) of LEE attachment to the LUMO leading to strand break formation provides one pathway that TNI formation can lead to DNA strand breaks, but it ignores the role of shape resonances that are available at 1–2 eV and above (estimated from the vertical attachment energies of the UMOs of 5'-dTMPH, see Fig. 8 and Kumar and Sevilla (2007)). It is well known that during electron–molecule interactions, a variety of excited states can be involved. Therefore, the study of excited states of the TNI, which model resonance states formed on the electron attachment to the parent molecule, can be pursued to better understand LEE interactions in DNA. With this in mind, we studied the excited states of anion radical of 5'-dTMPH (as a model of DNA) using time-dependent (TD) density functional theory (Kumar and Sevilla 2008b). The geometries of 5'-dTMPH in

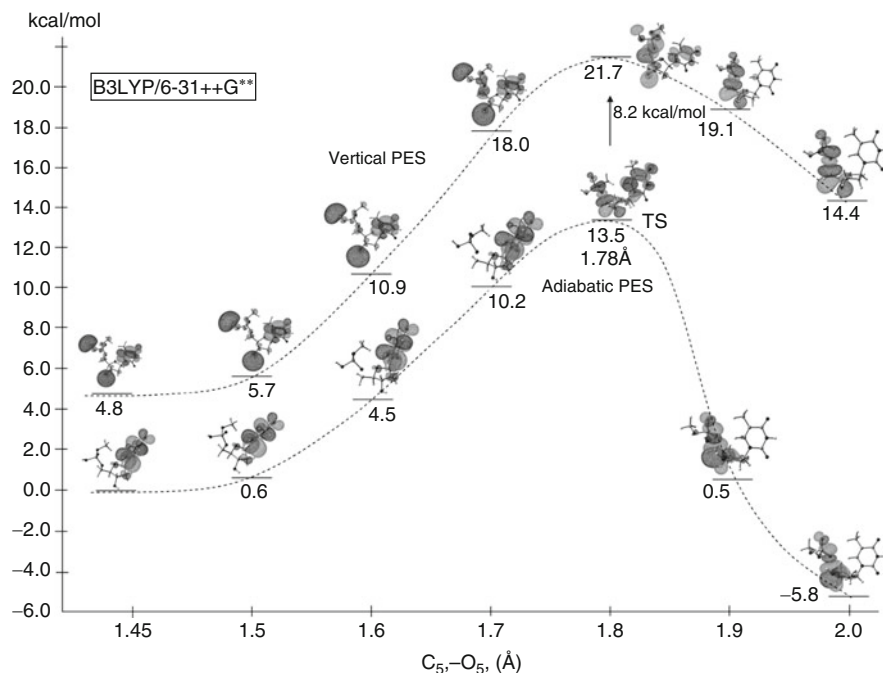


Fig. 16 B3LYP/6-31++G** calculated adiabatic and vertical potential energy surfaces (PESs) of $C_5'-O_5'$ bond dissociation of 5'-dTMPH radical anion. Energies and distances are given in kcal/mol and angstroms ($^{\circ}A$), respectively. The singly occupied molecular orbital (SOMO) is also shown (Reprinted with permission from Kumar and Sevilla (2007). © (2007) American Chemical Society)

their neutral and anion radical state were fully optimized using the B3LYP/6-31G* method. The lowest excited states of the TNI were calculated using TD-B3LYP and TD-BH and HLYP methods using the 6-31G* basis set. In this study, a compact basis set (6-31G*) was employed to avoid the mixing of the valence states with the dipole-bound states and continuum (Li et al. 2002a). The assumption of this work is that the TNI (resonance) formation is equivalent to an excited state of the electron adduct of the parent molecule. Further, the applicability of both functionals was tested for a number of anion radicals for which experimental and theoretical values were available (for details see the supporting information of Kumar and Sevilla (2008b)). Also, the transition energies, calculated using both the methods, were compared with the available shape resonance energies of adenine, thymine, cytosine, and uracil, as determined from the ETS experiment (Aflatooni et al. 1998), and 5'-dTMPH with their scaled vertical attachment energies (see Table 6). From Table 6, we see that both the methods provide reasonable estimates of shape resonance energies of the bases. However, for 5'-dTMPH radical anion, the B3LYP method was found to largely underestimate the transition energies, while the BH and HLYP method was found to give more realistic values. The problem with the B3LYP

Table 6 Vertical excitation energies (ΔE , eV) of transient negative ion (TNI) of DNA/RNA bases calculated using TD-B3LYP/6-31G* and TD-BH and HLYP/6-31G* methods and their comparison with available experimental values^a

Transition	Molecule	ΔE		
		B3LYP	BH and HLYP	Exp ^{b,c}
–	Uracil	–	–	0.22 (π_1^*)
$\pi \rightarrow \pi^*$		1.33	1.85	1.58 (π_2^*)
$\pi \rightarrow \pi^*$		4.27	4.73	3.83 (π_3^*)
–	Thymine	–	–	0.29 (π_1^*)
$\pi \rightarrow \pi^*$		1.38	1.89	1.71 (π_2^*)
$\pi \rightarrow \pi^*$		3.86	4.46	4.05 (π_3^*)
–	Cytosine	–	–	0.32 (π_1^*)
$\pi \rightarrow \pi^*$		1.55	1.91	1.53 (π_2^*)
$\pi \rightarrow \pi^*$		4.47	5.06	4.50 (π_3^*)
–	Adenine	–	–	0.54 (π_1^*)
$\pi \rightarrow \pi^*$		0.88	1.0	1.36 (π_2^*)
$\pi \rightarrow \pi^*$		1.89	1.86	2.17 (π_3^*)
–	5'-dTMPH ^d	–	–	(0.53 T) ^d
$\pi \rightarrow \pi^*$		1.16	1.68	(1.56 T) ^d
$\pi \rightarrow \sigma^{*e}$		0.18	1.42	(1.80 PO ₄) ^d
$\pi \rightarrow \sigma^{*e}$		0.70	2.06	(2.23 S) ^d

^aTransition energies of radical anions were calculated at the optimized neutral geometry of the molecules

^bEnergies of the shape resonances in the electron transmission spectroscopy (ETS) experiment (Afatooni et al. 1998)

^c π_1^* corresponds to the energy of the singly occupied molecular orbital (SOMO) and its difference with π_2^* and π_3^* orbital energies gives the estimate of the transition energies

^dScaled B3LYP orbital energies (VOE) (Kumar and Sevilla 2007)

^eElectron transfers from thymine (π) to PO₄ and sugar (σ) part of 5'-dTMPH (Reprinted with permission from Kumar and Sevilla (2008b)). © (2008) American Chemical Society)

method in underestimating the long-range Coulomb interactions in charge transfer states is well documented in the literature, and this arises from the small contribution of the Hartree–Fock exchange (20 %) in the B3LYP functional (Dreuw and Head-Gordon 2005). However, for this case, the charge transfer state is not created during the excitation because only the excess electron transfers between the nominally neutral portions of the molecule; this is another limitation of the B3LYP functional. In view of this, we employed the BHandHLYP/6-31G* method to compute the excitation energies of 5'-dTMPH radical anion in the vertical and adiabatic states.

In this study, the ground- and excited-state potential energy surfaces were generated by scanning the C_{5'}–O_{5'} bond from their equilibrium bond length in neutral and anion radical state to 2° Å in the step size of 0.1° Å. At each fixed C_{5'}–O_{5'} bond length, the lowest excited states were calculated. The excited states of the TNI of 5'-dTMPH, calculated at the optimized geometry of the neutral 5'-dTMPH, are shown in Fig. 17. The TD-BHandHLYP/6-31G* calculated three lowest transition

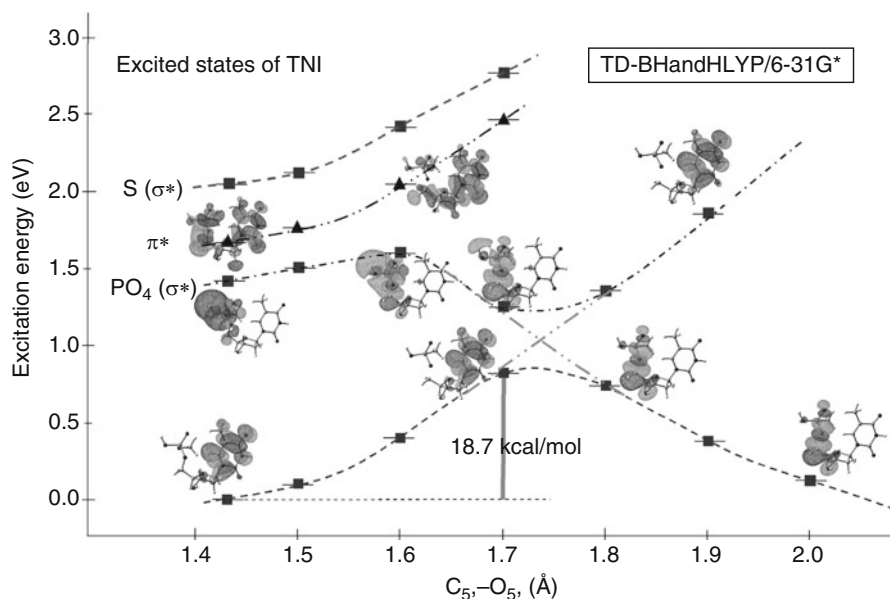


Fig. 17 Lower curve: Potential energy surface (PES) of the 5'-dTMPH transient negative ion (TNI) calculated in the neutral optimized geometry of 5'-dTMPH with $C_{5'}-O_{5'}$ bond elongation. SOMO is shown at selected points. Upper curves: calculated vertical excitation energies of the radical anion at each point along the PES; MOs involved in excitations are also shown. Energies and distances are given in electronvolts and angstroms, respectively. The lowest $\pi\pi^*$ state (triangles) and lowest $\pi\sigma^*$ states (square) are shown (Reprinted with permission from Kumar and Sevilla (2008b). © (2008) American Chemical Society)

energies of TNI were found to be $\pi(T) \rightarrow \sigma(\text{PO}_4)^*$, $\pi(T) \rightarrow \pi(T)^*$, and $\pi(T) \rightarrow \sigma(S)^*$ in nature, and the corresponding transition energies are 1.42, 1.68, and 2.06 eV, respectively. The calculated $\pi(T) \rightarrow \pi(T)^*$ transition energy 1.68 eV is in good agreement with those calculated using ETS experiment of thymine (see Table 6 and Fig. 17). From the ground-state surface, it is evident that SOMO is localized on the base as found from the ESR (electron spin resonance) experiments (Sevilla et al. 1991). From the nature of the PES in the ground state, we see that the ground-state surface (lowest surface, in Fig. 17) shows the bound character and attains the transition state at 1.7 Å with a barrier height of 18.7 kcal/mol (see Fig. 17). The second and third surfaces corresponding to $\pi(T) \rightarrow \pi(T)^*$ and $\pi(T) \rightarrow \sigma(S)^*$ transitions also exhibit the bound character. However, the first surface, which is a $\pi(T) \rightarrow \sigma(\text{PO}_4)^*$ type, shows a dissociative nature.

Such a dissociative nature of $\pi \rightarrow \sigma^*$ transitions in excited states is well documented in the literature (Domcke and Sobolewski 2003; Sobolewski and Domcke 2002; Ashfold et al. 2006). The presence of $\pi(T) \rightarrow \sigma(\text{PO}_4)^*$ transition < 2 eV clearly indicates that LEEs with this energy can directly attach to the sugar-phosphate group during TNI formation and result in a prompt SSB formation in

DNA. This transition is dark in nature and cannot be populated directly; however, it can be populated through vibronic coupling when the neighboring $\pi(T) \rightarrow \pi(T)^*$ state lies at 1.68 eV. The coupling of π^* and σ^* surfaces leads to a facile strand break formation within picoseconds. Sanche and coworkers (Martin et al. 2004; Panajotovic et al. 2006) studied the yields of SSB and DSB induced by 0–4 eV electrons and found two peaks near 0.8 ± 0.3 and 2.2 eV, respectively. Our calculated transitions at 1.4 and 1.7 eV are in reasonable agreement to the experimental observation. Further, the σ^* nature of shape resonance localized on the phosphate group in 5'-dTMPH radical anion is also supported from the recent experimental study carried out by Burrow et al. (2008) for the temporary anion states of trimethyl phosphate and several P=O groups containing compounds. The electron-scattering study of the A and B forms of DNA concluded that electrons with energies between 2 and 5 eV had the largest probability of electron capture on the phosphate group versus on the other components of the DNA (Caron and Sanche 2005). Experiments reported by Illenberger and coworkers (König et al. 2006) showed that C–O bond rupture is caused by LEE that directly attacks to the phosphate group.

The above study was extended to radical anions of a series of nucleoside diphosphate: 2'-deoxyguanosine-3',5'-diphosphate(3',5'-dGDP), 2'-deoxyadenosine-3',5'-diphosphate (3',5'-dADP), 2'-deoxythymidine-3',5'-diphosphate(3',5'-dTDP), and 2'-deoxycytidine-3',5'-diphosphate (3',5'-dCDP). The excited states of these species in gas phase and in the condensed phase in their TNI and adiabatic states were calculated using the BH and HLYP/6-31G* method (Kumar and Sevilla 2009). The geometries of each species (3',5'-dGDP, 3',5'-dADP, 3',5'-dTDP, and 3',5'-dCDP) in their neutral and radical anionic states, in the gas phase, and in the presence of two and three water molecules were fully optimized at the B3LYP/6-31G* level of theory. In this work, the nature of the resonances under the influence of the bulk solvent was modeled by the use of integral equation formalism-polarized continuum model (IEF-PCM) on the optimized hydrated geometries in the neutral and radical anion states. To the best of our information, this was the first study that showed that while in the gas phase the dissociative (σ^*) state is accessible by LEEs, in solution these states are blue-shifted toward higher energy and are not accessible by LEEs having energy under 4 eV. This phenomenon is well understood from the photoelectron spectroscopy of the solvated anions. It is well known that under solvation, anions become energetically more stable (see discussion in section “Effect of Solvation (Condensed Media)”) and a higher detachment energy (VDE) is needed in solution in comparison with the gas phase. As a result, the photoelectron spectra of solvated anions have their peaks blue-shifted (Simons 2008; Schiedt et al. 1998; Eustis et al. 2007) in comparison with their gas-phase spectra.

In this work, the lowest vertical $\pi \rightarrow \pi^*$, $\pi \rightarrow \sigma^*$ excited states and core excitation of each of the molecules in gas phase and in solution were calculated. As an example, we present the excited states of TNI of 3', 5'-dADP in the gas phase and in solvation in Fig. 18. In the gas phase, the lowest three transitions of 3',5'-dADP are $\pi(A) \rightarrow \pi(A)^*$, $\pi(A) \rightarrow \sigma(3'-PO_4)^*$, and $\pi(A) \rightarrow \sigma(5'-PO_4)^*$ in nature and the corresponding transition energies are 0.99, 1.07, and 1.14 eV, respectively.

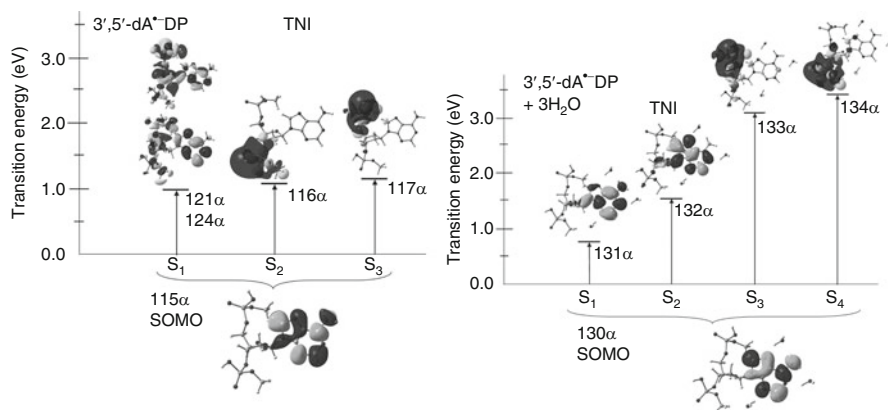
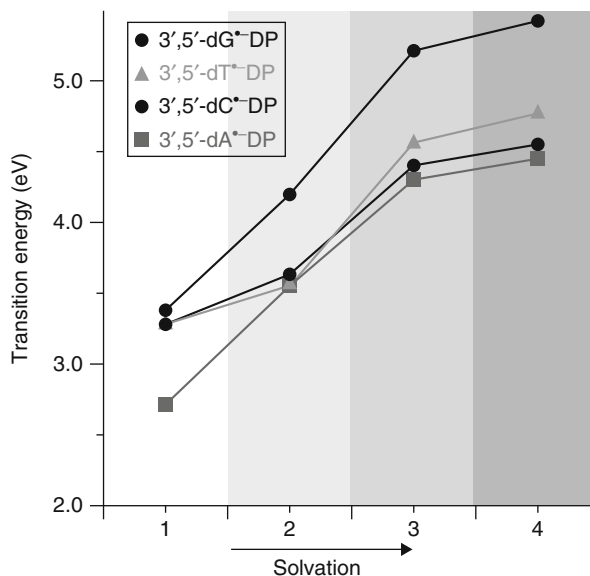


Fig. 18 TD-BH and HLYP/6-31G* computed transition energies in TNI state of 3',5'-dADP in gas phase (*left*) and aqueous solution (*right*). The effect of bulk water solvent was considered using IEF-PCM model on the trihydrated 3',5'-dA⁺-DP system. Transition from SOMO to different MOs (*shape resonance*) is shown. Transition energies are given in electronvolts (Reprinted with permission from Kumar and Sevilla (2009). © (2009) Wiley-VCH Verlag GmbH & Co. KGaA Weinheim)

The first transition, S₁, occurs from SOMO(π) \rightarrow (121 α and 124 α) MOs on the adenine base. The other two transitions, S₂ and S₃, occur from SOMO(π) to 3'- and 5'-ends of the PO₄ group. These transitions are singly dominant transitions and each has ca. 90 % contribution. In Fig. 18, we present the excited states of TNI of fully solvated 3',5'-dADP + 3H₂O. In this case, we found the two lowest transitions (S₁ and S₂) are $\pi(A) \rightarrow \pi(A)^*$ in nature, and the corresponding transition energies are 0.76 and 1.54 eV, respectively. Interestingly, the transitions $\pi(A) \rightarrow \sigma(5'-\text{PO}_4)^*$ and $\pi(A) \rightarrow \sigma(3'-\text{PO}_4)^*$ are blue-shifted under solvation toward high energy in comparison with their corresponding gas-phase values (see Fig. 18). This phenomenon is observed for all the cases studied here (for details see Kumar and Sevilla (2009)).

Since $\pi \rightarrow \sigma^*$ states, localized on 5'- and 3'-ends of the PO₄ group, are dissociative in nature, they are of particular interest. Therefore, we also calculated the variation of the transition energies of these states in (1) gas phase, (2) with discrete water molecules, and (3) and (4) in the presence of bulk solvent with two dielectric constant values $\epsilon = 7.0$ and 78.4, respectively, for TNI and their adiabatic states. As an example, the plot of the variation of transition energies of $\pi \rightarrow \sigma(5'-\text{PO}_4)^*$ with increasing solvation is shown in Fig. 19. From Fig. 19, it is evident that transition energies for the $\pi \rightarrow \sigma^*$ states increase with increasing solvation and become less accessible by LEEs < 4 eV. This conclusion is also supported by the earlier conclusion of Sevilla et al. (1994) that diffuse states are destabilized under the influence of the solvation. Therefore, from this work we conclude that σ^* states lie in the range of 1.1–1.7 eV (in the gas phase) and are in close agreement with the experimental threshold value of 0.8 ± 0.3 eV for creating SSBs (Martin et al. 2004;

Fig. 19 Variation of transition energies (eV) of the $\pi \rightarrow \sigma(5'-\text{PO}_4)^*$ excited state of nucleotides in adiabatic state with increasing solvation. Different solvation levels labeled on X-axis are 1 – gas phase, 2 – three discrete water molecules, 3 – three waters and a continuous dielectric ($\epsilon = 7.0$) and 4 – three waters and a continuous dielectric ($\epsilon = 78.4$) (Reprinted with permission from Kumar and Sevilla (2009), © (2009) Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim)



Panajotovic et al. 2006). However, the most important point we can draw from this work is that in solution these dissociative (σ^*) states increase in energy by several electronvolts. This clearly suggests that solvation of DNA would reduce the direct cleavage of DNA by LEEs in the energy range of 0–4 eV.

Leszczynski and coworkers (Bao et al. 2006; Gu et al. 2006) used the B3LYP/DZP++ method to study the $\text{C}_{5'}-\text{O}_{5'}$ and $\text{C}_{3'}-\text{O}_{3'}$ bond cleavage in pyrimidine nucleotides anion radical in their adiabatic states. This work supported the mechanism suggested by Simons and coworkers (Barrios et al. 2002). An alternative mechanism using the concept of electron-induced proton transfer mechanism was proposed by Dabkowska et al. (2005) considering a cytosine nucleotide as a model of DNA for creating SSBs in DNA. In this model, two electrons and a proton add to the cytosine base, which then undergoes proton transfer from the sugar C_2 site to the base subsequently resulting in strand cleavage. Such a mechanism may be active in radiation damage regions where several excess electrons and protons are produced in a small spatial region, that is, a spur. Using first principle molecular dynamics simulations, Kohanoff and coworkers (Gu et al. 2014a) showed that amino acids can protect DNA from damage by LEE. In their model, they used hydrogen-bonded thymine and glycine dimer and attach an excess electron to it. In this study, they found that (i) LEEs initially attached to thymine produce a thymine radical anion and then a barrierless proton transfer takes place from glycine to thymine and (ii) LEEs are also captured by the glycine, which acts as an electron scavenger thus protecting the thymine to form thymine radical anion.

Recently, Kopyra (2012) reported DEA to the gas-phase dCMP ($2'$ -deoxycytidine $5'$ -monophosphate) using a crossed electron molecular beam arrangement experiment. From the yield of the different ions and resonant electron attachment profile to

dCMP, they concluded that both direct electron attachment to backbone and excess electron transfer from cytosine (base) to backbone contribute the SSB. They further estimated the relative contribution of these mechanisms leading to SSB. According to their estimate, (a) 60 % electron localizes at the phosphate unit, (b) 25 % electron localizes at the sugar unit, and (c) 15 % electron initially localizes at the base and transfers to backbone. These estimates clearly suggest that direct localization of LEE to backbone leading to SSB is a major pathway, while *base-to-backbone* excess electron transfer leading to SSB is likely only a minor pathway. This experimental study supports the theoretically proposed mechanism of direct LEE attachment to phosphate group leading to SSB (Li et al. 2003a; Kumar and Sevilla 2008b; Kumar and Sevilla 2009).

We note that most of the theoretical investigations used density functional theory, particularly B3LYP method to estimate the barriers for $C_{3'}-O_{3'}$ and $C_{5'}-O_{5'}$ bonds dissociation due to LEE electron attachment to the DNA units. Generally, the barrier computed by the B3LYP method is underestimated because at transition state electron overly delocalizes from base to sugar phosphate region (Kumar and Sevilla 2008). This excess delocalization in B3LYP method mainly arises due to self-interaction error. This discrepancy in the B3LYP method in underestimating the barrier for C–O bond dissociation was also raised in a very recent study by Chen et al. (2014), which is described below.

In the recent work, Chen et al. (2014) benchmarked the different DFT methods to compute the barrier for $C_{3'}-O_{3'}$, $C_{5'}-O_{5'}$, and $C_{1'}-N$ bonds dissociation on LEE attachment to a nucleobase. The benchmark was carried out relative to the most accurate CCSD(T) method. The study showed that the glycosidic ($C_{1'}-N$) bond cleavage was well described by the DFT methods, but the barriers for $C_{3'}-O_{3'}$, $C_{5'}-O_{5'}$ bonds dissociation were severely underestimated. The error for $C_{3'}-O_{3'}$, $C_{5'}-O_{5'}$, and $C_{1'}-N$ bonds dissociation relative to CCSD(T) is shown in Fig. 20. From Fig. 20, it is evident that M06-2X and ω B97XD density functionals are quite accurate in predicting the $C_{3'}-O_{3'}$, $C_{5'}-O_{5'}$ bonds dissociation and the absolute maximum error relative to CCSD(T) is ca. 2 kcal/mol (Chen et al. 2014). The most used B3LYP method underestimated the C–O bond dissociation barrier by about 7 kcal/mol relative to CCSD(T), see Fig. 20. On the basis of this benchmark study, authors also concluded that *base-to-backbone* electron transfer leading to SSB does not occur as easily as proposed in earlier studies. Alternatively, the direct LEE localization to phosphate group to produce SSB is a possible mechanism.

Recently, using DFT (M06-2X/6-31+G(d,P) quantum mechanics/molecular mechanics (MM/QM) molecular dynamics simulations, Cauët et al. (2013) considered solvated 12-mer fragment of B-DNA ((a) 3'-TTTTTTTTTTTTT-5' (b) 3'-CCCCCCCCCCC-5' (c) 3-GCGCTATGCGA-5') to study LEE-induced strand breaks. In this study, 2-base pairs in the center of the 12-mer fragment were considered by the QM, while rest of the system was considered by the MM. From the analyses of the electronic and the nuclear Fukui functions, they found that after vertical electron attachment to the DNA fragment (12-mer B-DNA) the position and dynamics of the atomic nuclei in DNA are largely modified. The LEE interaction

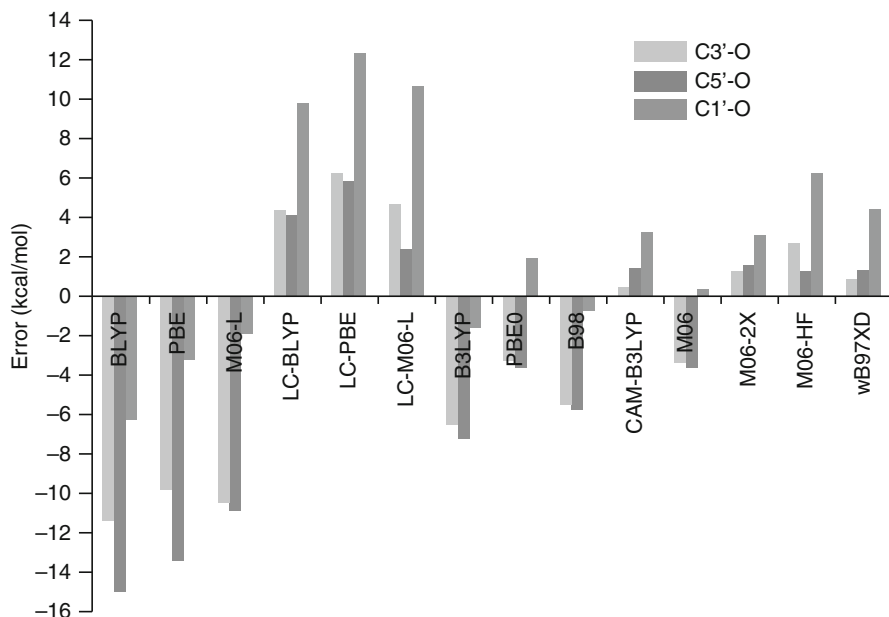


Fig. 20 Errors in $C_{3'}-O_{3'}$, $C_{5'}-O_{5'}$, and $C_{1'}-N$ bonds dissociation by different DFT methods relative to CCSD(T) calculations (Reprinted with permission from Chen et al. (2014), *J. Phys. Chem.* © (2014). American Chemical Society)

induces nuclear motions that involve the elongation of P–O and C–C (deoxyribose) bonds. If the acquired kinetic energy becomes high enough, the neighboring $C_{3'}-O_{3'}$ or $C_{5'}-O_{5'}$ phosphodiester bonds may break within a 15–30 fs timescale. In other words, DNA strand breaks can take place before electron localization on the base. Such fast reaction to induce strand breaks in DNA needs experimental verification as strand break formation will be in competition with electron localization at the base and subsequent protonation of the anionic base radical from solvent.

In addition to these works, a number of other studies related to the LEE-induced DNA damage have appeared over the last few years (Zhang et al. 2008; Xie et al. 2008; Takayanagi et al. 2009; Solomun et al. 2009; Solomun and Skalicky 2008; Gu et al. 2011; Smyth and Kohanoff 2012; Baccarelli et al. 2011; González-Ramírez et al. 2012).

Profile of DNA Damage Produced by LEE and Ionizing Radiation

Very recently, Sanche and coworkers (Park et al. 2013) experimentally compared the DNA damage produced by LEE (<10 eV) and high-energy photons. In this study, they considered solid TpTpT trinucleotide as a model of DNA. The radiation-produced damage was measured by high-performance liquid chro-

matography (HPLC) with UV detection and HPLC coupled with tandem mass spectrometry. Their analyses found thymine release, strand breaks, and the formation of base modifications, such as 5,6-dihydrothymine, 5-hydroxymethyluracil, and 5-formyluracil. The damage by high-energy X-rays was found to be 20 % (base release), 19 % (strand break), and 61 % (base modifications), respectively, while the corresponding damage produced by LEE (≤ 10 eV) was found to be 35 %, 13 %, and 51 %, respectively. X-rays give lower base release in comparison with LEE. From the analyses, it was found that LEEs overall give a similar profile of products compared to ionizing radiation. The similarity in damage profiles suggests that LEE maybe a significant contributor in X-ray damage.

LEE-Induced Base Release

In addition to LEE-induced SSBs and DSBs in DNA, glycosidic bond cleavage (base release) induced by LEE has also been experimentally observed by Sanche and coworkers (Zheng et al. 2004, 2005, 2008) as well as by Illenberger and coworkers (Ptasińska et al. 2006; Abdoul-Carime et al. 2004b). LEEs with 4–15 eV energy were found to be effective in causing base release in thymidine, as well as in the DNA sequence 5'-TpTpT-3' (Zheng et al. 2004, 2008). The initial quantum yield for thymine release in thymidine was found to be 3.2×10^{-2} preincident electron (Zheng et al. 2004). The conversion of thymidine into thymine and sugar was proposed due to the localization of LEEs into the antibonding MOs of $N_1-C_{1'}$ (glycosidic) bond, which leads either to homolytic or heterolytic cleavage of the $N_1-C_{1'}$ bond joining the thymine and sugar ring (Zheng et al. 2004).

Using the B3LYP/DZP++ method, Schaefer and coworkers (Gu et al. 2005) calculated the ground-state adiabatic PESs for the $N_1-C_{1'}$ bond dissociation considering the radical anions of 2'-deoxyribothymidine (dT) and 2'-deoxyribocytidine (dC). Subsequently, a detailed analysis of the base release process in dC, dT, and deoxyadenosine in two conformations, that is, dA1 (with internal hydrogen bonding) and dA2 (without hydrogen bond), was carried out by Li, Sanche, and Sevilla (Li et al. 2006) using B3LYP/6-31+G(d) level of theory. This work provided some insight into the relative energies of $N_1-C_{1'}$ (glycosidic) bond dissociation of these molecules in their neutral and anion radical states in gas phase as well as in the presence of bulk solvent via the PCM method. This study showed that LEEs in fact greatly weakened the $N_1-C_{1'}$ bond strength of these molecules in comparison with their corresponding bond strength in neutral state. The bond dissociation energy of the $N_1-C_{1'}$ bond in gas phase is found to be exergonic for dA ($\Delta G = -14$ kcal/mol) and dT ($\Delta G = -6$ kcal/mol) and is endergonic for dC ($\Delta G = 1$ kcal/mol). The PESs for the $N_1-C_{1'}$ bond dissociation of these molecules are shown in Fig. 21. From Fig. 21, it is clearly evident that the barrier for the $N_1-C_{1'}$ bond dissociation is located between 19 and 23 kcal/mol for all the molecules. The experimental value for LEE-induced thymine release in dT of about 1 eV (ca. 23 kcal/mol) (Zheng et al. 2004, 2005, 2008; Ptasińska et al. 2006; Abdoul-Carime et al. 2004b) is in good agreement with the calculated value of barrier for thymine release in dT

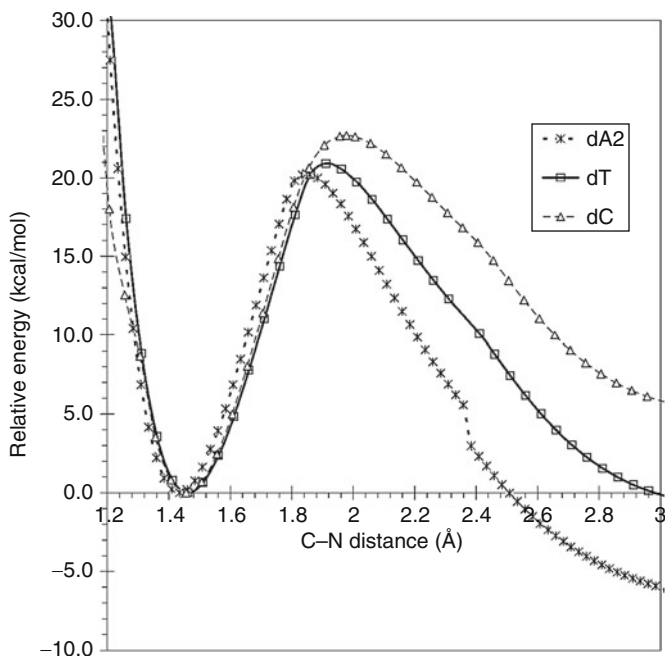


Fig. 21 Anionic N1-C potential energy surfaces of deoxyadenosine (*dA2*), thymidine (*dT*), and deoxycytidine (*dC*). *dA1* has a similar profile to *dA2* (Reprinted with permission from Li et al. (2006). © (2006) Radiation Research Society)

(19 kcal/mol) (see Fig. 21). From the SOMO, it was found that in adiabatic states all the molecules showed the valence-bound π^* state. Also, on the PES as the $N_1-C_{1'}$ bond extends from its equilibrium value, the π^* (valence state) crosses the antibonding σ^* dissociative state at transition state and excess electron localized on $N_1-C_{1'}$ bond region creating a negative nitrogen site on the base and a free radical at $C_{1'}$ site on the sugar ring.

Effect of Solvation on Strand Break Formation

In comparison with gas-phase investigations, studies concerning electron-induced strand breaks under the influence of solvation are few in number. From pulse radiolysis experiments, it is well known that solvated electrons do not cause strand breaks upon attachment to DNA in an aqueous environment (von Sonntag 1987; Nabben et al. 1982); this is discussed in detail in section “Effect of Solvation (Condensed Media).” Most theoretical studies to date have assumed adiabatic systems and have attempted to explain the effect of solvation on strand breaks using polarized continuum model (Simons 2006; Barrios et al. 2002; Li et al. 2003a, 2006; Bao et al. 2006; Gu et al. 2006). Though the PCM model is adequate for

bulk solvent effects, it neglects the specific local interactions and hydrogen-bonding energetics, which are substantial contributions to solvation energetics and are critical in determining reaction pathways. This level of sophistication is usually avoided owing to the increase in complexity of the treatment.

In order to gain an understanding of the influence of specific hydrogen-bonding effects in an aqueous environment on potential strand breaks, we considered (Kumar and Sevilla 2007) a 5'-dTMP that includes a solvation shell of 11 waters and a sodium counterion to neutralize the phosphate group. Using this hydrated (5'-dTMPNa + 11 H₂O) model, we calculated the adiabatic PES of C_{5'}-O_{5'} bond dissociation in the neutral and radical anion states using the B3LYP/6-31G** level of theory. On the PES of radical anion state, we also plotted the SOMO at each step of the C_{5'}-O_{5'} bond elongation. This calculation showed that the solvent has a pronounced effect, which would essentially eliminate strand break formation. The barrier for the C_{5'}-O_{5'} bond dissociation in the radical state was found to be 29 kcal/mol. This is in accord with the experimental finding that aqueous electrons do not cause strand breaks.

Very interestingly, we found that SOMO initially localized on the π* MO on thymine was still localized on the thymine at the TS. However, in gas phase (see Figs. 15 and 16), the SOMO transferred to the σ* C_{5'}-O_{5'} bond at the TS. The B3LYP/6-31G** calculated AEA and VDE of 5'-dTMPNa + 11 H₂O in full solvent using PCM was found to be 2.17 and 3.20 eV, respectively. This study confirmed that in comparison with the gas phase, the vertical detachment energies (VDEs) of the 5'-dTMP radical anion in the solvated environment are enhanced appreciably (3.20 eV), which shows that the excess electron in the solvated anion radical is highly stable toward electron detachment. Recently, the 3'- and 5'-C-O bond rupture in cytosine monophosphate was studied by Schyman et al. (2008) in the aqueous environment using a molecular dynamics study. The authors also found a higher barrier for the 3'- and 5'-C-O bond rupture. In another study, Schyman and Laaksonen (2008) studied the C_{3'}-O_{3'} bond dissociation in guanine 3'-monophosphate (3'-GMPH) radical anion in gas phase as well as in the presence of 21 water molecules using the B3LYP/DZP++ method. In this calculation, the authors found dipole-bound state for gas phase, which was expected. Surprisingly they also reported dipole-bound state for the hydrated system. As has already been mentioned, solution diffuse states such as dipole-bound states are destabilized relative to valence-bound states, which become lower in energy (see Fig. 10 and Li et al. (2002a)). This is already well established experimentally from the photoelectron spectroscopy of nucleic acid bases (Schiedt et al. 1998) as well as from theoretical calculations (Li et al. 2002a). Thus, the calculation of Schyman and Laaksonen (2008) for hydrated 3'-GMPH radical anion is considered to be problematic.

Loos et al. (2009) calculated the C_{3'}-O_{3'} bond cleavage of the radical anion of 2-deoxycytidine-3'-monophosphate moiety (3'-dCMPH) in single- and double-stranded DNA (5'-CCC-3' sequence) in gas phase using the QM/MM approach. The authors also reported the dipole-bound diffuse state localized near the phosphate group. They did not, however, report the initial attachment of the excess electron

on the cytosine. Thus, the approach which may be applicable to the gas phase is not predictive for the condensed phase. A similar calculation was already done by Simons and coworkers (Anusiewicz et al. 2004) considering a single-strand CCC sequence. Due to the large size of the system, they considered the ONIOM approach. The central nucleotide was treated with the UHF method with a 6-311+G* basis set; however, the two terminal nucleotides were treated at the lower level using a 4-31G* basis set. In this calculation, the initial attachment of an excess electron was found to be localized on the π^* MOs of the middle cytosine, showing valence-bound nature as expected from ESR experiments (Sevilla et al. 1991). Kobyłeczka et al. (2008) also showed a valence-bound anion formation on thymine in the DNA π -stacked system using the MP2/aug-cc-pVDZ method.

Park et al. (2011) investigated the DNA damage caused by LEE for nucleotides and short oligonucleotides in the solid state. The samples of TpTpT were irradiated with monoenergetic electrons of energy 10 eV and subsequently analyzed by high-performance liquid chromatography (HPLC) and mass spectrometry. The study showed that about 67 % of the total damage was identified as 5,6-dihydrothymine and 33 % damage involved the phosphodiester–sugar (C–O) bond and the *N*-glycosidic bond (CN) dissociation. However, in a similar type of study (Li et al. 2010), the total damage (base release and C–O bond cleavage) was found to undergo a twofold decrease for the following trimer sequences in the order: TTT > TCT > TAT > TGT. The nonmodified base release was found to be 17–24 % of the total damage and mainly occurred from the terminal sites of the trimers. The phosphodiester–sugar (C–O) bond cleavage was found to be only 9–20 %. Both the studies certainly conclude that an excess electron initially attached to the base forms a stable radical anion, which protonates to form 5,6-dihydrothymine as observed in a variety of ESR experiments on hydrated DNA samples (Sevilla et al. 1991; Gregoli et al. 1989). Once localized to the base, no transfer to phosphodiester–sugar bond region is likely, as discussed in sections “Effect of Solvation (Condensed Media)” and “Proposed Theoretical Models of DNA Damage.” Thus, SSBs are not produced from the adiabatic base anion radicals by reductive cleavage.

Using B3LYP/DZP++ level of theory, Gu et al. (2010a, b) reported the $C_{3'}-O_{3'}$, $C_{5'}-O_{5'}$ σ bond and *N*-glycosidic bond dissociation energies of anion radicals of 2'-deoxycytidine-3',5'-diphosphate (3',5'-dCMP), 2'-deoxythymidine-3',5'-diphosphate (3',5'-dTMP), and 2'-deoxyguanosine-3',5'-diphosphate (3',5'-dGMP) in gas phase and in aqueous media using the PCM model. For 3',5'-dCMP and 3',5'-dTMP, the order of bond dissociation energy was found to be $C_{3'}-O_{3'} < C_{5'}-O_{5'} < N$ -glycosidic bond, and the calculated energies lie in the range ca. 6–29 kcal/mol for both gas phase and in aqueous solution. For 3',5'-dGMP, the energies for breaking the corresponding bonds in the gas phase were calculated to be 11–24 kcal/mol, while in aqueous solution, the energies decrease appreciably than their gas-phase values and lie in the range ca. 1–10 kcal/mol, which indicates that $C_{5'}-O_{5'}$ bonds are more susceptible for cleavage than $C_{3'}-O_{3'}$ bonds. However, these activated processes are unlikely in aqueous solutions since the faster protonation reactions of the anion radicals will outcompete with such processes as pointed out above.

ESR work by Becker et al. (1996, 2003) has found evidence for both phosphodiester–sugar (C–O) bond cleavage and P–O bond cleavage (see Scheme 1). In these experiments, free radicals formed from 60 MeV/nucleon $^{16}\text{O}^{8+}$ ion beam irradiation of DNA samples at 77 K were analyzed by the ESR, and a new phosphorus-centered radical species was identified by its large ^{31}P parallel hyperfine coupling of about 780 G (Becker et al. 1996). The radical was identified as a phosphoryl radical, which was produced by the direct dissociation of the P–O bond. However, the yield of this radical was found to be very small (0.1 % of the total radicals). In other experiments, DNA samples were irradiated by ^{36}Ar and ^{40}Ar ion beams, and these two types of radicals were again identified by ESR at 77 K (Becker et al. 2003). These radicals were produced from dissociation of the C–O and P–O bonds at the $\text{C}_{3'}$ of the sugar moiety. The yield of the $\text{C}_{3'}$ -dephos radical produced by C–O bond dissociation was estimated as 95 %, while the ROPO_2^- produced by P–O bond dissociation was estimated to be 5 %.

Recently, a mechanism of SSB by P–O bond fragmentation in DNA was proposed by Rak et al. (2011) using B3LYP/6-31++G** level of theory. In their model, the authors proposed that LEE initially attached to the thymine, which is subsequently protonated at the C5 position of the thymine to produce the dihydro-6'-thymyl radical. In the next event, intramolecular hydrogen atom transfer takes place from $\text{C}_{3'}$ or $\text{C}_{5'}$ sites of sugar ring to the C6 site of the thymine, and finally the P–O bond is broken. However, from the calculated kinetic and thermodynamic barriers related to the P–O bond fragmentation, the authors concluded that the proposed mechanism was not favorable over the LEE-induced C–O bond fragmentation. ESR studies in solution suggest protonation at C6 of the thymine anion radical is favored, and the most favored reaction path for the $\text{C}_{5'}$ radical in aqueous media is by phosphate loss, that is, C–O bond cleavage.

LEE Interaction with 5-Bromouracil (5-BrU) and Radiosensitization Activity

Uracil (U) and thymine are very similar in structure and differ only by substitution at the C5-position (H for uracil and CH_3 for thymine). Thus, in DNA, 5-halogenated uracils are easily incorporated in place of thymine (Djordjevic and Szybalski 1960; Bagshaw et al. 1967; Doiron et al. 1999). These halogenated uracils are very sensitive to radiation such as γ - and X-rays and UV light. Because of this property (photosensitivity), 5-halogenated uracils were considered an efficient radiosensitizers in anticancer therapy. Particularly, 5-BrU and 5-Idouracil (5-IU) showed increased sensitivity in DNA to UV and X-ray radiation, and 5-BrU is used as lead drug for radiotherapy (Djordjevic and Szybalski 1960). 5-BrU is most easily incorporated in DNA as the bromine atom is very close in size to a methyl group. It is also observed that irradiated 5-BrU labeled DNA showed more damage to DNA (SSB and DSB) in comparison with native DNA (Webb et al. 1993; Dextraze et al. 2007; Li et al. 2011; Park et al. 2012; Hashiya et al. 2014; Wiczek et al. 2013; Zdrowowicz et al. 2014; Ling and Ward 1990). A number of

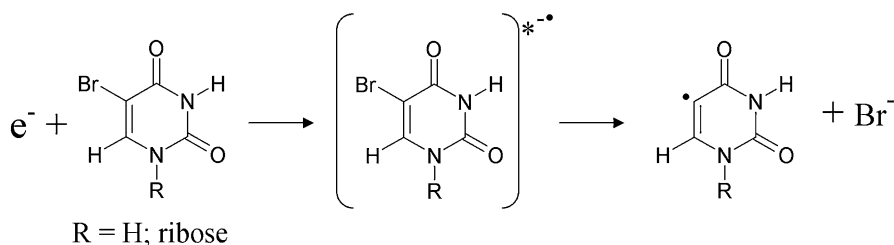


Fig. 22 Attachment of an excess electron to 5-BrU leads to transient negative ion (5-BrU)^{*-} formation which produces uracil-5-yl radical and bromide anion via DEA mechanism

theoretical and experimental studies confirmed that its radiosensitization activity is primarily related to the rapid reaction of the 5-BrU with secondary electrons that are generated during radiation (Wieczór et al. 2014; Golon et al. 2014; Sevilla et al. 1974; Li et al. 2002b, 2003b; Riederer et al. 1978; Voit et al. 1986; Razskazovskii et al. 1997; Wang et al. 2006; Wang and Lu 2007, 2010). The attachment of the secondary electrons to 5-BrU leads to the formation of very reactive uracil-5-yl radical (U-5-yl[•]) and bromide anion (Br⁻) via dissociative electron attachment mechanism as shown in Fig. 22. The U-5-yl[•] has been suggested to abstract a H atom from a neighboring 2'-deoxyribose unit in double-stranded (ds) DNA, which then leads to strand break formation, that is, SSB and DSB, and neutral U base release (Dextraze et al. 2007; Li et al. 2011; Park et al. 2012; Hashiya et al. 2014; Wiczek et al. 2013). However, very recently, in an interesting study (Chomicz et al. 2014) regarding the reactivity of U-5-yl[•], the authors point out that to produce a DNA strand break the sugar C3', C4', or C5' hydrogens must be abstracted; however, in dsDNA these hydrogens are too far from the adjacent 5-U-yl[•], to be readily involved in this process. These workers suggest an alternative mechanism that 5-U-yl[•] reacts with a water molecule to produce uracil and OH radical, which is described below.

Hydration of the Uracil-5-yl Radical

A joint experimental (electron spin resonance (ESR)) and theoretical (DFT) approach was used to study the hydration of U-5-yl[•] [244]. In this study, the ESR spectra of γ -irradiated aqueous glasses (7.5 M LiCl in H₂O/D₂O) containing 5-Bromouridine (5-BrdU) at 77 K show the initial formation of 5-BrdU π -anion radical. After 15-min annealing at 145 K, U-5-yl[•] was formed due to loss of Br anion. Further, after 15-min annealing at 160 K, hydrated U-5-yl[•] (5-UHOH[•]) was formed and was confirmed by the four line components in the ESR spectrum. The experimental ESR spectra are shown in Fig. 23. The reaction of U-5-yl[•] with water to form 5-UHOH[•] (hydrate radical) was also reported earlier by Sevilla et al. (1974) using ESR. The formation of 5-UHOH[•] was proposed due to addition of

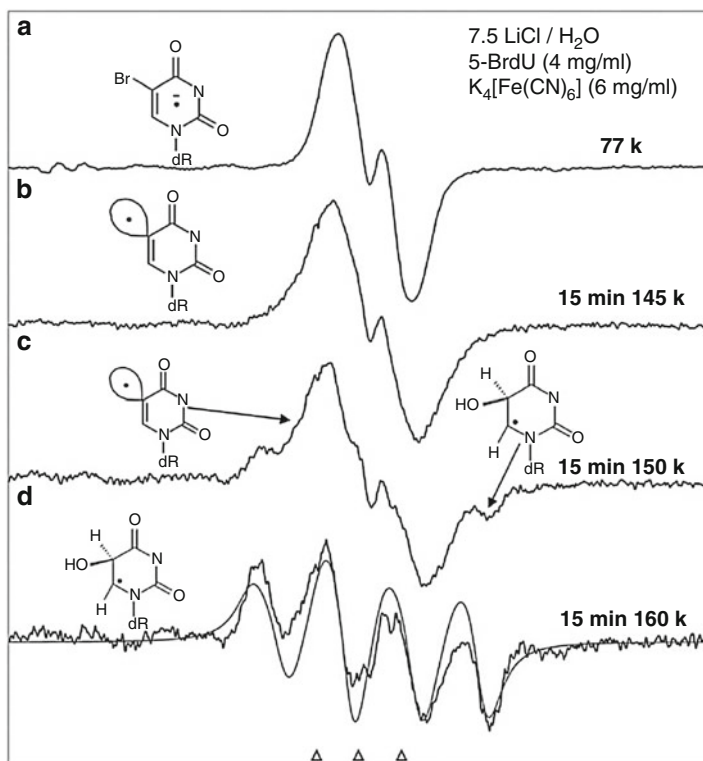


Fig. 23 ESR spectra of BrdU in 7.5 M LiCl in H₂O with K₄[Fe(CN)₆] as a hole scavenger (a) immediately after γ -irradiation; (b–d) obtained after annealing the sample for 15 min; all recorded at 77 K. Structure of the radical assigned to each spectrum is shown. The simulated (gray) spectrum (d) employed is a combination of ca. 70 % quartet from 5-UHOH[•] (H-axial to the radical site p-orbital) and of ca. 30 % doublet from 5-UHOH[•] (H-equatorial to the radical site p-orbital) as well as from 6-UHOH[•] and is superimposed on the experimental (black) spectrum for comparison. The three triangular field markers are individually separated by 13.09 G with the central marker at $g = 2.0056$ (Reprinted with permission from Chomicz et al. (2014), *Chem. Commun.*, © (2014), The Royal Society of Chemistry)

HOH to the C5 radical site of neutral U-5-yl radical. To support this experiment, DFT calculations using M06-2X/6-31++G(d, p) method was used to calculate the reaction energy between U-5-yl[•] and a water molecule to directly form the hydrate radical 5-UHOH[•] or abstract from water to form uracil and OH[•] (U-5-yl[•] + H₂O → U + OH[•]). In the calculation, the effect of full solvent was considered through the use of polarized continuum model (PCM). Considering two possible reaction modes of a water molecule with U-5-yl[•], see Fig. 24, the energies of the reactants, transition state, and product were calculated. From the calculation, the forward lowest barrier ca. 12 kcal/mol was calculated to form uracil and OH[•] (product), and the product was calculated to be nearly isoenergetic with respect to reactants (U-5-yl[•] + H₂O). Subsequently uracil and OH[•] were suggested to rapidly react

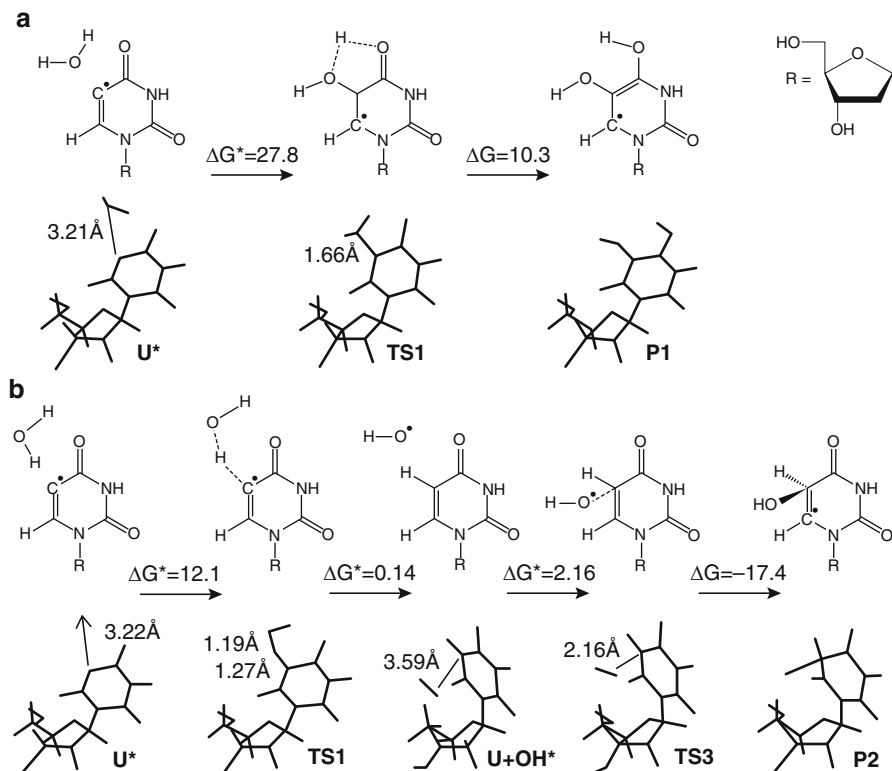


Fig. 24 Possible reaction paths calculated at the M06-2X/6-31++G(d,p) level for 2'-deoxy-5-U-yl^{*} and the water molecule system. ΔG and ΔG^* stand for reaction free energy and activation free energy, respectively. All values are expressed in kcal/mol. For 2D models, R refers to the 2'-deoxyribose ring. P1 and P2 are tautomers with P2 being the most stable (Reprinted with permission from Chomicz et al. (2014), *Chem. Commun.* © (2014), The Royal Society of Chemistry)

to form C5-adduct radical (5-UHOH^{*}), involving a small activation barrier of ca. 2 kcal/mol, see Fig. 24. The C5-adduct radical (5-UHOH^{*}) was found to be more stable than the reactants (U-5-yl^{*} + H₂O) by ca. 17 kcal/mol. This interesting work provides an alternative pathway to the proposed hypothesis that U-5-yl^{*} abstracts hydrogen atom from neighboring 2'-deoxyribose moiety yielding strand breaks (Dextraze et al. 2007; Li et al. 2011; Park et al. 2012; Hashiya et al. 2014; Wiczek et al. 2013). It is also known that in ds-B-DNA, U-5-yl^{*} could abstract hydrogen atom from C_{1'}-H and C_{2'}-H of the adjacent sugar moiety as these two sites are close to the U-5-yl^{*} (Sugiyama et al. 1996; Tashiro and Sugiyama 2010). It is well known that C_{1'} and C_{2'} do not produce strand breaks in ds-B-DNA. Thus, in this situation, the produced U-5-yl^{*} can also react with surrounding water molecule to produce uracil and OH radical. The produced OH radical can abstract H atom from 2'-deoxyribose to yield strand breaks in ds-B-DNA.

Modeling 5-Substituted Uracil as an Efficient DNA Radiosensitizer

Recently, eight 5-substituted 1-methyl-uracil derivatives (5-R-U) with R=SCN, NCS, SH, CN, OCN, NCO, NNN, andNOO (see Fig. 25) were considered by theory and experiment. The reaction of these molecules with an excess electrons was studied using the B3LYP/6-31++G(d,p) method, including full effect of solvent through PCM (Chomicz et al. 2013). The aim of this study is to search a potent DNA radiosensitizer which has similar reaction mechanism as 5-BrU. In devising a potent DNA radiosensitizer, authors considered two criterions: (i) the molecule should have large electron affinity and form stable anion radical after electron attachment, and (ii) the chemical bond that connects a substituent (R) to the C5 atom of uracil ideally should be quite weak to produce U-5-yl[•], after electron attachment (see Fig. 25). Thus, to meet these two requirements, authors (i) studied the electron capturing strength of these derivatives by calculating their AEA and VDE in the gas phase and in aqueous solution, and (ii) estimated the reaction barrier to produce U-5-yl[•] and R⁻ from (R-U)^{•-}. From the calculations in the aqueous phase, it was found that all the considered molecules have positive AEAs in the range 2.26 eV to 3.55 eV and the calculated VDEs lie in the range 2.66 eV to 3.86 eV. The large positive AEAs and VDEs clearly indicate that all the molecules are able to capture hydrated electrons and are quite stable against electron detachment. The reliability of the calculated VDE for these molecular anion radicals was tested by comparison with the results of photo electron spectroscopy experiments, which were in good agreement. The calculated reaction barrier to produce U-5-yl[•] and R⁻ from (R-U)^{•-} for all the considered molecules lies in the range 2.5 kcal/mol to 41.2 kcal/mol (see Fig. 26). From the analyses of the AEAs, VDEs, and the reaction barriers, SCN-U and OCN-U were proposed to be the possible potent radiosensitizers. The calculated AEA, VDE, and barrier height for SCN-U are 2.70 eV, 3.07 eV, and 3.4 kcal/mol, respectively, and the corresponding values for OCN-U are 2.62 eV, 3.07 eV, and 2.5 kcal/mol, respectively. This study demonstrates that solvated electron after attachment to SCN-U and OCN-U is predicted to dissociate the C5-S and C5-O bonds of SCN-U and OCN-U to produce U-5-yl[•] involving a small barrier of ca. 3 kcal/mol (see Fig. 26). For 5-BrU, the AEA, VDE and the barrier for

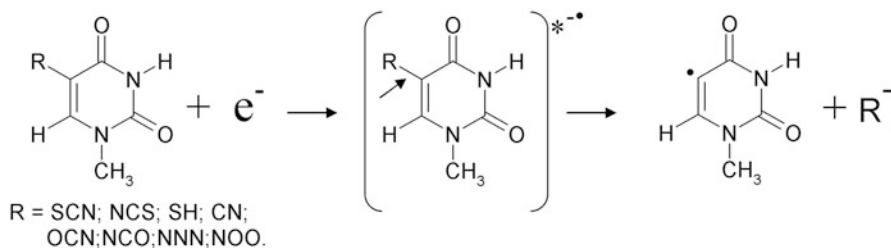


Fig. 25 Schematic diagram showing the reaction of excess electron with 5-substituted modified uracils. *Arrow* points the C5-R bond considered for dissociation

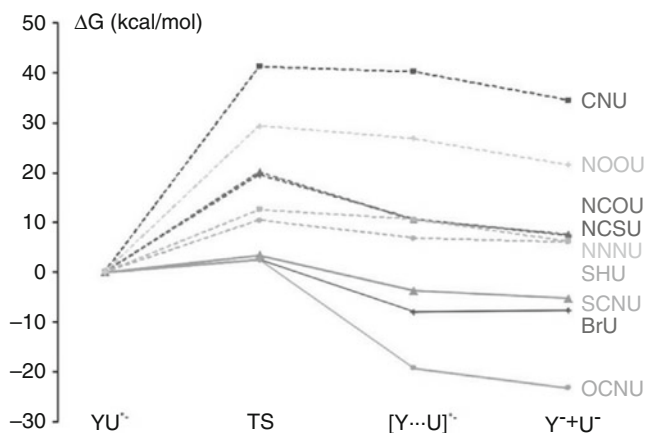


Fig. 26 B3LYP-PCM/6-31++G(d,p) calculated free energy profiles in kcal/mol for electron-induced degradation of 5-substituted modified uracils (Reprinted with permission from Chomicz et al. (2013), *J. Phys. Chem. Lett.* © (2013), American Chemical Society)

C5–Br bond dissociation are 2.48 eV, 2.74 eV, and 2.5 kcal/mol, respectively (see Fig. 26). Further, the mass spectrum of SCN-U anion radical shows strong signal intensity originating from SCN⁻, which confirms the theoretical prediction. We also note that in this study authors investigated the dissociation of only the C5–R bond (see Fig. 25) and did not check the barrier for other possible bond fragmentation. For example, there are few reports (Hamed et al. 2006; Houmam et al. 2003a, b; Houmam 2008; Ciszek et al. 2004) in the literature that one electron reduction of thiocyanate (SCN)-substituted benzyl and aryl compounds leads to the dissociation of the S–CN bond, that is, R–SCN + e⁻ → RS[•] + CN⁻ (R = benzyl; aryl). The RS[•] further reacted to form R-S-S-R dimer. Thus, an experimental verification was needed to further characterize electron-induced bond cleavage site in SCN-U, and very recently, this process has been tested by Zdrowowicz et al. (2015) employing ESR and product analysis experiments along with theory. These workers found that both S–CN bond cleavage and C–SCN bond cleavage occurred in a 10 to 1 ratio favoring S–CN bond cleavage to form U–S[•] + CN⁻. The major radical produced U–S[•] was found to dimerized to the U-S-S-U product, which was identified by LC-MS.

Conclusion

Our understanding of the interaction of low-energy electrons with DNA has increased substantially over the past few years. From experiment, it is clear that single- and double-strand breaks, base damage, and base release are all direct results of DEA. For strand breaks, addition of an electron to the base and the sugar phosphate backbone can lead to strand breaks from shape and core-excited

resonances. From their higher energies and nature, core-excited resonances are those likely to be involved in double-strand breaks. Simple vibrational excitations have been shown to provide the necessary excitation when combined with an even 0 eV electron to result in strand cleavage. While much work presented in this chapter has focused on the DNA components without surroundings (“gas-phase” calculations), these calculations often give diffuse states (dipole-bound electrons and continuum states) as ground states for the system. Theoretical treatments of DNA systems, relevant to biology, must include the aqueous phase, including specific water hydrogen bonding, as well as the base stacking and counterion placement. In such systems, diffuse states are greatly destabilized and valence-bound states dominate. Theoretical methods such as DFT with various solvation models are now able to treat more complex and more biologically relevant systems. The reaction of excess electrons with DNA depends on its energy, phase, and state of solvation, that is, the LEEs in the gas phase (0 – 20 eV) and prehydrated electron (e^-_{pre}) (–0.5 to –1.0 eV) are expected to be more damaging and dissociate chemical bond, while fully solvated electron (e^-_{aq}) is considerably less damaging and is captured by DNA bases to form base anion radicals which subsequently are protonated from surrounding solvent.

A variety of compounds have been employed as radiosensitizers to treat cancer, and 5-Bromouracil has been of considerable interest in this regard. The radiosensitization activity of 5-bromouracil is due to its reaction with electron which rapidly produces bromide anion and U-5-yl^{*}. The highly reactive U-5-yl^{*} formed is believed to then abstract hydrogen from neighboring sugar moieties in DNA to create strand breaks. While this mechanism is quite likely, other mechanisms have been proposed and discussed in this review. Initiative has been taken to develop more efficient radiosensitizers employing accurate DFT methods. In the future, improved techniques should provide a better understanding of the actions of LEEs with the DNA in a biological environment.

Acknowledgment This work was supported by the NIH NCI under Grant R01CA045424. The authors thank Prof. P.D. Burrow for helpful discussions and advice.

Bibliography

- Abdoul-Carime, H., Huels, M. A., Illenberger, E., & Sanche, L. (2001). Sensitizing DNA to secondary electron damage: Resonant formation of oxidative radicals from 5-halouracils. *Journal of the American Chemical Society*, *123*, 5354.
- Abdoul-Carime, H., Gohlke, S., & Illenberger, E. (2004a). Site-specific dissociation of DNA bases by slow electrons at early stages of irradiation. *Physical Review Letters*, *92*, 168103.
- Abdoul-Carime, H., Gohlke, S., Fischbach, E., Scheike, J., & Illenberger, E. (2004b). Thymine excision from DNA by subexcitation electrons. *Chemical Physics Letters*, *387*, 267.
- Abel, B., Buck, U., Sobolewski, A. L., & Domcke, W. (2012). On the nature and signatures of the solvated electron in water. *Physical Chemistry Chemical Physics*, *14*, 22.
- Adhikary, A., Malkhasian, A. Y. S., Collins, S., Koppen, J., Becker, D., & Sevilla, M. D. (2005). UVA-visible photo-excitation of guanine radical cations produces sugar radicals in DNA and model structures. *Nucleic Acids Research*, *33*, 5553.

- Adhikary, A., Collins, S., Koppen, J., Becker, D., & Sevilla, M. D. (2006a). C5'- and C3'-sugar radicals produced via photo-excitation of one-electron oxidized adenine in 2'-deoxyadenosine and its derivatives. *Nucleic Acids Research*, *34*, 1501.
- Adhikary, A., Kumar, A., & Sevilla, M. D. (2006b). Photo-induced hole transfer from base to sugar in DNA: Relationship to primary radiation damage. *Radiation Research*, *165*, 479.
- Adhikary, A., Collins, S., Khanduri, D., & Sevilla, M. D. (2007). Sugar radicals formed by photoexcitation of guanine cation radical in oligonucleotides. *The Journal of Physical Chemistry B*, *111*, 7415.
- Adhikary, A., Khanduri, D., Kumar, A., & Sevilla, M. D. (2008). Photoexcitation of adenine cation radical [A⁺⁺] in the near UV-vis region produces sugar radicals in adenosine and in its nucleotides. *The Journal of Physical Chemistry B*, *112*, 15844.
- Adhikary, A., Khanduri, D., Pottiboyina, V., Rice, C. T., & Sevilla, M. D. (2010). Formation of aminyl radicals on electron attachment to AZT: Abstraction from the sugar phosphate backbone versus one-electron oxidation of guanine. *The Journal of Physical Chemistry B*, *114*, 9289.
- Aflatooni, K., Gallup, G. A., & Burrow, P. D. (1998). Electron attachment energies of the DNA bases. *The Journal of Physical Chemistry: A*, *102*, 6205.
- Alizadeh, E., & Sanche, L. (2012). Precursors of solvated electrons in radiobiological physics and chemistry. *Chemical Reviews*, *112*, 5578.
- Allan, M. (1989). Study of triplet states and short-lived negative ions by means of electron impact spectroscopy. *Journal of Electron Spectroscopy and Related Phenomena*, *48*, 219.
- Anusiewicz, I., Berdys, J., Sobczyk, M., Skurski, P., & Simons, J. (2004). Effects of base π -stacking on damage to DNA by low-energy electrons. *The Journal of Physical Chemistry: A*, *108*, 11381.
- Ashfold, M. N. R., Cronin, B., Devine, A. L., Dixon, R. N., & Nix, G. D. (2006). The role of pi sigma* excited states in the photodissociation of heteroaromatic molecules. *Science*, *312*, 1637.
- Baccarelli, I., Gianturco, F. A., Grandi, A., Sanna, N., Lucchese, R. R., Bald, I., et al. (2007). Selective bond breaking in β -D-ribose by gas-phase electron attachment around 8 eV. *Journal of the American Chemical Society*, *129*, 6269.
- Baccarelli, I., Bald, I., Gianturco, F. A., Illenberger, E., & Kopyra, J. (2011). Electron-induced damage of DNA and its components: Experiments and theoretical models. *Physics Reports*, *508*, 1.
- Bachorz, A. R., Klopper, W., & Gutowski, M. (2007). Coupled-cluster and explicitly correlated perturbation-theory calculations of the uracil anion. *The Journal of Chemical Physics*, *126*, 085101.
- Bagshaw, M. A., Dogget, R. L. S., Kendric, C., Smith, K. C., Kaplan, H. S., & Nelsen, T. S. (1967). Intra-arterial 5-bromodeoxyuridine and x-ray therapy. *American Journal of Roentgenology*, *99*, 886.
- Baker, J., Wolinski, K., Malagoli, M., Kinghorn, D., Wolinski, P., Magyarfalvi, G., et al. (2009). Quantum chemistry in parallel with PQS. *Journal of Computational Chemistry*, *30*, 317.
- Bald, I., Kopyra, J., & Illenberger, E. (2006). Selective excision of C5 from D-ribose in the gas phase by low-energy electrons (0–1 eV): Implications for the mechanism of DNA damage. *Angewandte Chemie International Edition*, *45*, 4851.
- Bald, I., Dąbkowska, I., & Illenberger, E. (2008). Probing biomolecules by laser-induced acoustic desorption: Electrons at near zero electron volts trigger sugar-phosphate cleavage. *Angewandte Chemie International Edition*, *47*, 8518.
- Bao, X., Wang, J., Gu, J., & Leszczynski, J. (2006). DNA strand breaks induced by near-zero-electronvolt electron attachment to pyrimidine nucleotides. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 5658.
- Barrios, R., Skurski, P., & Simons, J. (2002). Mechanism for damage to DNA by low-energy electrons. *The Journal of Physical Chemistry B*, *106*, 7991.
- Bartels, D. M. (2001). Moment analysis of hydrated electron cluster spectra: Surface or internal states? *The Journal of Chemical Physics*, *115*, 4404.

- Bartels, D. M., Takahashi, K., Cline, J. A., Marin, T. W., & Jonah, C. D. (2005). Pulse radiolysis of supercritical water. 3. Spectrum and thermodynamics of the hydrated electron. *The Journal of Physical Chemistry A*, *109*, 1299.
- Becker, D., & Sevilla, M. D. (1993). The chemical consequences of radiation damage to DNA. *Advances in Radiation Biology*, *17*, 121.
- Becker, D., & Sevilla, M. D. (2008). EPR studies of radiation damage to DNA and related molecules. *Electron Paramagnetic Resonance*, *21*, 33.
- Becker, D., Razskazovskii, Y., Callaghan, M. U., & Sevilla, M. D. (1996). Electron spin resonance of DNA irradiated with a heavy-ion beam ($^{16}\text{O}^{8+}$): Evidence for damage to the deoxyribose phosphate backbone. *Radiation Research*, *146*, 361.
- Becker, D., Bryant-Friedrich, A., Trzasko, C., & Sevilla, M. D. (2003). Electron spin resonance study of DNA irradiated with an argon-ion beam: Evidence for formation of sugar phosphate backbone radicals. *Radiation Research*, *160*, 174.
- Becker, D., Adhikary, A., & Sevilla, M. D. (2007). The role of charge and spin migration in DNA radiation damage. In T. Chakraborty (Ed.), *Charge migration in DNA* (p. 139). Berlin/Heidelberg: Springer.
- Becker, D., Adhikary, A., & Sevilla, M. D. (2010). Mechanisms of radiation-induced DNA damage: Direct effects. In B. S. M. Rao & J. Wishart (Eds.), *Recent trends in radiation chemistry* (p. 509). Singapore/New Jersey/London: World Scientific Publishing.
- Boag, J. W., & Hart, E. J. (1963). Absorption spectra in irradiated water and some solutions, absorption spectra of 'hydrated' electron. *Nature*, *197*, 45.
- Boudaïffa, B., Cloutier, P., Hunting, D., Huels, M. A., & Sanche, L. (2000). Resonant formation of DNA strand breaks by low-energy (3 to 20 eV) electrons. *Science*, *287*, 1658.
- Brun, E., Cloutier, P., Sicard-Roselli, C., Fromm, M., & Sanche, L. (2009). Damage induced to DNA by low-energy (0–30 eV) electrons under vacuum and atmospheric conditions. *The Journal of Physical Chemistry B*, *113*, 10008.
- Burrow, P. D., Gallup, G. A., Scheer, A. M., Denifl, S., Ptasnińska, S., Märk, T., et al. (2006). Vibrational Feshbach resonances in uracil and thymine. *The Journal of Chemical Physics*, *124*, 124310.
- Burrow, P. D., Gallup, G. A., & Modelli, A. (2008). Are there π^* shape resonances in electron scattering from phosphate groups? *The Journal of Physical Chemistry: A*, *112*, 4106.
- Burrows, C. J., & Muller, J. G. (1998). Oxidative nucleobase modifications leading to strand scission. *Chemical Reviews*, *98*, 1109.
- Buxton, G. V., Greenstock, C. L., Helman, W. P., & Ross, A. B. (1988). Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms and hydroxyl radicals ($\bullet\text{OH}/\bullet\text{O}^-$) in aqueous solution. *Journal of Physical and Chemical Reference Data*, *17*, 513.
- Cadet, J., Douki, T., & Ravanat, J. L. (2008). Oxidatively generated damage to the guanine moiety of DNA: Mechanistic aspects and formation in cells. *Accounts of Chemical Research*, *41*, 1075.
- Caron, L., & Sanche, L. (2003). Low-energy electron diffraction and resonances in DNA and other helical macromolecules. *Physical Review Letters*, *91*, 113201.
- Caron, L., & Sanche, L. (2004). Diffraction in resonant electron scattering from helical macromolecules: A- and B-type DNA. *Physical Review A*, *70*, 032719.
- Caron, L., & Sanche, L. (2005). Diffraction in resonant electron scattering from helical macromolecules: Effects of the DNA backbone. *Physical Review A*, *72*, 032726.
- Caron, L., & Sanche, L. (2006). Temporary electron localization and scattering in disordered single strands of DNA. *Physical Review A*, *73*, 062707.
- Caron, L., Sanche, L., Tonzani, S., & Greene, C. H. (2008). Diffraction in low-energy electron scattering from DNA: Bridging gas-phase and solid-state theory. *Physical Review A*, *78*, 042710.
- Casey, J. R., Kahros, A., & Schwartz, B. J. (2013). To be or not to be in a cavity: The hydrated electron dilemma. *The Journal of Physical Chemistry B*, *117*, 14173.
- Cauët, E., Bogatko, S., Liévin, J., De Profit, F., & Geerlings, P. (2013). Electron-attachment-induced DNA damage: Instantaneous strand breaks. *The Journal of Physical Chemistry B*, *117*, 9669.

- Chen, D., & Gallup, G. A. (1990). The relationship of the virtual orbitals of self-consistent-field theory of temporary negative ions in electron scattering from molecules. *The Journal of Chemical Physics*, *93*, 8893.
- Chen, H.-Y., Yang, P.-Y., Chen, H.-F., Kao, C.-L., & Liao, L.-W. (2014). DFT reinvestigation of DNA strand breaks induced by electron attachment. *The Journal of Physical Chemistry B*, *118*, 11137.
- Chomicz, L., Zdrowowicz, M., Kasprzykowski, F., Rak, J., Buonaugurio, A., Wang, Y., & Bowen, K. H. (2013). How to Find out whether a 5-substituted uracil could be a potential DNA radiosensitizer. *The Journal of Physical Chemistry Letters*, *4*, 2853.
- Chomicz, L., Petrovici, A., Archbold, I., Adhikary, A., Kumar, A., Sevilla, M. D., & Rak, J. (2014). An ESR and DFT study of hydration of the 2'-deoxyuridine-5-yl radical: a possible hydroxyl radical intermediate. *Chemical Communications*, *50*, 14605.
- Ciszek, J. W., Stewart, M. P., & Tour, J. M. (2004). Spontaneous assembly of organic thiocyanates on gold surfaces. Alternative precursors for gold thiolate assemblies. *Journal of the American Chemical Society*, *126*, 13172.
- Coe, J. V. (2001). Fundamental properties of bulk water from cluster ion data. *International Reviews in Physical Chemistry*, *20*, 33.
- Coe, J. V., Lee, G. H., Eaton, J. G., Arnold, S. T., Sarkas, H. W., Bowen, K. H., Ludewigt, C., Haberland, H., & Worsnop, D. R. (1990). Photoelectron spectroscopy of hydrated electron cluster anions, $(\text{H}_2\text{O})_n^-$, $n=2-69$. *The Journal of Chemical Physics*, *92*, 3980.
- Coe, J. V., Williams, S. M., & Bowen, K. H. (2008). Photoelectron spectra of hydrated electron clusters vs. cluster size: Connecting to bulk. *International Reviews in Physical Chemistry*, *27*, 27.
- Colson, A. O., Besler, B., & Sevilla, M. D. (1992). Ab initio molecular orbital calculations on DNA base pair radical ions: Effect of base pairing on proton-transfer energies, electron affinities, and ionization potentials. *The Journal of Physical Chemistry*, *96*, 9787.
- Compton, R. N., Yoshioka, Y., & Jordan, K. D. (1980). Comment on semi-empirical calculations of electron affinities. *Theoretica Chimica Acta*, *54*, 259.
- Crowell, R. A., & Bartels, D. M. (1996). Multiphoton ionization of liquid water with 3.0–5.0 eV photons. *Journal of Physical Chemistry*, *100*, 17940.
- Curtiss, L. A., Redfern, P. C., & Raghavachari, K. (2007). Gaussian-4 theory. *The Journal of Chemical Physics*, *126*, 084108.
- Dabkowska, I., Rak, J., & Gutowski, M. (2005). DNA strand breaks induced by concerted interaction of H radicals and low-energy electrons – A computational study on the nucleotide of cytosine. *European Physical Journal D: Atomic, Molecular, Optical and Plasma Physics*, *35*, 429.
- Denifl, S., Sulzer, P., Huber, D., Zappa, F., Probst, M., Märk, T. D., et al. (2007). Influence of functional groups on the site-selective dissociation of adenine upon low-energy electron attachment. *Angewandte Chemie International Edition*, *46*, 5238.
- Desfancois, C., Abdoul-Carime, H., & Schermann, J. P. (1996). Electron attachment to isolated nucleic acid bases. *The Journal of Chemical Physics*, *104*, 7792.
- Desfancois, C., Periquet, V., Bouteiller, Y., & Schermann, J. P. (1998). Valence and dipole binding of electrons to uracil. *The Journal of Physical Chemistry. A*, *102*, 1274.
- Dessent, C. E. H., Kim, J., & Johnson, M. A. (2000). Spectroscopic observation of vibrational Feshbach resonances in near-threshold photoexcitation of $\text{X}^- \text{CH}_3\text{NO}_2$ ($\text{X}^- = \text{I}^-$ and Br^-). *Faraday Discussion*, *115*, 395.
- Dextraze, M.-E., Wagner, J. R., & Hunting, D. J. (2007). 5-Bromodeoxyuridine radiosensitization: Conformation-dependent DNA damage. *Biochemistry*, *46*, 9089.
- Djordjevic, B., & Szybalski, W. J. (1960). Genetics of human cell lines. III. Incorporation of 5-bromo- and 5-iododeoxyuridine into the deoxyribonucleic acid of human cells and its effect on radiation sensitivity. *The Journal of Experimental Medicine*, *112*, 509.
- Doiron, A., Yapp, D. T. T., Olivares, M., Zhu, J. X. X., & Lehnert, S. (1999). Tumor radiosensitization by sustained intratumoral release of bromodeoxyuridine. *Cancer Research*, *59*, 3677.

- Domcke, W., & Sobolewski, A. L. (2003). Chemistry-unraveling the molecular mechanisms of photoacidity. *Science*, *302*, 1693.
- Dora, A., Tennyson, J., Bryjko, L., & van Mourik, T. (2009). R-matrix calculation of low-energy electron collisions with uracil. *The Journal of Chemical Physics*, *130*, 164307.
- Dreuw, A., & Head-Gordon, M. (2005). Single-reference ab Initio methods for the calculation of excited states of large molecules. *Chemical Reviews*, *105*, 4009.
- Eustis, S., Wang, D., Lyapustina, S., & Bowen, K. H. (2007). Photoelectron spectroscopy of hydrated adenine anions. *The Journal of Chemical Physics*, *127*, 224309.
- Feng, D.-F., & Kevan, L. (1980). Theoretical models for solvated electrons. *Chemical Reviews*, *80*, 1.
- Galduel, Y., Pommeret, S., Migus, A., & Antonetti, A. (1989). Femtosecond dynamics of geminate pair recombination in pure liquid water. *The Journal of Physical Chemistry*, *93*, 3880.
- Gianturco, F. A., & Lucchese, R. R. (2004). Radiation damage of biosystems mediated by secondary electrons: Resonant precursors for uracil molecules. *The Journal of Chemical Physics*, *120*, 7446.
- Gianturco, F. A., Sebastianelli, F., Lucchese, R. R., Baccarelli, I., & Sanna, N. (2008). Ring-breaking electron attachment to uracil: Following bond dissociations via evolving resonances. *The Journal of Chemical Physics*, *128*, 174302.
- Golon, Ł., Chomicz, L., & Rak, J. (2014). Electron-induced single strand break in the nucleotide of 5- and 6-bromouridine. A DFT study. *Chemical Physics Letters*, *612*, 289.
- González-Ramírez, I., Segarra-Martí, J., Serrano-Andrés, L., Merchán, M., Rubio, M., & Roca-Sanjuán, D. (2012). On the N1-H and N3-H bond dissociation in uracil by low energy electrons: A CASSCF/CASPT2 study. *Journal of Chemical Theory and Computation*, *8*, 2769.
- Gregoli, S., Olast, M., & Bertinchamps, A. (1989). Radiolytic pathways in γ -irradiated DNA: Influence of chemical and conformational factors. *Radiation Research*, *89*, 238.
- Gu, J., Xie, Y., & Schaefer, H. F. (2005). Glycosidic bond cleavage of pyrimidine nucleosides by low-energy electrons: A theoretical rationale. *Journal of the American Chemical Society*, *127*, 1053.
- Gu, J., Wang, J., & Leszczynski, J. (2006). Electron attachment-induced DNA single strand breaks: C₃-O_{3'} σ -bond breaking of pyrimidine nucleotides predominates. *Journal of the American Chemical Society*, *128*, 9322.
- Gu, J. D., Wang, J., & Leszczynski, J. (2010a). Electron attachment-induced DNA single-strand breaks at the pyrimidine sites. *Nucleic Acids Research*, *38*, 5280.
- Gu, J., Wang, J., & Leszczynski, J. (2010b). Comprehensive analysis of DNA strand breaks at the guanosine site induced by low-energy electron attachment. *Chemphyschem*, *11*, 175.
- Gu, J., Wang, J., & Leszczynski, J. (2011). Low energy electron attachment to the adenosine site of DNA. *The Journal of Physical Chemistry B*, *115*, 14831.
- Gu, J., Leszczynski, J., & Schaefer, H. F. (2012). Interactions of electrons with bare and hydrated biomolecules: From nucleic acid bases to DNA segments. *Chemical Reviews*, *112*, 5603.
- Gu, B., Smyth, M., & Kohanoff, J. (2014a). Protection of DNA against low-energy electrons by amino acids: a first-principles molecular dynamics study. *Physical Chemistry Chemical Physics*, *16*, 24350.
- Gu, J., Xie, Y., & Schaefer, H. F. (2014b). Benchmarking the electron affinity of uracil. *Journal of Chemical Theory and Computation*, *10*, 609.
- Hamed, E. M., Doai, H., McLaughlin, C. K., & Houmam, A. (2006). Regioselective bond cleavage in the dissociative electron transfer to benzyl thiocyanates: The role of radical/ion pair formation. *Journal of the American Chemical Society*, *128*, 6595.
- Hammer, N. I., Shin, J.-W., Headrick, J. M., Diken, E. G., Roscioli, J. R., Weddle, G. H., & Johnson, M. A. (2004). How do small water clusters bind an excess electron? *Science*, *306*, 675.
- Han, P., & Bartels, D. M. (1990). Reevaluation of arrhenius parameters for $\text{H}\cdot + \text{OH}^- \rightarrow (\text{e}^-)_{\text{aq}} + \text{H}_2\text{O}$ and the enthalpy and entropy of hydrated electrons. *The Journal of Physical Chemistry*, *94*, 7294.

- Hanel, G., Gstir, B., Denifl, S., Scheier, P., Probst, M., Farizon, B., et al. (2003). Electron attachment to uracil: Effective destruction at subexcitation energies. *Physical Review Letters*, *90*, 188104.
- Haranczyk, M., & Gutowski, M. (2005). Valence and dipole-bound anions of the most stable tautomers of guanine. *Journal of the American Chemical Society*, *127*, 699.
- Hart, E. J., & Boag, J. W. (1962). Absorption spectrum of the hydrated electron in water and in aqueous solutions. *Journal of the American Chemical Society*, *84*, 4090.
- Hashiya, F., Saha, A., Kizaki, S., Li, Y., & Sugiyama, H. (2014). Locating the uracil-5-yl radical formed upon photoirradiation of 5-bromouracil-substituted DNA. *Nucleic Acids Research*. doi:10.1093/nar/gku1133.
- Hendricks, J. H., Lyapustina, S. A., de Clercq, H. L., Snodgrass, J. T., & Bowen, K. H. (1996). Dipole bound, nucleic acid base anions studied via negative ion photoelectron spectroscopy. *The Journal of Chemical Physics*, *104*, 7788.
- Hendricks, J. H., Lyapustina, S. A., de Clercq, H. L., & Bowen, K. H. (1998). The dipole bound-to-covalent anion transformation in uracil. *The Journal of Chemical Physics*, *108*, 8.
- Herbert, J. M., & Head-Gordon, M. (2005). Calculation of electron detachment energies for water cluster anions: An appraisal of electronic structure methods, with application to $(\text{H}_2\text{O})_{20}^-$ and $(\text{H}_2\text{O})_{24}^-$. *The Journal of Physical Chemistry A*, *109*, 5217.
- Hickel, B., & Sehested, K. (1985). Activation energy for the reaction $\text{H} + \text{OH}^- \rightarrow \text{e}_{\text{aq}}^-$. kinetic determination of the enthalpy and entropy of solvation of the hydrated electron. *The Journal of Physical Chemistry*, *89*, 5271.
- Hotop, H., Ruf, M. W., Allan, M., & Fabrikant, I. I. (2003). Resonance and threshold phenomena in low-energy electron collisions with molecules and clusters. *Advances in Atomic, Molecular, and Optical Physics*, *49*, 85.
- Houmam, A. (2008). Electron transfer initiated reactions: Bond formation and bond dissociation. *Chemical Reviews*, *108*, 2180.
- Houmam, A., Hamed, E. M., Hapiot, P., Motto, J. M., & Schwan, A. L. (2003a). Regioselective bond cleavage in the dissociative electron transfer to benzyl thiocyanates. *Journal of the American Chemical Society*, *125*, 12676.
- Houmam, A., Hamed, E. M., & Still, I. W. J. (2003b). A unique autocatalytic process and evidence for a concerted-stepwise mechanism transition in the dissociative electron-transfer reduction of aryl thiocyanates. *Journal of the American Chemical Society*, *125*, 7258.
- Huels, M. A., Boudaïffa, B., Cloutier, P., Hunting, D., & Sanche, L. (2003). Single, double, and multiple double strand breaks induced in DNA by 3–100 eV electrons. *Journal of the American Chemical Society*, *125*, 4467.
- Hush, N. S., & Cheung, A. S. (1975). Ionization potentials and donor properties of nucleic acid bases and related compounds. *Chemical Physics Letters*, *34*, 11.
- Illenberger, E. (1992). Electron-attachment reactions in molecular clusters. *Chemical Reviews*, *92*, 1589.
- International Commission on Radiation Units and Measurements. (1979). ICRU Report No. 31. ICRU, Washington, DC.
- Jacobson, L. D., & Herbert, J. M. (2011). Comment on “Does the Hydrated Electron Occupy a Cavity?”. *Science*, *331*, 1387.
- Jordan, K. D. (2004). A fresh look at electron hydration. *Science*, *306*, 618.
- Jordan, K. D., & Burrow, P. D. (1987). Temporary anion states of polyatomic hydrocarbons. *Chemical Reviews*, *87*, 557.
- Jortner, J., & Noyes, R. M. (1966). Some thermodynamic properties of the hydrated electron. *The Journal of Physical Chemistry*, *70*, 770.
- Kevan, L. (1981). Solvated electron structure in glassy matrices. *Accounts of Chemical Research*, *14*, 138.
- Khanduri, D., Collins, S., Kumar, A., Adhikary, A., & Sevilla, M. D. (2008). Formation of sugar radicals in RNA model systems and oligomers via excitation of guanine cation radical. *The Journal of Physical Chemistry B*, *112*, 2168.

- Kheir, J., Chomicz, L., Engle, A., Rak, J., & Sevilla, M. D. (2013). Presolvated low energy electron attachment to peptide methyl esters in aqueous solution: C–O bond cleavage at 77 K. *The Journal of Physical Chemistry B*, *117*, 2872.
- Kobyłecka, M., Leszczynski, J., & Rak, J. (2008). Valence anion of thymine in the DNA π -stack. *Journal of the American Chemical Society*, *130*, 15683.
- König, C., Kopyra, J., Bald, I., & Illenberger, E. (2006). Dissociative electron attachment to phosphoric acid esters: The direct mechanism for single strand breaks in DNA. *Physical Review Letters*, *97*, 018105.
- Kopyra, J. (2012). Low energy electron attachment to the nucleotide deoxycytidine monophosphate: direct evidence for the molecular mechanisms of electron-induced DNA strand breaks. *Physical Chemistry Chemical Physics*, *14*, 8287.
- Kumar, A., & Sevilla, M. D. (2007). Low-energy electron attachment to 5'-thymidine monophosphate: Modeling single strand breaks through dissociative electron attachment. *The Journal of Physical Chemistry B*, *111*, 5464.
- Kumar, A., & Sevilla, M. D. (2008a). Radiation effects on DNA: Theoretical investigations of electron, hole and excitation pathways to DNA damage. In M. K. Shukla & J. Leszczynski (Eds.), J. Leszczynski (Series Ed.), *Radiation induced molecular phenomena in nucleic acids: Vol. 5. Challenges and advances in computational chemistry and physics* (p. 577). Amsterdam: Springer.
- Kumar, A., & Sevilla, M. D. (2008b). The role of $\pi\sigma^*$ excited states in electron-induced DNA strand break formation: A time-dependent density functional theory study. *Journal of the American Chemical Society*, *130*, 2130.
- Kumar, A., & Sevilla, M. D. (2009). Role of excited states in low-energy electron (LEE) induced strand breaks in DNA model systems: Influence of aqueous environment. *Chem Phys Chem*, *10*, 1426.
- Kumar, A., & Sevilla, M. D. (2010a). Proton-coupled electron transfer in DNA on formation of radiation-produced ion radicals. *Chemical Reviews*, *110*, 7002.
- Kumar, A., & Sevilla, M. D. (2010b). Theoretical modeling of radiation-induced DNA damage. In M. Greenberg (Ed.), *Radical and radical ion reactivity in nucleic acid chemistry* (p. 1). Hoboken: Wiley. ISBN 978-0-470-25558-2.
- Kumar, A., Walker, J., Bartels, D., & Sevilla, M. D. (2015). A simple ab initio model for the hydrated electron that matches experiment. *The Journal of Physical Chemistry A*, *119*, 9148.
- Lane, N. F. (1980). The theory of electron-molecule collisions. *Reviews of Modern Physics*, *52*, 29.
- Larsen, R. E., Glover, W. J., & Schwartz, B. J. (2010). Does the hydrated electron occupy a cavity? *Science*, *329*, 65.
- Levesque, P. L., Michaud, M., & Sanche, L. (2003). Cross sections for low energy (1–12 eV) inelastic electron scattering from condensed thymine. *Nuclear Instruments and Methods in Physics Research B*, *208*, 225.
- Li, X., & Sevilla, M. D. (2007). DFT treatment of radiation produced radicals in DNA model systems. *Advances in Quantum Chemistry*, *52*, 59.
- Li, X. F., Cai, Z. L., & Sevilla, M. D. (2002a). DFT calculations of the electron affinities of nucleic acid bases: Dealing with negative electron affinities. *The Journal of Physical Chemistry. A*, *106*, 1596.
- Li, X., Sanche, L., & Sevilla, M. D. (2002b). Dehalogenation of 5-halouracils after low energy electron attachment: A density functional theory investigation. *The Journal of Physical Chemistry. A*, *106*, 11248.
- Li, X., Sevilla, M. D., & Sanche, L. (2003a). Density functional theory studies of electron interaction with DNA: Can zero eV electrons induce strand breaks? *Journal of the American Chemical Society*, *125*, 13668.
- Li, X., Sevilla, M. D., & Sanche, L. (2003b). DFT Investigation of dehalogenation of adenine-halouracil base pairs upon low-energy electron attachment. *Journal of the American Chemical Society*, *125*, 8916.

- Li, X. F., Sevilla, M. D., & Sanche, L. (2004). Hydrogen atom loss in pyrimidine DNA bases induced by low-energy electrons: Energetics predicted by theory. *The Journal of Physical Chemistry B*, *108*, 19013.
- Li, X., Sanche, L., & Sevilla, M. D. (2006). Base release in nucleosides induced by low-energy electrons: A DFT study. *Radiation Research*, *165*, 721.
- Li, Z., Zheng, Y., Cloutier, P., Sanche, L., & Wagner, J. R. (2008). Low energy electron induced DNA damage: Effects of terminal phosphate and base moieties on the distribution of damage. *Journal of the American Chemical Society*, *130*, 5612.
- Li, Z., Cloutier, P., Sanche, L., & Wagner, J. R. (2010). Low-energy electron-induced DNA damage effect of base sequence in oligonucleotide trimers. *Journal of the American Chemical Society*, *132*, 5422.
- Li, Z., Cloutier, P., Sanche, L., & Wagner, J. R. (2011). Low-energy electron-induced damage in a trinucleotide containing 5-bromouracil. *The Journal of Physical Chemistry B*, *115*, 13668.
- Ling, L. L., & Ward, J. F. (1990). Radiosensitization of Chinese-hamster v79 cells by bromodeoxyuridine substitution of thymidine-enhancement of radiation-induced toxicity and DNA strand break production by monofilar and bifilar substitution. *Radiation Research*, *121*, 76.
- Loos, P. F., Dumont, E., Laurent, A. D., & Assfeld, X. (2009). Important effects of neighbouring nucleotides on electron induced DNA single-strand breaks. *Chemical Physics Letters*, *475*, 120.
- Lübcke, A., Buchner, F., Heine, N., Hertel, I. V., & Schultz, T. (2010). Time-resolved photoelectron spectroscopy of solvated electrons in aqueous NaI solution. *Physical Chemistry Chemical Physics*, *12*, 14629.
- Madarász, Á., Rossky, P. J., & Turi, L. (2009). Interior- and surface-bound excess electron states in large water cluster anions. *The Journal of Chemical Physics*, *130*, 124319.
- Marsalek, O., Uhlig, F., Vandevondele, J., & Jungwirth, P. (2012). Structure, dynamics, and reactivity of hydrated electrons by ab initio molecular dynamics. *Accounts of Chemical Research*, *45*, 23.
- Martin, F., Burrow, P. D., Cai, Z., Cloutier, P., Hunting, D., & Sanche, L. (2004). DNA strand breaks induced by 0–4 eV electrons: The role of shape resonances. *Physical Review Letters*, *93*, 068101.
- McConkey, J. W., Malone, C. P., Johnson, P. V., Winstead, C., McKoy, V., & Kanik, I. (2008). Electron impact dissociation of oxygen-containing molecules – A critical review. *Physics Reports*, *466*, 1.
- Modelli, A. (2003). Electron attachment and intramolecular electron transfer in unsaturated chloroderivatives. *Physical Chemistry Chemical Physics*, *5*, 2923.
- Modelli, A., & Martin, H. D. (2002). Temporary anions and empty level structure in cyclobutanediones: Through-space and through-bond interactions. *The Journal of Physical Chemistry. A*, *106*, 7271.
- Nabben, F. J., Karman, J. P., & Loman, H. (1982). Inactivation of biologically active DNA by hydrated electrons. *International Journal of Radiation Biology*, *42*, 23.
- Neese, F., Hansen, A., Wennmohs, F., & Grimme, S. (2009). Accurate theoretical chemistry with coupled pair models. *Accounts of Chemical Research*, *42*, 641.
- Orlando, T. M., Oh, D., Chen, Y., & Aleksandrov, A. B. (2008). Low-energy electron diffraction and induced damage in hydrated DNA. *The Journal of Chemical Physics*, *128*, 195102.
- Orlov, V. M., Smirnov, A. N., & YaM, V. (1976). Ionization potentials and electron-donor ability of nucleic acid bases and their analogues. *Tetrahedron Letters*, *17*, 4315.
- Oyler, N. A., & Adamowicz, L. (1993). Electron attachment to uracil. Theoretical ab initio study. *The Journal of Physical Chemistry*, *97*, 11122.
- Oyler, N. A., & Adamowicz, L. (1994). Theoretical ab initio calculations of the electron affinity of thymine. *Chemical Physics Letters*, *219*, 223.
- Panajotovic, R., Martin, F., Cloutier, P. C., Hunting, D., & Sanche, L. (2006). Effective cross sections for production of single-strand breaks in plasmid DNA by 0.1 to 4.7 eV electrons. *Radiation Research*, *165*, 452.

- Park, Y., Li, Z., Cloutier, P., Sanche, L., & Wagner, J. R. (2011). DNA damage induced by low-energy electrons: Conversion of thymine to 5,6-dihydrothymine in the oligonucleotide trimer TpTpT. *Radiation Research*, 175, 240.
- Park, Y., Polska, K., Rak, J., Wagner, J. R., & Sanche, L. (2012). Fundamental mechanisms of DNA radiosensitization: Damage induced by low-energy electrons in brominated oligonucleotide trimers. *The Journal of Physical Chemistry B*, 116, 9676.
- Park, Y., Peoples, A. R., Madugundu, G. S., Sanche, L., & Wagner, J. R. (2013). Side-by-side comparison of DNA damage induced by low-energy electrons and high-energy photons with solid TpTpT trinucleotide. *The Journal of Physical Chemistry B*, 117, 10122.
- Periquet, V., Moreau, A., Carles, S., Schermann, J. P., & Desfrançois, C. (2000). Cluster size effects upon anion solvation of N-heterocyclic molecules and nucleic acid bases. *Journal of Electron Spectroscopy and Related Phenomena*, 106, 141.
- Petrovici, A., Adhikary, A., Kumar, A., & Sevilla, M. D. (2014). Presolvated electron reactions with methyl acetoacetate: Electron localization, proton-deuteron exchange, and h-atom abstraction. *Molecules*, 19, 13486.
- Pimblott, S. M., & LaVerne, J. A. (2007). Production of low-energy electrons by ionizing radiation. *Radiation Physics and Chemistry*, 76, 1244.
- Pimblott, S. M., Laverne, J. A., & Mozumber, A. (1996). Monte Carlo simulation of range and energy deposition by electrons in gaseous and liquid water. *The Journal of Physical Chemistry*, 100, 8595.
- Ptasińska, S., & Sanche, L. (2007a). Dissociative electron attachment to a basic DNA. *Physical Chemistry Chemical Physics*, 9, 1730.
- Ptasińska, S., & Sanche, L. (2007b). Dissociative electron attachment to hydrated single DNA strands. *Physical Review E*, 75, 031915.
- Ptasińska, S., Denifl, S., Scheier, P., Illenberger, E., & Märk, T. D. (2005a). Bond- and site-selective loss of H atoms from nucleobases by very-low-energy electrons (<3 eV). *Angewandte Chemie International Edition*, 44, 6941.
- Ptasińska, S., Denifl, S., Grill, V., Märk, T. D., Illenberger, E., & Scheier, P. (2005b). Bond- and site-selective loss of H[⋅] from pyrimidine bases. *Physical Review Letters*, 95, 093201.
- Ptasińska, S., Denifl, S., Grill, V., Märk, T. D., Scheier, P., Gohlke, S., Huels, M. A., & Illenberger, E. (2005c). Bond-selective H⁻ ion abstraction from thymine. *Angewandte Chemie International Edition*, 44, 1647.
- Ptasińska, S., Denifl, S., Gohlke, S., Scheier, P., Illenberger, E., & Märk, T. D. (2006). Decomposition of thymidine by low-energy electrons: Implications for the molecular mechanisms of single-strand breaks in DNA. *Angewandte Chemie International Edition*, 45, 1893.
- Puiatti, M., Vera, D. M. A., & Pierini, A. B. (2009). In search for an optimal methodology to calculate the valence electron affinities of temporary anions. *Physical Chemistry Chemical Physics*, 11, 9013.
- Rak, J., Mazurkiewicz, K., Kobylecka, M., Storoniak, P., Haranczyk, M., Dąbkowska, I., Bachorz, R. A., Gutowski, M., Radisic, D., Stokes, S. T., Eustis, S. N., Wang, D., Li, X., Ko, Y. J., & Bowen, K. H. (2008). Stable valence anions of nucleic acid bases and DNA strand breaks induced by low energy electrons. In M. K. Shukla & J. Leszczynski (Eds.), *J. Leszczynski (Series Ed.), Radiation induced molecular phenomena in nucleic acids: Vol. 5. Challenges and advances in computational chemistry and physics* (p. 619). Amsterdam: Springer.
- Rak, J., Kobylecka, M., & Storoniak, P. (2011). Single strand break in DNA coupled to the O–P bond cleavage. A computational study. *The Journal of Physical Chemistry B*, 115, 1911.
- Razskazovskii, Y., Swarts, S. G., Falcone, J. M., Taylor, C., & Sevilla, M. D. (1997). Competitive electron scavenging by chemically modified pyrimidine bases in bromine-doped DNA: Relative efficiencies and relevance to intrastrand electron migration distances. *The Journal of Physical Chemistry B*, 101, 1460.
- Riederer, H., Hüttermann, J., & Symons, M. C. R. (1978). Sigma-star-electron addition to 5-halogenouracils in neutral glasses. *Journal of the Chemical Society, Chemical Communications*, 7, 313.

- Ryabukhin, A. G. (1977). The standard entropy of the electron in aqueous solution. *Russian Journal of Physical Chemistry*, 51, 573; translated from *Zhurnal Fizicheskoi Khimii*, 1977, 51, 968.
- Sanche, L. (1995). Interactions of low-energy electrons with atomic and molecular-solids. *Scanning Microscopy*, 9, 619.
- Sanche, L. (2005). Low energy electron-driven damage in biomolecules. *European Physical Journal D: Atomic, Molecular, Optical and Plasma Physics*, 35, 367.
- Sanche, L. (2008) Low energy electron damage to DNA. In M. K. Shukla & J. Leszczynski (Eds.), J. Leszczynski (Series Ed.), *Radiation induced molecular phenomena in nucleic acids: Vol. 5. Challenges and advances in computational chemistry and physics* (p. 531). Amsterdam: Springer.
- Sanche, L. (2009). Role of secondary low energy electrons in radiobiology and chemoradiation therapy of cancer. *Chemical Physics Letters*, 474, 1.
- Sanche, L. (2010). Low-energy electron interaction with DNA: Bond dissociation and formation of transient anions, radicals, and radical anions. In M. Greenberg (Ed.), *Radical and radical ion reactivity in nucleic acid chemistry* (p. 239). Hoboken: Wiley. ISBN 978-0-470-25558-2.
- Scheer, A. M., Aflatooni, K., Gallup, G. A., & Burrow, P. D. (2004). Bond breaking and temporary anion states in uracil and halouracils: Implications for the DNA bases. *Physical Review Letters*, 92, 068102.
- Schiedt, J., Weinkauff, R., Neumark, D. M., & Schlag, E. W. (1998). Anion spectroscopy of uracil, thymine and the amino-oxo and amino-hydroxy tautomers of cytosine and their water clusters. *Chemical Physics*, 239, 511.
- Schulz, G. J. (1973a). Resonances in electron impact on atoms. *Reviews of Modern Physics*, 45, 378.
- Schulz, G. J. (1973b). Resonances in electron impact on diatomic molecules. *Reviews of Modern Physics*, 45, 423.
- Schwabe, T., & Grimme, S. (2008). Theoretical thermodynamics for large molecules: Walking the thin line between accuracy and computational cost. *Accounts of Chemical Research*, 41, 569.
- Schwarz, H. A. (1981). Free radicals generated by radiolysis of aqueous solutions. *Journal of Chemical Education*, 58, 101.
- Schwarz, H. A. (1991). Enthalpy and entropy of formation of the hydrated electron. *The Journal of Physical Chemistry*, 95, 6697.
- Schyman, P., & Laaksonen, A. (2008). On the effect of low-energy electron induced DNA strand break in aqueous solution: A Theoretical study indicating guanine as a weak link in DNA. *Journal of the American Chemical Society*, 130, 12254.
- Schyman, P., Laaksonen, A., & Hugosson, H. W. (2008). Phosphodiester bond rupture in 5' and 3' cytosine monophosphate in aqueous environment and the effect of low-energy electron attachment: A Car–Parrinello QM/MM molecular dynamics study. *Chemical Physics Letters*, 462, 289.
- Seidel, C. A. M., Schulz, A., & Sauer, M. H. M. (1996). Nucleobase-specific quenching of fluorescent dyes. 1. Nucleobase one-electron redox potentials and their correlation with static and dynamic quenching efficiencies. *The Journal of Physical Chemistry*, 100, 5541.
- Sevilla, M. D., & Bernhard, W. A. (2008). Mechanisms of direct radiation damage to DNA. In M. Spothem-Maurizot, M. Mostafavi, T. Douki, & J. Belloni (Eds.), *Mechanisms of direct radiation damage to DNA. In radiation chemistry from basics to applications in material and life sciences* (p. 191). Cedex A, France: EDP Sciences.
- Sevilla, M. D., Failor, R., & Zorman, G. (1974). Radicals formed after electron attachment to 5-halouracils in aqueous glasses. *The Journal of Physical Chemistry*, 78, 696.
- Sevilla, M. D., Becker, D., Yan, M., & Summerfield, S. R. (1991). Relative abundances of primary ion radicals in γ -irradiated DNA: Cytosine vs. thymine anions and guanine vs. adenine cations. *The Journal of Physical Chemistry*, 95, 3409.
- Sevilla, M. D., Besler, B., & Colson, A. O. (1994). Comment on “electron attachment to uracil. Theoretical *ab Initio* study.”. *The Journal of Physical Chemistry*, 98, 2215.

- Sevilla, M. D., Besler, B., & Colson, A. O. (1995). Ab initio molecular orbital calculations of DNA radical ions. 5. Scaling of calculated electron affinities and ionization potentials to experimental values. *The Journal of Physical Chemistry*, *99*, 1060.
- Shao, Y., Molnar, L. F., Jung, Y., Kussmann, J., Ochsenfeld, C., Brown, S. T., et al. (2006). Advances in methods and algorithms in a modern quantum chemistry program package. *Physical Chemistry Chemical Physics*, *8*, 3172.
- Shkrob, I. A. (2007). The structure of the hydrated electron. Part 1. Magnetic resonance of internally trapping water anions: A density functional theory study. *The Journal of Physical Chemistry A*, *111*, 5223.
- Shreve, A. T., Yen, T. A., & Neumark, D. M. (2010). Photoelectron spectroscopy of hydrated electrons. *Chemical Physics Letters*, *493*, 216.
- Shukla, L. I., Pazdro, R., Huang, J., DeVreugd, C., Becker, D., & Sevilla, M. D. (2004). The formation of DNA sugar radicals from photoexcitation of guanine cation radicals. *Radiation Research*, *161*, 582.
- Shukla, L. I., Pazdro, R., Becker, D., & Sevilla, M. D. (2005). Sugar radicals in DNA: Isolation of neutral radicals in gamma-irradiated DNA by hole and electron scavenging. *Radiation Research*, *163*, 591.
- Siefermann, K. R., & Abel, B. (2011). The hydrated electron: A seemingly familiar chemical and biological transient. *Angewandte Chemie International Edition*, *50*, 5264.
- Siefermann, K. R., Liu, Y., Lugovoy, E., Link, O., Faubel, M., Buck, U., Winter, B., & Abel, B. (2010). Binding energies, lifetimes and implications of bulk and interface solvated electrons in water. *Nature Chemistry*, *2*, 274.
- Simons, J. (2006). How do low-energy (0.1–2 eV) electrons cause DNA-strand breaks? *Accounts of Chemical Research*, *39*, 772.
- Simons, J. (2008). Molecular anions. *The Journal of Physical Chemistry. A*, *112*, 6401.
- Simons, J., & Jordan, K. D. (1987). Ab Initio electronic structure of anions. *Chemical Reviews*, *87*, 535.
- Smyth, M., & Kohanoff, J. (2012). Excess electron interactions with solvated DNA nucleotides: Strand breaks possible at room temperature. *Journal of the American Chemical Society*, *134*, 9122.
- Sobolewski, A. L., & Domcke, W. (2002). On the mechanism of nonradiative decay of DNA bases: Ab initio and TDDFT results for the excited states of 9H-adenine. *European Physical Journal D: Atomic, Molecular, Optical and Plasma Physics*, *20*, 369.
- Solomon, T., & Skalicky, T. (2008). The interaction of a protein–DNA surface complex with low-energy electrons. *Chemical Physics Letters*, *453*, 101.
- Solomon, T., Seitz, H., & Sturm, H. (2009). DNA damage by low-energy electron impact: Dependence on guanine content. *The Journal of Physical Chemistry B*, *113*, 11557.
- Staiey, S. W., & Strnad, J. T. (1994). Calculation of the energies of π^* negative ion resonance states by the use of Koopmans' theorem. *The Journal of Physical Chemistry*, *98*, 116.
- Steenken, S., & Jovanovic, S. V. (1997). How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution. *Journal of the American Chemical Society*, *119*, 617.
- Sugiyama, H., Fujimoto, K., Saito, I., Kawashima, E., Sekine, T., & Ishido, Y. (1996). Evidence for intrastrand C2' hydrogen abstraction in photoirradiation of 5-halouracil-containing oligonucleotides by using stereospecifically C2'-deuterated deoxyadenosine. *Tetrahedron Letters*, *37*, 1805.
- Sulzer, P., Ptasnińska, S., Zappa, F., Mielewska, B., Milosavljevic, A. R., Scheier, P., et al. (2006). Dissociative electron attachment to furan, tetrahydrofuran, and fructose. *The Journal of Chemical Physics*, *125*, 044304.
- Svozil, D., Frigato, T., Havlas, Z., & Jungwirth, P. (2005). Ab initio electronic structure of thymine anions. *Physical Chemistry Chemical Physics*, *7*, 840.
- Swarts, S. G., Sevilla, M. D., Becker, D., Tokar, C. J., & Wheeler, K. T. (1992). Radiation-induced DNA damage as a function of hydration. I. Release of unaltered bases. *Radiation Research*, *129*, 333.

- Swiderek, P. (2006). Fundamental processes in radiation damage of DNA. *Angewandte Chemie International Edition*, 45, 4056.
- Takayanagi, T., Asakura, T., & Motegi, H. (2009). Theoretical study on the mechanism of low-energy dissociative electron attachment for uracil. *The Journal of Physical Chemistry A*, 113, 4795.
- Tang, Y., Shen, H., Sekiguchi, K., Kurahashi, N., Mizuno, T., Suzuki, Y.-I., & Suzuki, T. (2010). Direct measurement of vertical binding energy of a hydrated electron. *Physical Chemistry Chemical Physics*, 12, 3653.
- Tashiro, R., & Sugiyama, H. (2010) Reactivity of 5-halopyrimidines in nucleic acids. In: M. Greenberg (Ed.), *Radical and radical ion reactivity in nucleic acid chemistry* (p 163). Wiley, Hoboken, New Jersey ISBN: 978-0-470-25558-2.
- Tonzani, S., & Greene, C. H. (2006a). Low-energy electron scattering from DNA and RNA bases: Shape resonances and radiation damage. *The Journal of Chemical Physics*, 124, 054312.
- Tonzani, S., & Greene, C. H. (2006b). Radiation damage to DNA: Electron scattering from the backbone subunits. *The Journal of Chemical Physics*, 125, 094504.
- Turi, L., & Madarász, Á. (2011). Comment on “Does the Hydrated Electron Occupy a Cavity?”. *Science*, 331, 1387.
- Uhlig, F., Marsalek, O., & Jungwirth, P. (2012). Unraveling the complex nature of the hydrated electron. *The Journal of Physical Chemistry Letters*, 3, 3071.
- Vera, D. M. A., & Pierini, A. B. (2004). Species with negative electron affinity and standard DFT methods. *Physical Chemistry Chemical Physics*, 6, 2899.
- Verlet, J. R. R., Bragg, A. E., Kammrath, A., Cheshnovsky, O., & Neumark, D. M. (2005). Observation of large water-cluster anions with surface-bound excess electrons. *Science*, 307, 93.
- Voit, K., Oloff, H., Hüttermann, J., Köhnlein, W., Gräslund, A., & Rupprecht, A. (1986). Anion-radicals from 5-halouracil substituted oriented DNA after X-irradiation at low-temperatures. *Radiation and Environmental Biophysics*, 25, 175.
- von Sonntag, C. (1987). *The chemical basis of radiation biology*. London: Taylor and Francis.
- von Sonntag, C. (1991). The chemistry of free-radical-mediated DNA damage. In W. A. Glass & M. N. Varma (Eds.), *Physical and chemical mechanisms in molecular radiation biology* (p. 287). New York: Plenum.
- Vysotskiy, V. P., Cederbaum, L. S., Sommerfeld, T., Voora, V. K., & Jordan, K. D. (2012). Benchmark calculations of the energies for binding excess electrons to water clusters. *Journal of Chemical Theory and Computation*, 8, 893.
- Wagenknecht, H. A. (Ed.). (2005). *Charge transfer in DNA. From mechanism to application*. New York: Wiley-VCH, Verlag GmbH.
- Wang, C. R., & Lu, Q. B. (2007). Real-time observation of a molecular reaction mechanism of aqueous 5-halo-2'-deoxyuridines under UV/ionizing radiation. *Angewandte Chemie International Edition*, 46, 6316.
- Wang, C. R., & Lu, Q. B. (2010). Molecular mechanism of the DNA sequence selectivity of 5-halo-2'-deoxyuridines as potential radiosensitizers. *Journal of the American Chemical Society*, 132, 14710.
- Wang, C. R., Hu, A., & Lu, Q.-B. (2006). Direct observation of the transition state of ultrafast electron transfer reaction of a radiosensitizing drug bromodeoxyuridine. *The Journal of Chemical Physics*, 124, 241102.
- Wang, C. R., Nguyen, J., & Lu, Q. B. (2009). Bond breaks of nucleotides by dissociative electron transfer of nonequilibrium prehydrated electrons: A new molecular mechanism for reductive DNA damage. *Journal of the American Chemical Society*, 131, 11320.
- Wardman, P. (1989). Reduction potentials of one-electron couples involving free radicals in aqueous solution. *Journal of Physical and Chemical Reference Data*, 18, 1637.
- Webb, C. F., Jones, G. D. D., Ward, J. F., Moyer, D. J., Aguilera, J. A., & Ling, L. L. (1993). Mechanisms of radiosensitization in bromodeoxyuridine-substituted cells. *International Journal of Radiation Biology*, 64, 695.

- Wesolowski, S. S., Leininger, M. L., Pentchev, P. N., & Schaefer, H. F. (2001). Electron affinities of the DNA and RNA bases. *Journal of the American Chemical Society*, *123*, 4023.
- Wetmore, S. D., Boyd, R. J., & Eriksson, L. A. (2000). Electron affinities and ionization potentials of nucleotide bases. *Chemical Physics Letters*, *322*, 129.
- Wetzel, D. M., & Brauman, J. I. (1987). Electron photodetachment spectroscopy of trapped negative ions. *Chemical Reviews*, *87*, 607.
- Wiczak, J., Miloch, J., & Rak, J. (2013). UV-induced strand breaks in double-stranded DNA labeled with 5-bromouracil: Frank or secondary? *The Journal of Physical Chemistry Letters*, *4*, 4014.
- Wieczór, M., Wityk, P., Czub, J., Chomicz, L., & Rak, J. (2014). A first-principles study of electron attachment to the fully hydrated bromonucleobases. *Chemical Physics Letters*, *595*, 133.
- Winstead, C., & McKoy, V. (2006). Interaction of low-energy electrons with the purine bases, nucleosides, and nucleotides of DNA. *The Journal of Chemical Physics*, *125*, 244302.
- Winstead, C., & McKoy, V. (2008). Resonant interactions of slow electrons with DNA constituents. *Radiation Physics and Chemistry*, *77*, 1258.
- Xie, H. J., Wu, R. B., Xia, F., & Cao, Z. X. (2008). Effects of electron attachment on C-5'-O-5' and C-1'-N-1 bond cleavages of pyrimidine nucleotides: A theoretical study. *Journal of Computational Chemistry*, *29*, 2025.
- Yalunin, S., & Leble, S. B. (2007). Multiple-scattering and electron-uracil collisions at low energies. *European Physical Journal*, *144*, 115.
- Yang, X., Wang, X. B., Vorpapel, E. R., & Wang, L. S. (2004). Direct experimental observation of the low ionization potentials of guanine in free oligonucleotides by using photoelectron spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 17588.
- Yokoya, A., Shikazono, N., Fujii, K., Urushibara, A., Akamatsu, K., & Watanabe, R. (2008). DNA damage induced by the direct effect of radiation. *Radiation Physics and Chemistry*, *77*, 1280.
- Young, R. M., & Neumark, D. M. (2012). Dynamics of solvated electrons in clusters. *Chemical Reviews*, *112*, 5553.
- Younkin, J. M., Smith, L. J., & Compton, R. N. (1976). Semi-empirical calculations of p-electron affinities for some conjugated organic molecules. *Theoretica Chimica Acta*, *41*, 157.
- Zdrowowicz, M., Michalska, B., Zyllich-Stachula, A., & Rak, J. (2014). Photoinduced single strand breaks and intrastrand cross-links in an oligonucleotide labeled with 5-bromouracil. *The Journal of Physical Chemistry B*, *118*, 5009.
- Zdrowowicz, M., Chomicz, L., Żyndul, M., Wityk, P., Rak, J., Wiegand, T. J., Hanson, C. G., Adhikary, A., & Sevilla, M. D. (2015). 5-Thiocyanato-2'-deoxyuridine as a possible radiosensitizer: electron-induced formation of uracil-C5-thiyl radical and its dimerization. *Physical Chemistry Chemical Physics*, *17*, 16907.
- Zewail, A. H. (2000). Femtochemistry: Atomic-scale dynamics of the chemical bond. *The Journal of Physical Chemistry. A*, *104*, 5660.
- Zhan, C.-G., & Dixon, D. A. (2003). The nature and absolute hydration free energy of the solvated electron in water. *The Journal of Physical Chemistry B*, *107*, 4403.
- Zhang, R. B., Zhang, K., & Eriksson, L. A. (2008). Theoretical studies of damage to 3'-uridine monophosphate induced by electron attachment. *Chemistry A European Journal*, *14*, 2850.
- Zhao, Y., & Truhlar, D. G. (2008). Density functionals with broad applicability in chemistry. *Accounts of Chemical Research*, *41*, 157.
- Zheng, Y., Cloutier, P., Hunting, D. J., Wagner, J. R., & Sanche, L. (2004). Glycosidic bond cleavage of thymidine by low-energy electrons. *Journal of the American Chemical Society*, *126*, 1002.
- Zheng, Y., Cloutier, P., Hunting, D. J., Sanche, L., & Wagner, J. R. (2005). Chemical basis of DNA sugar-phosphate cleavage by low-energy electrons. *Journal of the American Chemical Society*, *127*, 16592.
- Zheng, Y., Wagner, J. R., & Sanche, L. (2006). DNA damage induced by low-energy electrons: Electron transfer and diffraction. *Physical Review Letters*, *96*, 208101.
- Zheng, Y., Hunting, D. J., Ayotte, P., & Sanche, L. (2008). Role of secondary low-energy electrons in the concomitant chemoradiation therapy of cancer. *Physical Review Letters*, *100*, 198101.

Jiří Šponer, Manoj K. Shukla, Jing Wang, and Jerzy Leszczynski

Contents

Introduction	1804
Hydrogen Bonding and Stacking Interactions in Nucleic Acids	1809
Level of Computations	1811
Gold Standard	1812
Other Approaches	1812
Geometries	1814
Interpreting the Computations	1817
Conclusion	1820
Bibliography	1820

Abstract

A comprehensive analysis of the benefits and pitfalls of quantum chemical methods used to determine the structures, properties, and functions of DNA and RNA fragments is presented. Main emphasis is given to the application of different ab initio quantum chemical methods. An overview of computations reveals that quantum chemical methods provide an important means to investigate structures and interactions in nucleic acids. However, judicious selection of

J. Šponer (✉)

Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno, Czech Republic
e-mail: sponer@ncbr.chemi.muni.cz

M.K. Shukla

US Army Engineer Research and Development Center, Vicksburg, MS, USA
e-mail: Manoj.K.Shukla@usace.army.mil

J. Wang • J. Leszczynski

Department of Chemistry and Biochemistry, Interdisciplinary Center for Nanotoxicity, Jackson State University, Jackson, MS, USA
e-mail: jingw@icnanotox.org; jerzy@icnanotox.org

computational approach is necessary, depending upon the nature of the problem under investigation.

Introduction

Over the past several decades, investigations of the structure and property of nucleic acids have been an important subject of scientific research. Such investigations have been motivated by the fundamental roles played by ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) in biology.

DNA has been considered to be the central biology molecule, being the depository of genetic information, where hereditary information (of higher-level life forms) is encoded in the form of specific sequences of hydrogen bonds formed between the purine (adenine and guanine) and complementary pyrimidine (thymine and cytosine) bases. Obviously, any permanent variation in the hydrogen bonding pattern may change the function and can even be lethal (especially when specific mutations accumulate). On the other hand all kinds of beneficial mutations, ranging from small-scale mutations (such as point mutations, insertions and deletions) to large-scale mutations (such as chromosomal translocations), are the driving force of evolution. This demands that the structural integrity of DNA be maintained for the identity of each and every organism. It is, however, notable that the biology of DNA does not depend only on the “digital” information of the base sequence. In reality, key aspects of DNA storage in the chromatin and all major aspects of DNA-based control of gene expression are regulated by the subtle sequence-dependent variability of conformational and physicochemical properties of DNA double helix, which is definitely everything else than just a regular double helix.

RNA was, until the 1980s, considered a boring and unimportant cousin of DNA. However, since earth-shaking discovery of RNA catalysis in 1982 (Guerrier-Takada et al. 1983; Kruger et al. 1982) (1989 Nobel Prize for chemistry, Cech and Altman), major RNA discoveries keep coming one after another. We now assume that the RNA molecule is likely the primary molecule of life, the first modern replicator, which in later stages of primary evolution created chemically more stable DNA for better coding and proteins for more diverse catalysis. Since the early stages of evolution, RNA has kept control over many key processes in cellular life while also acquiring new functions. Thus, in the last two decades in biology and biochemistry much of the research focus shifted from DNA to RNA (not reflected by adequate efforts in the field of computational chemistry), as exemplified, e.g., by the 2006 Nobel Prize in Physiology and Medicine awarded to Fire and Mello for their 1998 discovery of RNA interference (Fire et al. 1998) and 2009 Chemistry Nobel Prize to Ramakrishnan, Steitz, and Yonath for solving (in 2000) the atomic resolution structures of the most formidable molecular machine, the ribosome (Ban et al. 2000; Wimberly et al. 2000). A ribosome is a large RNA assembly that has been in the course of evolution supplemented by dozens of ribosomal proteins. It is now well established that while less than 2 % of the genomic DNA codes protein sequences, over 80 % of the genome is actually transcribed into RNA during the cellular life,

most of it (obviously besides genes coding, e.g., ribosomal and transfer RNAs) as noncoding RNA molecules of yet unknown function. In other words, what was not a long time ago considered as “junk” DNA with no obvious role has emerged as DNA template coding for critically important regulatory RNAs that probably affect every corner of gene expression and regulation. One example is the discovery of small RNA molecules (microRNA) that regulate gene expression at multiple levels. The structural versatility and complexity of RNA molecules is incomparably larger than the variability of DNA.

Although the research related to nucleic acids has grown manifold, giving rise to various distinct research fields, the age-old question as to the origin of life on earth remains to be answered. We do not know if it was a spontaneous process which, over the period of time, evolved into the current form on earth or it was endowed from outer space. In fact, we still do not know all the species present on our planet. The existence of different simple molecules, which can be precursors of genetic monomers, such as water, carbon mono- and dioxides, formaldehyde, nitrogen, hydrogen cyanide, hydrogen sulfide, and methane, have been shown in cometary comas (Mix 2006). The purine base adenine has been observed in asteroids and comets. The existence of significant amounts of HCN and HNC molecules in the interstellar space is well known (Ishii et al. 2006). Tennekes et al. (2006) have measured the distribution of these isomers (HCN and HNC) in the protostellar dust core. Smith et al. (2001) have discussed the formation of small HCN-oligomers in the interstellar clouds. It has been demonstrated experimentally that under certain conditions adenine can be formed from the pentamerization of HCN in the solid, liquid, and gas phases (Miller and Urey 1959; Ponnampertuma et al. 1963). Based on the appearance of the brown-orange color as the consequence of impacts of comet P/Shoemaker-Levy 9 on the planet in 1994, the presence of HCN polymers has been speculated on Jupiter (Matthews 1997). The coloration of the Saturn has also been speculated to be due to the presence of HCN polymer. To understand prebiotic adenine synthesis, Glaser et al. (2007) recently performed theoretical calculations on the pyrimidine ring formation of monocyclic HCN-pentamers and found that the key steps proceed without any catalysts producing the purine ring under photolytic conditions and no activation barrier was involved. Roy et al. (2007) proposed a step-by-step pathway for formation of adenine from 2,3-diaminomaleonitrile (DAMN) and 4-amino-5-cyanoimidazole (4-aminoimidazole-5-carbonitrile, AICN) under prebiotic conditions. Experiments for abiotic adenine and purine synthesis from formamide were also reported (Hudson et al. 2012), and a related self-catalyzed mechanisms for the reactions were investigated theoretically (Wang et al. 2013a–c). Based on GC-MS, high-resolution FTIR spectroscopic results and theoretical calculations, an interdisciplinary effort by Ferus et al. (2015) simulated the high-energy synthesis of nucleobases from formamide during the extraterrestrial impact body of the Late Heavy Bombardment (LHB) period. This demonstrated that nucleobases may be synthesized in an impact plasma through reactions of the dissociation products of formamide, without any catalyst.

It is a valid assumption to speculate that life on earth probably evolved under acute harsh environments including the presence of different kinds of irradiations.

Therefore, it is expected that several structural transformations/refinements with respect to genetic code preservation must have taken place. Survival of the fittest prevailed over the period of time in evolution, thereby bringing the purine and pyrimidine bases as the genetic molecules. Further, since these molecules absorb ultraviolet (UV) irradiation, some sort of mechanism was needed to avoid the excited state photo reactions. This was probably achieved through the ultrashort excited state lifetimes of nucleic acid bases (de La Harpe and Kohler 2011; Middleton et al. 2009; Serrano-Andres and Merchan 2009; Shukla and Leszczynski 2007, 2008). Recent state-of-the-art investigations have suggested that such ultrafast excitation processes are achieved through internal conversion where excited and ground state potential energy surfaces conically intersect (Barbatti et al. 2010; Bisgaard et al. 2009; de Vries and Hobza 2007; de La Harpe and Kohler 2011; Kohler 2010; Middleton et al. 2009; Serrano-Andres and Merchan 2009; Shukla and Leszczynski 2007, 2008; Yamazaki et al. 2008). And thus the absorbed energy is efficiently dissipated in the form of heat. Therefore, it is not unexpected that accurate structural determination of nucleic acids and their fragments have been one of the fundamental areas of research. Another obvious precondition that has to be fulfilled in the initial selection of a nucleobase is its inability to tautomerize in aqueous solution, as tautomers would bias any nucleobase-based genetic code as well as RNA folding; thus none of the native nucleobases tautomerize under biochemical conditions to any appreciable amount.

Computational quantum chemical techniques are fast becoming an attractive alternative to the expensive and time-consuming experimental methods in determining the structures and activities of molecular systems. Although we would like to state that modeling of the exact experimental environment, in particular the large biological systems *in vivo*, is not yet possible, computational methods can still provide reliable predictions and thus can be useful to experimentalists. Theoretical methods are especially attractive in the area where experimental measurements are still not possible, e.g., the determination of excited state geometries of complex molecular systems. For the smaller molecular species, one can routinely apply a high level of electron correlated methods and large basis sets; however, for larger molecular systems one has to make a compromise between the level of theory and the basis set and thus with computational accuracy. Experimental determination of excited state geometries of complex molecules like nucleic acid bases is still not possible; only some limited information such as possibility of excited state nonplanarity can be deduced. On the other hand, quantitative information about the excited state geometries can be obtained using the reliable theoretical level, although one has to make a compromise between the theoretical method and the size of the system under investigation. Further, while there was some indication about the amino group nonplanarity in nucleic acid bases in the crystal environment (McMullan et al. 1980), quantitative prediction about amino group nonplanarity was obtained through the quantum mechanical calculations about a decade ago (Leszczynski 1992). It should be noted that such nonplanarity in the gas phase of molecules using experiment was obtained by Dong and Miller in 2002 (Dong and Miller 2002). However, we believe that theoretical and experimental methods are

complementary to each other, and a judicious decision is needed for their efficient application. One of the classical examples would be the tautomerism in guanine.

It is well known that due to the lowest ionization energy among nucleic acid bases, guanine is the primary target for nucleic acid damage by ionizing irradiation (Crespo-Hernandez et al. 2004; Lin et al. 1980). Further, low-energy electrons can also cause the strand break (Boudaffa et al. 2000; Gu et al. 2006, 2007a, b, 2010a–d, 2011a, b; Kumar and Sevilla 2007; Cauet et al. 2014). A combination of molecular dynamics and density functional theory approaches was applied to simulate the reactions that can damage DNA by the attachment of a low energy electron to the nucleobase (McAllister et al. 2015). A low-energy electron interaction with the phosphate group in DNA molecule demonstrates a single-strand break pathways (Bhaskaran and Sarma 2015). DNA labeled with electrophilic nucleobases was reported to be damaged by ionizing or UV radiation (Rak et al. 2015). A recent experimental study on single-strand DNA oligonucleotide suggests that there is a linear correlation between the low-energy electron-induced DNA damage and the presence of the guanine molecules in the sequence (Solomon et al. 2009). Further, guanine can potentially form the most diverse set of energetically accessible rare tautomers in nonpolar environments. Initially, based upon the infrared (IR) spectroscopic analysis of guanine in the argon matrix the presence of equal proportions of keto and enol forms was suggested (Sheina et al. 1987). But, the canonical form of guanine dominates in the polar solvent (Leszczynski 2000). Theoretical methods have generally predicted that the keto-N7H tautomer of guanine is the most stable in the gas phase, but the keto-N9H tautomer dominated in the water solution. At the MP2 and CCSD(T) levels along with several large basis sets, the four low-energy tautomers of guanine (keto-N9H, keto-N7H and cis and trans forms of enol-N9H) have been shown to be within 1 kcal/mol of energy (Gorb et al. 2005). The assignments of resonance-enhanced multiphoton ionization spectra of laser-desorbed, jet-cooled guanine have suggested the presence of four tautomers of guanine (Mons et al. 2002; Nir et al. 2001). Based upon the comparison of IR spectra of thermally vaporized guanine trapped in helium droplets with that of computed vibrational frequencies at the MP2 level with the 6-311++G(d,p) and aug-cc-pVDZ basis sets, Choi and Miller (2006) have assigned the presence of keto-N9H, keto-N7H, and cis and trans forms of enol-N9H tautomer of guanine. The results of guanine in helium droplets necessitated the reassignment of earlier R2PI data and accordingly, based upon the comparison of experimental and theoretical results, Mons et al. (2006) found the presence of enol-N9H-trans, enol-N7H, and two rotamers of the keto-N7H-imino tautomers of guanine in the supersonic jet-beam. Thus, in the new assigned R2PI spectra the stable keto-N9H and keto-N7H tautomers of guanine are not present. The high-level of theoretical calculations (Chen and Li 2006; Marian 2007) on guanine tautomers also supported the reassignment of the R2PI spectra of guanine tautomers in the supersonic jet-cooled beam and suggested the presence of efficient nonradiative deactivation channels as the reason for the missing of spectral origins of the stable tautomers in the R2PI experiments. Recently, Zhou et al. (2009) have performed a comprehensive investigation of guanine tautomers using the VUV photoionization technique, where

the gas phase of guanine was obtained at both the thermal vaporization and laser desorption methods. It was revealed that the method used to generate the gas phase sample of guanine has significant influence on the population of tautomers in the experiment. Consequently, it was found that in the thermal vaporization, a maximum of five most stable tautomers are populated and these results are in agreement with that obtained in the helium droplet experiment. On the other hand, when the laser desorption technique was used to make a gas phase sample of guanine, up to seven tautomers are populated. As noted above, however, guanine for obvious reasons does not tautomerize in biochemically relevant environments, as convincingly shown more than a decade ago also by advanced computational methods (Colominas et al. 1996). Eventual computations suggesting formation of tautomers of natural bases in water are to be dismissed and are at odds with all other experimental data.

Another unique feature of QM methods is their capability to reveal relation between molecular structures and molecular energies at the level of direct (contact, gas phase, electronic structures) interactions. This concerns mainly the two fundamental interactions, base stacking and base pairing. Stacking interactions play an important role in the biological structure, providing both thermodynamics stability and structuring of nucleic acids. For example, base stacking is assumed to be the primary determinant of sequence dependence of B-DNA structure and flexibility, which is the single most important feature of DNA that enters all of DNAs molecular interactions, genetic material storage, replication, and gene expression. The sequence dependence of DNA (and the role of base stacking) has been intensely studied. Despite this, all the experimental and theoretical research failed to provide clear rules allowing to predict the fine B-DNA structural variability from sequence (Calladine 1982; Dickerson and Drew 1981; Šponer and Kypr 1991; Suzuki et al. 1997; Wing et al. 1980; Yanagi et al. 1991). The research in this area has been stalled for some time and perhaps improved description of stacking with the help of modern QM methods could bring some new ideas.

Base pairing (extended beyond the base-base interactions) is especially interesting in large functional RNAs, where it determines their architectures and gives very strict constraints on RNA evolutionary patterns (Leontis et al. 2002; Šponer et al. 2010; Stombaugh et al. 2009; Zirbel et al. 2009). Large RNA molecules are organized as complex (and often very dynamical) jigsaw puzzles, where the exact shapes of the base pairs determine function and thus also allowable mutations (isostericity principle). Nevertheless, recent studies also indicate a nonnegligible role of energy of molecular interactions supplementing the basic isostericity principle (Zirbel et al. 2009).

Although stacking interactions are important and high-level *ab initio* methods are needed to account for such interactions, it is only possible to use these methods to model systems and very small fragments of large biomolecules, even using the large computational resources. Therefore, especially force field methods (and QM/MM method to certain extent) are mainly used to study larger biopolymers and other large biological systems. Nevertheless, QM calculations remain instrumental in reference studies on the nature and magnitude of molecular interactions in nucleic

acids and in verification of the other methods. Quantum chemical calculations provided the ultimate answer about the physicochemical nature of base stacking and characterized many other features of nucleobase interactions (Hobza and Sponer 1999; Morgado et al. 2009; Svozil et al. 2010).

Hydrogen Bonding and Stacking Interactions in Nucleic Acids

The most fundamental roles of nucleic acid bases (nucleobases) in biology and chemistry are their involvement in two qualitatively different mutual interactions: hydrogen bonding (base pairing) and aromatic base stacking.

The base pairing is utilizing the H-bond donor and acceptor capabilities of nucleobase exocyclic groups and ring nitrogen atoms. In RNA molecules, the base pairing also involves the sugar ribose which, in contrast to DNA deoxyribose, possesses hydroxyl group in the 2' position (Leontis et al. 2002; Sponer et al. 2005a, 2007, 2009, 2010; Stombaugh et al. 2009). The 2' -hydroxyl group is a powerful donor and acceptor of hydrogen bonds. In fact, key RNA base pair families utilize the 2' -OH group for base pairing, and these extended base pairs are known as sugar-edge (SE) base pairs or interactions. Many important "SE" base pairs include no direct base to base H-bonds, and yet they are crucially important for folding of complex functional noncoding RNA molecules and ribonucleoprotein particles. By noncoding RNAs we mean RNA molecules that are not translated to proteins via messenger RNA and perform different functions instead. Note that recent research highlighted that while less than 2 % of the human genome directly codes for proteins, at least 80 % of the genomic DNA is actually transcribed into RNA. Thus, majority of the genome encodes noncoding RNAs that play absolutely essential roles in life and evolution (many of the RNA functions have yet to be discovered, but they are assumed to be key players in fine regulation of gene expression), which is a finding that has truly revolutionized biology in recent years. The largest noncoding RNAs are, obviously, ribosomal RNAs. The most important RNA tertiary interactions (A-minor and P-interactions) are base pairs, triads, and quartets mediated by base-sugar and sugar-sugar interactions (Sponer et al. 2007). Recently, the RNA base pairing classification was extended to include base-phosphate (BPh) interactions, after recognizing that $\sim 12\%$ of nucleotides in ribosome are involved with BPh interactions with other proximal or distal nucleotides while these interactions bring important evolutionary constraints (Zirbel et al. 2009).

Base stacking occurs between the aromatic faces of the nucleic acid bases and is at least equally as important as base pairing, for both thermodynamics stabilization and shaping of nucleic acids. Stacking is responsible for the local conformational variations and other sequence-dependent properties of B-DNA. With the help of eight micro-solvation water molecules, the dinucleoside phosphate deoxyguanylyl-3',5'-deoxycytidine dimer octahydrate, [dGpdC]₂ illustrative minimal unit of the DNA double helix has been constructed and its geometry fully optimized (B3LYP/6-31 + G(d,p) level) (Gu et al. 2011a). Similarly, four water molecules microhydrated

dApdT, dTpdA, dGpdC, and dCpdG species have been predicted to have reasonable stacking structure using both B3LYP and M05-2x functionals (Gu et al. 2009, 2010c). The Minnesota density functionals, M05-2X and M06-2X, level calculations were carried out to evaluate the stacking and H-bonding patterns for the structures of the nucleotide oligomers (dGpdC, dGpdG, and dGpdCpdG) for single-strand DNA. This level of calculations shows excellent agreement with the results from the MP2 theory (Gu et al. 2011a, 2012a, b). Unambiguous classification of base stacking is missing. One of the reasons is flexibility of base stacking, as the stacked bases can always slide and twist over a range of mutual stacked geometries, being not fixed by the individual H-bonds (note, however, that many complex RNA base pairs also possesses complex conformational space with multiple competing conformations). While classification of base pairing could have been done purely by considering structural data (i.e., geometries seen in X-ray structures), classification of base stacking will likely require appropriate energy analyses.

The *ab initio* QM technique can be used to determine optimal structures of molecular clusters and to calculate energies for any single geometry of the cluster. QM calculations provide molecular wave functions, which can be used to derive physicochemical properties, such as vibrational spectra, dipole and higher multipole moments, polarizabilities, proton affinities, NMR parameters, and others. Nevertheless, the main achievement of QM calculations was the description of the nature and energetics of nucleobase interactions. This is because the leading experimental approaches of structural biology, that is, mainly X-ray crystallography, provide purely structural data. Information about energetics of molecular interactions can be inferred only indirectly, while the interpretation of structural data ignoring energetics of molecular interactions is often misleading.

QM calculations can help to understand the role of molecular interactions in nucleic acids because of their capability to give a direct link between structures and energies. Nevertheless, QM calculations are always done on small systems and typically in the gas phase that is far from real environments and structural contexts. To make meaningful QM computations with biological relevance, we need to follow several rules (Svozil et al. 2010).

First, we need to make the strategic decision whether our computations aim to be indirectly or directly relevant to biology. By indirect relevance we mean, for example, the use of QM calculations for parameterization or validation of other methods (mainly the molecular mechanical force fields) or for basic understanding of the physical chemistry of interactions. By direct relevance we mean applications that range from calculations of some specific interaction patterns seen in structural studies (Sponer et al. 2003) through combined QM-bioinformatics studies aimed at classifying interactions (Sponer et al. 2003; Zirbel et al. 2009) up to QM/MM calculations of RNA catalysis (Banáš et al. 2009). Then, there are at least three tactical issues that need to be very carefully decided. We need to select the level of calculations, i.e., primarily the method and basis set. Equally important is the appropriate choice or preparation of geometries used in computations. Inappropriate geometries may easily blow up the whole effort and result in computations that are misleading (Svozil et al. 2010). Finally, the results should be properly interpreted. We also need to separately consider applications that deal with nucleic acid bases

but are directed to other areas of science (basic physicochemical experiments, origin of life studies, adsorption on surfaces, supramolecular assemblies, etc.)

Level of Computations

Let us assume that we have a dimer of two nucleic acid bases, A and B , with a given geometry. For that geometry, the interaction energy between A and B , $\Delta E^{A \cdots B}$, is the energy difference between the total electronic energy of the dimer $E^{A \cdots B}$ and the electronic energies E^A and E^B of isolated bases.

$$\Delta E^{AB} = E^{AB} - E^A - E^B \quad (1)$$

The interaction energy reflects a hypothetical dimerization process at 0 K and is not measurable. In order to be related to experimental dissociation energies D_0 and enthalpies of formation, the deformation energy of monomers and the zero-point vibration energy must be included. The zero-point energies and enthalpy and entropy contributions at nonzero temperature are usually calculated in the harmonic approximation. Since base pair complexes are weak, anharmonicity can play an important role, especially for stacked systems and particularly at higher temperatures. Nevertheless, except for direct comparison with gas phase experiment, interaction energy evaluation is the sufficient outcome of QM analysis.

Methods suitable for base stacking and base pairing calculations have been discussed many times (Morgado et al. 2009; Spomer et al. 2008; Gu et al. 2008; Svozil et al. 2010) and will thus be only briefly noted here. With modern computers, we have methods that are satisfactorily accurate.

Base stacking stabilization is dominated by the intermolecular electron correlation effects (i.e., the dispersion energy). Therefore, stacking calculations must be done with inclusion of electron correlation effects and with large basis sets of atomic orbitals. The dispersion energy is created in the space between the interacting monomers that are separated by $\sim 3.3\text{--}3.4$ Å. This space needs to be covered by atomic orbitals, dictating the use of diffuse-polarized basis sets. H-bonded complexes are not dominated by the dispersion energy, albeit it is still a very significant contribution. Thus, HF calculations or computations with “dispersion-neglecting” DFT methods, while not being accurate, are not entirely incorrect. When including electron correlation, higher angular momentum functions are important for base pairing, since the space between the interacting monomers is bridged by H-atoms so the requirement for the diffusivity of atomic orbitals is not as strict as for stacking.

The best accuracy is achieved by complete basis set (CBS) extrapolation methods, when two systematically improved basis sets are applied and the data is then extrapolated. The interaction energy computations, even with large basis sets, need to be corrected for basis set superposition error (BSSE). We oppose suggestions to ignore the BSSE correction or to attempt only its partial inclusion while assuming that the numbers can be correct due to error cancellation. This is a risky game. It is much better to provide BSSE-corrected numbers where a solid

estimate of the underestimation of the interaction is typically possible. Fortunately, the CBS calculations are intrinsically BSSE-free. Similarly, computations with modern parameterized DFT-D methods (see below) do not require BSSE correction, since it is indirectly (effectively) included via parameterization.

H-bonded complexes are rather well described by the MP2 method while for aromatic stacking this method typically significantly overshoots the stabilization. Thus, for stacking higher-order electron correlations are quite important.

Gold Standard

Within the variation (supramolecular) approach, definitely the method of choice for interaction energies would be the coupled cluster CCSD(T) method (in which the single and double excitations are evaluated iteratively while the triple excitations are included in a noniterative way). The CCSD(T) method yields a significant portion of the correlation energy. The MP2 method, including the double electron excitations at the second order of perturbation theory, overestimates the correlation interaction energy for stacking, as noted above.

The determination of a CBS limit of CCSD(T) calculations is still difficult. Until recently, the CCSD(T) calculations for larger complexes were performed only with medium basis sets (e.g., 6-31G*) and even these calculations were at the computer limits. Thus, the gold standard in base stacking and pairing calculations is the method sometimes abbreviated as CBS(T), which utilizes the similar basis set dependence of the CCSD(T) and MP2 energies. Thus, the difference between CCSD(T) and MP2 interaction energies ($\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}}$) exhibits small basis set dependence, and the CCSD(T)/CBS interaction energy can be approximated as

$$\Delta E_{\text{CBS}}^{\text{CCSD(T)}} = \Delta E_{\text{CBS}}^{\text{MP2}} + \left(\Delta E^{\text{CCSD(T)}} - \Delta E^{\text{MP2}} \right) \Big|_{\text{medium basis set}},$$

which is abbreviated as CBS(T) to distinguish from full CCSD(T)/CBS computation (Sponer et al. 2006, 2008). Various extrapolation schemes have been suggested for the determination of the $\Delta E_{\text{CBS}}^{\text{MP2}}$ term; the one proposed by Halkier et al. (1999) is the most widely used.

For aromatic stacking interactions the Δ CCSD(T) correction term is systematically nonnegligible (repulsive) and should never be omitted. For H-bonding interactions, the Δ CCSD(T) correction term is typically very small (Sponer et al. 2004, 2006, 2008).

Other Approaches

Earlier calculations on base stacking were done with the MP2 method utilizing the modified 6-31G*(0.25) basis set (Hobza and Sponer 1999; Sponer et al. 1996b, 1997). The polarization d-functions of the standard 6-31G* basis with an exponent

of 0.8 were replaced by more diffuse ones with an exponent of 0.25, allowing inclusion of a major part of the dispersion energy. While the MP2/6-31G * (0.25) method is now outdated, the main conclusions reached by the MP2/6-31G * (0.25) studies remain valid.

DFT methods were for years not recommended for stacking calculations, because common DFT methods (based on the local density, its gradient, and the local kinetic-energy density) notoriously fail to capture the (nonlocal) dispersion energy (Hobza et al. 1995; Kristyán and Pulay 1994).

This is a common feature of all LDA and GGA functionals, not excluding even the most advanced meta-GGA functionals. Many recent DFT methods provide much better results for dispersion-controlled complexes (Zhao and Truhlar 2008).

Nevertheless, we still suggest use of caution (and testing) in their application to stacking complexes. An alternative (which can achieve, at least for now, better accuracy and major speed up) is based on augmenting the DFT energy by an empirical London dispersion energy term (Elstner et al. 2001; Grimme 2004; Jurecka et al. 2007).

To correct for the overlap effects, the dispersion energy is damped by distance-dependent damping function. The dispersion energy, represented by the C_6/R^6 formula, is calculated separately from the DFT calculation and is simply added to the DFT energy. The disadvantage of DFT-D methods is obviously the need to combine electronic structure calculations with classical “force field” correction term, which also affects the transferability of these methods. Thus these methods are expected to be surpassed in the future by “true” DFT-based dispersion-including methods; however, for the moment it seems to us that DFT-D is more satisfactory for routine calculations of nucleobase interactions. One present difficulty is that we have so many new methods in the literature that is difficult to choose. This issue is beyond the scope of this chapter and we refer the reader to specialized literature (Banaš et al. 2009; Spöner et al. 2008). We hope that standard (optimal) wide-spectrum dispersion-including DFT methods will soon be identified.

Interaction energies can also be obtained by perturbation methods, as a sum of perturbation contributions. The symmetry adapted perturbation theory (SAPT) provides the interaction energy as a sum of first-, second-, and higher-order perturbation terms (Heßelmann et al. 2005; Jeziorski et al. 1994). The first-order contribution contains the electrostatic and exchange energies, while the second-order term includes induction and dispersion energies. The charge transfer energy is included in the second-order induction energy and higher-order contributions. SAPT (with extended basis sets) yields accurate values of the energy components and also of the total interaction energies. The determination of the interaction energy is straightforward and is not biased by additional theoretical problems, such as the BSSE inherent to variational methods. The broad use of SAPT is, however, hampered by large computer requirements. A significant improvement was reached by the combination of SAPT and DFT theories (Jeziorski et al. 1994). The DFT-SAPT approach has been rather routinely used for base-base calculations. When making SAPT decompositions, it is extremely important to use well-defined geometries, as the SAPT components are exceptionally sensitive to inter-monomer

separation, much more than the total interaction energies. SAPT decompositions can be spoiled by inappropriate choice of geometries (Sponer et al. 2008). Note also that from the biological point of view, what primarily matters are the interaction energies. Thus the usefulness of decompositions should not be overinterpreted (Sponer et al. 2008).

Geometries

Quantum-chemical calculations can provide meaningful data only when the energies are derived at appropriate geometries.

The easiest systems to deal with are well-behaved base pairs where gradient optimization leads to relevant structures. Modern QM programs allow easy gradient optimizations of base pairs, where all coordinates (or parameters) are optimized. Standard optimizations are not corrected for BSSE. Some earlier studies where base pairs were optimized in a step-by-step manner are of historical interest only. Since the optimization itself is more computer-demanding than the subsequent interaction energy evaluation, very often a better level of theory (level X) is used for interaction energy calculation than for optimization (level Y). This is abbreviated as X/Y. For example, the abbreviation MP2/aug-cc-pVTZ/MP2/cc-pVDZ indicates that the optimized structure was obtained at the MP2/cc-pVDZ level, while the energies were derived for this optimized geometry at the MP2/aug-cc-pVTZ level.

The gradient optimization is good for systems with well-defined local minima, while the minima correspond to the biochemically relevant structures. Stacking patterns seen in nucleic acids do not correspond to minima on the potential energy surfaces of isolated stacked dimers and thus conformational scanning is preferred. Further, gradient optimizations of dispersion-controlled clusters are affected by enormous BSSE, unless a very large basis sets are used. With lower quality methods the structures are unstable and convert to H-bonded ones. In addition, gradient optimizations of stacked dimers lead to puckering of the aromatic ring (Hobza and Sponer 1998). This is usually not desirable since in real environments the nucleobases have some interactions at both their sides, probably preventing such puckering. Thus, stacking calculations are mostly carried out as a series of single points with fixed geometries and rigid monomers (Sponer et al. 1996b, 1997).

An attractive option is to take structures from experimental (X-ray) studies (Sponer et al. 1997). However, usual accuracy of these experiments does not guarantee their straightforward utilization in QM energy computations. First, it is not advisable to directly use monomer geometries from PDB files of NA X-ray structures. Due to limited resolution the monomer geometries carry limited experimental information about monomer geometries, while the bases are often deformed after the refinement. Such deformed monomer geometries frustrate the electronic structure. It is necessary to replace (for example via overlay) the monomers from the PDB files via QM-optimized monomers.

In addition, intermolecular X-ray geometries may cause substantial errors in calculations. Especially drastic distortions of the calculated energies can be introduced

by steric clashes in the refined crystal structures (Sponer and Kypr 1993). A real nightmare occurs when the X-ray base stacks are effectively vertically compressed or extended due to inaccurate determination of the interbase angles (propeller twist, base pair roll, etc). This requires a case-by-case judgment and some experience with crystallography. Note that a rather small error in the interbase distances (which may still be tolerable from the geometry point of view) can lead to a considerable energy artifact. This is when the geometry falls into a region of interatomic distances where the short-range repulsion starts to dominate. The calculated energy is a highly nonlinear function of the interatomic distance (Sponer et al. 2008).

Similarly, H-bonded base pairs are sensitive to experimental geometry errors due to the genuine close contact between H-bond partners. Besides data and refinement errors a bad geometry can result from the presence of two or more local substates. Substates cannot be distinguished except as having nominal resolution better than $\sim 1 \text{ \AA}$. The refined geometry reflects an averaged geometry which may have very poor energy. Fiber diffraction models cannot be recommended for direct calculations (Svozil et al. 2010).

We would like to caution against using averaged (3D-bioinformatics) geometries, as they can represent unrealistic single structures from the energy point of view. It is always advisable to generate a range of structures around such geometries and to analyze the properties of the potential energy surfaces. Actually, an open question remains whether the base stacking can be characterized by some single representative geometry. Most likely stacking states correspond to a range of populated geometries, as evidenced, for example, by significant coordinate fluctuations seen for stacked bases in explicit solvent molecular dynamics simulations (Svozil et al. 2010).

We are interested in analyses of specific interactions which are neither stacking nor H-bonding and are substantially affected by the overall topology of the studied systems. The best approach is to fix the intermolecular geometry of interest (typically using a set of three dihedrals and two valence angles plus one intermonomer distance per each dimer) and then relax the monomers intramolecularly (Sponer and Hobza 1994; Sponer et al. 2003; Vlieghe et al. 1999). This approach has been applied in studies of cross-strand close amino group contacts in B-DNA, DNA-drug interactions, bifurcated H-bonds, out-of-plane H-bonds, and some other interactions. If a steric clash is suspected, then the intermonomer distance can again be varied (Sponer and Hobza 1994; Sponer et al. 2003; Vlieghe et al. 1999).

A specific problem is represented by the complex RNA base pairing patterns involving the sugar edges (Sponer et al. 2005a, b, 2007). Many of these base pairs have multiple minima. For many of them, the functional (observed) structures do not correspond to any intrinsic gas phase minimum energy structures, since they are constrained by other interactions and the overall RNA topology. Some base pairs can be intrinsically water-mediated. Huge problems in computations can be created by the sugar hydroxyl group in position 3', which normally is involved in the covalent backbone chain. One option is 3'-methylation. In some cases, the phosphate groups participate in the interactions and need to be included in computations. This creates problems due to a strong ionic nature of the associated interactions which are, in

real systems, obviously attenuated by solvent screening. Close to insurmountable electrostatic problems arise when more than one phosphate directly participates in the interaction. Thus, QM studies of RNA base pairs often require applications of sophisticated geometrical constraints which need to be implemented case by case. For nonneutral systems, even optimizations upon inclusion of continuum solvent could be a viable option. The situation is further complicated by very limited resolution of the experimental structures of folded RNAs (not mentioning ribosome) and their dynamical nature. This can lead to large coordinate errors (including poorly refined *syn* vs. *anti* orientation of the bases or incorrect sugar puckers). Thus, studies of RNA base pairs are far from routine. Studies of geometries that are substantially rearranged compared to the experimental structures are of a little value, similar to studies neglecting the sugar rings for base pairs where the 2'-OH groups are directly involved in base pairing.

When gradient optimization is carried out, the monomer geometries are changed upon complexation. This is due to mutual adaptations of the monomers that improve the intermolecular interaction at the expense of the intramolecular energy terms. Some of the deformations can be directly related to the binding strength. However, for larger systems, some monomer rearrangements reflect rather long-range effects. For example, there could be a substantial reorientation of the flexible sugar-phosphate backbone upon complex formation. Thus, real deformations consist of two contributions. Direct deformations (always present) reflect the strength of the binding and may be complemented by various indirect larger-scale conformational rearrangements. Besides real deformations, the BSSE contributes to the deformation when standard gradient optimization is applied. The BSSE contribution is obviously a computational artifact.

As explained above, the interaction energies of the optimized complexes should be a posteriori corrected for the BSSE using the geometry of the complex and dimer-centered basis set. Then we separately calculate the deformation energy using the monomer basis sets, as difference of monomer energies in the deformed (complex) and optimized (isolated) monomer geometries.

$$E_{\text{Def}}^A = E^{A(\text{dimer geometry})} - E^{A(\text{monomer geometry})}$$

Thus, the interaction energy of a dimer is defined in the following way:

$$\Delta E^{A\dots B} = E^{A\dots B} - (E^A + E^B) + E_{\text{Def}}^A + E_{\text{Def}}^B.$$

The first three energies are calculated in a dimer-centered basis set. The intramolecular deformation energy actually cancels a large part of the intermolecular energy improvement caused by mutual monomer adaptations.

In some of the literature, the authors include deformation energy formally as part of the BSSE correction. We consider this a weird option which may substantially spoil the interpretation of the results. Although it might look more sophisticated mathematically, this approach is misleading, mixing apples and oranges, and is especially unsuitable for larger systems such as base pairs and other fragments

of biopolymers (Sponer et al. 2004). In fact, the integrated expression is, after formal rearrangements, *entirely identical* to the above definition, which in addition is older (Sponer et al. 1996a), i.e., the correction was commonly known before researchers started to include deformations into BSSE correction. Second, while BSSE is a mathematical artifact, monomer deformations are not. They correspond to fundamental properties of the studied clusters including their vibrational spectra and polarization/charge transfer effects. Thus, it is useful to evaluate the magnitude of the monomer deformations explicitly. For flexible systems with large indirect rearrangements (as explained above) any formal inclusion of the deformation term into the BSSE correction is meaningless. Thus, albeit widespread, for base pairs and larger systems of chemical and biological interest, this approach is not appropriate (Szalewicz and Jeziorski 1998).

An alternative approach is to use counterpoise-corrected gradient optimization where the BSSE is removed in each gradient iteration, although with a substantial increase of the computer requirements. It eliminates the BSSE part of the deformation energies while true deformations persist.

In base pairing studies the deformation energy can be calculated either with respect to the planar monomers, thus neglecting the amino group nonplanarity, or with respect to nonplanar bases. These two numbers differ simply by the difference between energies of planar and nonplanar monomers and can be easily compared when needed (Sponer et al. 2004).

Interpreting the Computations

QM calculations (on nucleobase dimers) reveal the binding energy between two bases in the gas phase, i.e., in *complete isolation*. They thus describe the *intrinsic* interactions of the systems with no perturbation by external effects such as solvent. The intrinsic intermolecular stabilities are directly linked to molecular structures and can be derived in any selected geometry. However, the gas phase interaction energies do *not* correspond to the stability of the interactions in nucleic acids, as measured by thermodynamics experiments. It is not possible to easily correlate the QM calculations with measured base pairing and stacking stabilities in nucleic acids. The apparent (measured) strength of the base-base interactions in nucleic acids in various experiments is determined by a complex interplay of many factors and the intrinsic base-base term is only one of them. Many researchers incorrectly believe that the experiments reflect the “true” stabilities of base-base interactions and vice versa.

A textbook example of complexity of molecular interactions is stacking of consecutive protonated cytosines. This is a highly repulsive interaction in the gas phase due to a charge-charge repulsion (Sponer et al. 1996c). Nevertheless, in intercalated i-DNA quadruplex, stacking of a set of consecutive closely spaced protonated cytosines occurs (Gehring et al. 1993). The i-DNA tetraplex is paired via cytosine – protonated (N3) cytosine base pairs, each possessing charge +1. Both cytosines are equivalent in X-ray and NMR experiments suggesting rather fast intra-base pair proton switches (Chen et al. 1994). Two duplexes intercalate to form

the tetraplex. Stability of i-motif is due to this massive accumulation of closely spaced protonated base pairs. In i-DNA, the vertical repulsion between consecutive protonated base pairs is counterbalanced by solvent screening effects and possibly specific interactions with the anionic backbone (Spackova et al. 1998). Thus i-DNA indeed has, in contrast to other DNAs, *repulsive intrinsic stacking energy terms*. This example clearly demonstrates the actual magnitude of mutual compensation of molecular interactions in nucleic acids. The i-DNA stability contradicts the gas phase stacking energy calculations and demonstrates why we cannot use these calculations to directly predict DNA stability. However, the relation can be also reversed. It is not possible to unambiguously evaluate the *intrinsic* stacking energetics based on thermodynamics studies of nucleic acids. There is no unambiguous way to decompose the measured free energies into separate terms that would correspond to stacking, base pairing, etc. In other words, we cannot make straightforward extrapolation from gas phase to nucleic acids, while, conversely, studies of nucleic acids bring no unambiguous information about the intrinsic base-base terms.

To show the full complexity of molecular interactions, let us underline that the screening is specific for i-DNA. Strikingly contrasting i-DNA is the behavior of consecutive protonated cytosines in C + -G.C triples of Pyr-Pur.Pyr triplexes. Consecutive protonated cytosines would be needed to recognize consecutive guanines in the second strand (Soliva et al. 1998). The vertical position of protonated cytosines in triplex would adopt arrangement closely resembling i-DNA, and also planar H-bonding of the third-strand protonated cytosines to N7 of second-strand guanines resembles the i-DNA base pairing. However, this sharply destabilizes the DNA triplex, and even two or three consecutive CH^+ are not tolerated. This indicates that in triplex the screening of the vertical electrostatic repulsion by the backbone and solvent is less efficient than in i-DNA. Thus in this particular case of i-DNA and triplexes we cannot transfer experience concerning nucleobase interactions between two DNA forms. Each case should be studied separately. In other words, a given type of base pairing and base stacking may have entirely opposite roles in different nucleic acid forms. A given interaction may be a crucial stabilizing factor for one type of nucleic acid architecture (protonation of consecutive cytosines in the i-motif), while it may be even not tolerated in another architecture. This illustrates that there is no way to design some ultimate experiments to decide about the *common nature* of base stacking in nucleic acids. This simply is a wrong question. In order to understand the interactions in nucleic acids, we need to consider a wide range of systems, and the gas phase data represent an important part of the overall picture.

It is nevertheless clear that proper inclusion of solvent screening could help in interpretation of the QM data. Unfortunately, accurate inclusion of solvent effects into QM calculations is difficult. One option is to extend the studied system by a finite set of explicit water molecules. Such calculations still deal with a gas phase molecular cluster and do not correspond to bulk hydration. The cluster hydration patterns differ from those in water where the first shell waters interact with the second shell, etc., and the whole system is dynamical, as evidenced by large-scale explicit solvent simulations of nucleic acids. In a small cluster, individual

water molecules will form bridges and zippers between bases in order to maximize the number of H-bonds. The hydration picture in solution (MD data) and X-ray structures is different and reveals simple noncooperative in-plane hydration of the polar nucleobase sites (the nitrogens and oxygens). Water binding sites in common hydration sites around nucleic acids have water binding times ~ 50 – 500 ps. In complex RNA molecules or in molecular complexes, some hydration sites may be occupied by tightly bound water molecules (Rěblová et al. 2003). A substantial problem of cluster calculations is that the potential energy surface contains a large number of minima and, without an efficient sampling technique, it is virtually impossible to verify the true global minimum (Kabelác and Hobza 2001).

The other option is to include the solvent as a polarizable continuum. QM methods consider effects of the continuum on the electronic structure of the solute molecules, in contrast to classical continuum approaches. The outcomes are quite sensitive to the choice of parameters such as the atomic radii used to define the “solute” cavity; no universal accurate set of “true” radii can be established. The continuum calculations may be combined with cluster calculations, where the first hydration shell is treated explicitly. Even if QM continuum solvent calculation is properly performed, such calculations are not sufficient to achieve a direct correspondence with thermodynamics experiments. In practice we are neglecting, for example, the loss of degrees of freedom upon duplex formation, all effects associated with the presence of sugar-phosphate backbone including the only partial exposure of the bases to the solvent, etc. In addition, different base sequences are likely associated with different solute flexibilities (some base pair step sequences are stiff, others flexible) which will contribute to the free energy balance via solute entropy contributions. Reliable evaluation of these contributions is outside the applicability of available computational methods. Thus, the continuum solvent calculations should be considered just as rough estimates of the attenuation of the electrostatic contributions to the free energy upon solvation, which are still far from fully competent free energy computations.

All of the abovementioned considerations explain why the QM calculations on base stacking and base pairing do not (and should not) correlate with the measured thermodynamics properties of nucleic acids. Therefore, QM calculations and intrinsic interaction energies should never be interpreted as straightforward determinants of nucleic acid stabilities. This naive overinterpretation of otherwise very valuable computed data can discredit the computations. There is a plenty of evidence suggesting that the role of molecular interactions in thermodynamics stabilization of even the simplest duplex nucleic acids is more complex than usually assumed, and results from tiny, irregular, and case-specific interplays of all molecular forces, where literally a single specific hydration site or pocket can change the balance. Even simple base modifications and substitutions, as small as deletions of a single exocyclic group, may have complex and, at first sight, mutually contradicting context-dependent impacts on the measured thermodynamics stabilities that cannot be a priori predicted (Chen and Turner 2006; Siegfried et al. 2007).

We nevertheless suggest that this enormous complexity of molecular interactions defines new roles for modern computations, combing well-calibrated simulation

approaches and accurate QM calculations. The calculations can provide key insights into the tricky games of molecular interactions that are shaping up the molecules with their associated free energies and that are not fully understandable based on purely experimental approaches (Kopitz et al. 2008; Yildirim and Turner 2005; Yildirim et al. 2009).

Conclusion

QM calculations represent the leading tool to study intrinsic molecular interactions in nucleic acids, such as base stacking and base pairing. However, the QM data should not be overinterpreted, and any extrapolation to nucleic acids requires proper consideration of the gas phase nature of QM calculations. In addition, in order to obtain meaningful QM data, basic methodological requirements must be fulfilled. These include, in addition to the obvious selection of appropriate level of calculations, very careful selection or determination of geometries, which is discussed in detail in this chapter.

Acknowledgments JS acknowledges the support from the Grant Agency of the Academy of Sciences of the Czech Republic (grant IAA400040802), Grant Agency of the Czech Republic (grant 203/09/1476), Ministry of Education of the Czech Republic (LC06030), and Academy of Sciences of the Czech Republic (grants AV0Z50040507 and AV0Z50040702). MKS and JL are thankful for supports from NSF-CREST (grant HRD-0833178) and NSF EPSCoR (grant 362492-190200-01\NSFEPS-0903787) and computational support from Mississippi Center for Supercomputing Research (MCSR).

Bibliography

- Ban, N., Nissen, P., Hansen, J., Moore, P. B., & Steitz, T. A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 Angstrom resolution. *Science*, 289, 905.
- Banáš, P., Jurecka, P., Walter, N. G., Šponer, J., & Otyepka, M. (2009). Theoretical studies of RNA catalysis: Hybrid QM/MM methods and their comparison with MD and QM. *Methods*, 49, 202.
- Bao, X., Wang, J., Gu, J., & Leszczynski, J. (2006). DNA strand breaks induced by near-zero-electronvolt electron attachment to pyrimidine nucleotides. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 5658.
- Barbatti, M., Aquino, A. J. A., Szymczak, J. J., Nachtigallova, D., Hobza, P., & Lischka, H. (2010). Relaxation mechanisms of UV-photoexcited DNA and RNA nucleobases. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 21453.
- Bhaskaran, R., & Sarma, M. (2015). Low-energy electron interaction with the phosphate group in DNA molecule and the characteristics of single-strand break pathways. *The Journal of Physical Chemistry. A*, 119, 10130.
- Bisgaard, C. Z., Satzger, H., Ullrich, S., & Stolow, A. (2009). Excited-state dynamics of isolated DNA bases: A case study of adenine. *ChemPhysChem*, 10, 101.
- Boudaffa, B., Cloutier, P., Haunting, D., Huels, M. A., & Sanche, L. (2000). Resonant formation of DNA strand breaks by low-energy (3–20 eV) electrons. *Science*, 287, 1658.
- Calladine, C. R. (1982). Mechanics of sequence-dependent stacking of bases in B-DNA. *Journal of Molecular Biology*, 25, 343.

- Cauet, E., Bogatko, S., Lievin, J., De Proft, F., & Geerlings, P. (2014). Electron-attachment-induced DNA damage: Instantaneous strand breaks. *The Journal of Physical Chemistry. B*, *117*, 9669.
- Chen, H., & Li, S. (2006). Theoretical study on the excitation energies of six tautomers of guanine: Evidence for the assignment of the rare tautomers. *The Journal of Physical Chemistry. A*, *110*, 12360.
- Chen, G., & Turner, D. H. (2006). Consecutive GA pairs stabilize medium-size RNA internal loops. *Biochemistry*, *45*, 4025.
- Chen, L., Cai, L., Zhang, X., & Rich, A. (1994). Crystal structure of a four-stranded intercalated DNA: d(C4). *Biochemistry*, *33*, 13540.
- Choi, M. Y., & Miller, R. E. (2006). Four tautomers of isolated guanine from infrared laser spectroscopy in helium nanodroplets. *Journal of the American Chemical Society*, *128*, 7320.
- Colominas, C., Luque, F. J., & Orozco, M. (1996). Tautomerism and protonation of guanine and cytosine. Implications in the formation of hydrogen-bonded complexes. *Journal of the American Chemical Society*, *118*, 6811.
- Crespo-Hernandez, C. E., Arce, R., Ishikawa, Y., Gorb, L., Leszczynski, J., & Close, D. M. (2004). Ab initio ionization energy thresholds of DNA and RNA bases in gas phase and in aqueous solution. *The Journal of Physical Chemistry. A*, *108*, 6373.
- de La Harpe, K., & Kohler, B. (2011). Observation of long-lived excited states in DNA oligonucleotides with significant base sequence disorder. *Journal of Physical Chemistry Letters*, *2*, 133.
- de Vries, M. S., & Hobza, P. (2007). Gas-phase spectroscopy of molecular building blocks. *Annual Review of Physical Chemistry*, *58*, 585.
- Dickerson, R. E., & Drew, H. R. (1981). Structure of a B-DNA dodecamer. II. Influence of base sequence on helix structure. *Journal of Molecular Biology*, *15*, 761.
- Dong, F., & Miller, R. E. (2002). Vibrational transition moment angles in isolated biomolecules: A structural tool. *Science*, *298*, 1227.
- Elstner, M., Hobza, P., Frauenheim, T., Suhai, S., & Eftimios, K. (2001). Hydrogen bonding and stacking interactions of nucleic acid base pairs: A density-functional-theory based treatment. *Journal of Chemical Physics*, *114*, 5149.
- Ferus, M., Nesvomy, D., Sponer, J., Kubelik, P., Michalcíková, R., Shestivska, V., Sponer, J. E., & Civis, S. (2015). High-energy chemistry of formamide: A unified mechanism of nucleobase formation. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 657.
- Fire, A., Xu, S. Q., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, *391*, 806.
- Gehring, K., Leroy, J.-L., & Gueron, M. (1993). A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature*, *363*, 561.
- Glaser, R., Hodgen, B., Farrelly, D., & Mckee, E. (2007). Adenine synthesis in interstellar space: Mechanisms of prebiotic pyrimidine-ring formation of monocyclic HCN-pentamers. *Astrobiology*, *7*, 455.
- Gorb, L., Kaczmarek, A., Gorb, A., Sadlej, A. J., & Leszczynski, J. (2005). Thermodynamics and kinetics of intramolecular proton transfer in guanine. Post Hartree-Fock study. *Journal of Physical Chemistry B*, *109*, 13770.
- Grimme, S. (2004). Accurate description of van der Waals complexes by density functional theory including empirical corrections. *Journal of Computational Chemistry*, *25*, 1463.
- Gu, J., Wang, J., & Leszczynski, J. (2006). Electron attachment-induced DNA single strand breaks: C₃' - O₃' σ -bond breaking of pyrimidine nucleotides predominates. *Journal of American Chemical Society*, *128*, 9322.
- Gu, J., Wang, J., Rak, J., & Leszczynski, J. (2007a). Findings on the electron-attachment-induced abasic site in a DNA double helix. *Angewandte Chemie*, *119*, 3549.
- Gu, J., Xie, Y., & Schaefer, H. F. (2007b). Electron attachment to DNA single strands: Gas phase and aqueous solution. *Nucleic Acid Research*, *2007*(35), 5165.

- Gu, J., Wang, J., Leszczynski, J., Xie, Y., & Schaefer, H., III. (2008). To stack or not to stack: Performance of a new density functional for the uracil and thymine dimers. *Chemical Physics Letters*, 2008(459), 164.
- Gu, J., Xie, Y., & Schaefer, H. F. (2009). Electron attachment to oligonucleotide dimmers in water: Microsolvation-assisted base-stacking forms. *Chemical Physics Letters*, 473, 213.
- Gu, J., Wang, J., & Leszczynski, J. (2010a). Electron attachment induced DNA single strand breaks at the pyrimidine sites. *Nucleic Acids Research*, 38, 5280.
- Gu, J., Wang, J., & Leszczynski, J. (2010b). Comprehensive analysis of the DNA strand breaks at the guanosine site induced by low energy electron attachment. *ChemPhysChem*, 11, 175.
- Gu, J., Wong, N.-B., Xie, Y., & Schaefer, H. F. (2010c). Electron attachment to a hydrated DNA duplex: The dinucleoside phosphate deoxyguanytyl-3′/5′-deoxycytidine. *Chemistry – A European Journal*, 16, 13155.
- Gu, J., Xie, Y., & Schaefer, H. F. (2010d). Electron attachment to hydrated oligonucleotide dimers: Guanylyl-3′/5′-cytidine and cytidyl-3′/5′-guanosine. *Chemistry – A European Journal*, 16, 5089.
- Gu, J., Wang, J., & Leszczynski, J. (2011a). Stacking and H-bonding patterns of dGpdC and dGpdCpdG: Performance of the M05-2X and M06-2X minnesota density functionals for the single strand DNA. *Chemical Physics Letters*, 512, 108.
- Gu, J., Wang, J., & Leszczynski, J. (2011b). Low energy electron attachment to the adenosine site of DNA. *Journal of Physical Chemistry. B*, 115, 14831.
- Gu, J., Wang, J., & Leszczynski, J. (2012a). Electron attachment to the cytosine-centered DNA single strands: Does base stacking matter? *The Journal of Physical Chemistry. B*, 116, 1458.
- Gu, J., Liang, G., Xie, Y., & Schaefer, H. F. (2012b). Electron attachment to solvated dGpdG: Effects of stacking on base-centered and phosphate-centered valence-bound radical anions. *Chemistry – A European Journal*, 18, 5232.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., & Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, 35, 849.
- Halkier, A., Helgaker, T., Jorgensen, P., Klopper, W., & Olsen, J. (1999). Basis-set convergence of the energy in molecular Hartree-Fock calculations. *Chemical Physics Letters*, 302, 437.
- Heßelmann, A., Jansen, G., & Schütz, M. (2005). Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *Journal of Chemical Physics*, 122, 014103.
- Hobza, P., & Šponer, J. (1998). Significant structural deformation of nucleic acid bases in stacked base pairs: An ab initio study beyond Hartree-Fock. *Chemical Physics Letters*, 288, 7.
- Hobza, P., & Šponer, J. (1999). Structure, energetics, and dynamics of the nucleic acid base pairs: Nonempirical ab initio calculations. *Chemical Reviews*, 99, 3247.
- Hobza, P., Šponer, J., & Reschel, T. (1995). Density-functional theory and molecular clusters. *Journal of Computational Chemistry*, 16, 1315.
- Hudson, J. S., Eberle, J. F., Vachhani, R. H., Rogers, L. C., Wade, J. H., Krishnamurphy, R., & Springsteen, G. (2012). A Unified mechanism for abiotic adenine and purine synthesis in formamide. *Angewandte Chemie, International Edition*, 51, 5134.
- Ishii, K., Tajima, A., Taketsugu, T., & Yamashita, K. (2006). Theoretical elucidation of the unusually high [HNC]/[HCN] abundance ratio in interstellar space: Two-dimensional and two-state quantum wave packet dynamics study on the branching ratio of the dissociative recombination reaction $\text{HCNH}^+ + e^- \rightarrow \text{HNC/HCN} + \text{H}$. *The Astrophysical Journal*, 636, 927.
- Jeziorski, B., Moszynski, R., & Szalewicz, K. (1994). Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chemical Reviews*, 94, 1887.
- Jurecka, P., Cerny, J., Hobza, P., & Salahub, D. R. (2007). Density functional theory augmented with an empirical dispersion term. Interaction energies and geometries of 80 noncovalent complexes compared with ab initio quantum mechanics calculations. *Journal of Computational Chemistry*, 28, 555.
- Kabelác, M., & Hobza, P. (2001). At nonzero temperatures, stacked structures of methylated nucleic acid base pairs and microhydrated nonmethylated nucleic acid base pairs are favored

- over planar hydrogen-bonded structures: A molecular dynamics simulations study. *Chemistry – A European Journal*, 7, 2067.
- Kohler, B. (2010). Nonradiative decay mechanisms in DNA model systems. *Journal of Physical Chemistry Letters*, 1, 2047.
- Kopitz, H., Zivkovic, A., Engels, J. W., & Gohlke, H. (2008). Determinants of the unexpected stability of RNA fluorobenzene self pairs. *ChemBioChem*, 9, 2619.
- Kristyán, S., & Pulay, P. (1994). Can (semi)local density functional theory account for the London dispersion forces? *Chemical Physics Letters*, 229, 175.
- Kruger, K., Grabowski, P. J., Zaugg, A. J., Sands, J., Gottschling, D. E., & Cech, T. R. (1982). Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell*, 31, 147.
- Kumar, A., & Sevilla, M. D. (2007). Low-energy electron attachment to 5'-thymidine monophosphate: Modeling single strand breaks through dissociative electron attachment. *Journal of Physical Chemistry B*, 111, 5464.
- Leontis, N. B., Stombaugh, J., & Westhof, E. (2002). The non-Watson-Crick base pairs and their associated isostericity matrices. *Nucleic Acids Research*, 30, 3497.
- Leszczynski, J. (1992). Are the amino groups in the nucleic acid bases coplanar with the molecular rings? Ab Initio HF/631G and MP2/631G Studies. *International Journal of Quantum Chemistry: Quantum Biology Symposium*, 19, 43.
- Leszczynski, J. (2000). Isolated, solvated, and complexed nucleic acid bases: Structures and properties. In M. Hargittai & I. Hargittai (Eds.), *Advances in molecular structure research* (Vol. 6, p. 209). Stamford: JAI Press.
- Lin, J., Yu, C., Peng, S., Akiyama, I., Li, K., Lee, L. K., & LeBreton, P. R. (1980). Ultraviolet photoelectron studies of the ground-state electronic structure and gas-phase tautomerism of purine and adenine. *Journal of the American Chemical Society*, 102, 4627.
- Marian, C. M. (2007). The guanine tautomer puzzle: Quantum chemical investigation of ground and excited states. *The Journal of Physical Chemistry. A*, 111, 1545.
- Matthews, C. N. (1997). Hydrogen cyanide polymers from the impact of Comet P/Shoemaker/Levy 9 on Jupiter. *Advances in Space Research*, 19, 1087.
- McAllister, M., Smyth, M., Gu, B., Tribello, G. A., & Kohanoff, J. (2015). Understanding the interaction between low-energy electrons and DNA nucleotides in aqueous solution. *Journal of Physical Chemistry Letters*, 6, 3091.
- McMullan, R. K., Benci, P., & Craven, B. M. (1980). The neutron crystal structure of 9-methyladenine at 126 K. *Acta Crystallographica. Section B*, 36, 1424.
- Middleton, C. T., de La Harpe, K., Su, C., Law, Y. K., Crespo-Hernandez, C. E., & Kohler, B. (2009). DNA excited-state dynamics: From single bases to the double helix. *Annual Review of Physical Chemistry*, 60, 217.
- Miller, S. L., & Urey, H. C. (1959). Organic compound synthesis on the primitive earth. *Science*, 130, 245.
- Mix, L. J. (2006). The astrobiology primer: An outline of general knowledge – Version 1, 2006. *Astrobiology*, 6, 735.
- Mons, M., Dimicoli, I., Piuze, F., Tardivel, B., & Elhanine, M. (2002). Tautomerism of the DNA base guanine and its methylated derivatives as studied by gas-phase infrared and ultraviolet spectroscopy. *The Journal of Physical Chemistry. A*, 106, 5088.
- Mons, M., Piuze, F., Dimicoli, I., Gorb, L., & Leszczynski, J. (2006). Near-UV resonant two-photon ionization spectroscopy of gas phase guanine: Evidence for the observation of three rare tautomers. *The Journal of Physical Chemistry. A*, 110, 10921.
- Morgado, C. A., Jurecka, P., Svozil, D., Hobza, P., & Šponer, J. (2009). Balance of attraction and repulsion in nucleic-acid base stacking: CCSD(T)/complete-basis-set-limit calculations on uracil dimer and a comparison with the force-field description. *Journal of Chemical Theory and Computation*, 5, 1524.
- Nir, E., Janzen, C., Imhof, P., Kleinermanns, K., & de Vries, M. S. (2001). Guanine tautomerism revealed by UV–UV and IR–UV hole burning spectroscopy. *Journal of Chemical Physics*, 115, 4604.

- Ponnamperuma, C., Lemmon, R. M., Mariner, R., & Calvin, M. (1963). Formation of adenine by electron irradiation of methane, ammonia, and water. *Proceedings of the National Academy of Sciences of the United States of America*, *49*, 737.
- Rak, J., Chomicz, L., Wiczak, J., Westphal, K., Zdrochowicz, M., Wityk, P., Zyndul, M., Makurat, S., & Golon, L. (2015). Mechanisms of damage to DNA labeled with electrophilic nucleobases induced by ionizing or UV radiation. *The Journal of Physical Chemistry. B*, *119*, 8227.
- Réblová, K., Spackova, N., Štefl, R., Csaszar, K., Koa, J., Leontis, N. B., & Šponer, J. (2003). Non-Watson-Crick base pairing and hydration in RNA motifs: Molecular dynamics of 5S rRNA Loop E. *Biophysical Journal*, *84*, 3564.
- Roy, D., Najafian, K., & von Ragué Schleyer, P. (2007). Chemical evolution: The mechanism of the formation of adenine under prebiotic conditions. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 17272.
- Serrano-Andres, L., & Merchán, M. (2009). Are the five natural DNA/RNA base monomers a good choice from natural selection? *Journal of Photochemistry and Photobiology, C: Photochemistry Reviews*, *10*, 21.
- Sheina, G. G., Stepanian, S. G., Radchenko, E. D., & Blagoi, Y. P. (1987). IR spectra of guanine and hypoxanthine isolated molecules. *Journal of Molecular Structure*, *158*, 275.
- Shukla, M. K., & Leszczynski, J. (2007). Electronic spectra, excited state structures and interactions of nucleic acid bases and base assemblies. *Journal of Biomolecular Structure & Dynamics*, *25*, 93.
- Shukla, M. K., & Leszczynski, J. (2008). Computational study of UV-induced excitations of DNA fragments. In *Radiation induced molecular phenomena in nucleic acids* (Challenges and advances in computational chemistry and physics, Vol. 5). Dordrecht: Springer.
- Siegfried, N. A., Metzger, S. L., & Bevilacqua, P. C. (2007). Folding cooperativity in RNA and DNA is dependent on position in the helix. *Biochemistry*, *46*, 172.
- Smith, I. W. M., Talbi, D., & Herbst, E. (2001). The production of HCN dimer and more complex oligomers in dense interstellar clouds. *Astronomy & Astrophysics*, *369*, 611.
- Soliva, R., Laughon, C. A., Luque, F. J., & Orozco, M. (1998). Molecular dynamics simulations in aqueous solution of triple helices containing d(G.C.C) trios. *Journal of the American Chemical Society*, *120*, 11226.
- Solomon, T., Seitz, H., & Sturm, H. (2009). DNA damage by low-energy electron impact: Dependence on guanine content. *Journal of Physical Chemistry B*, *113*, 11557.
- Spackova, N., Berger, I., Egli, M., & Šponer, J. (1998). Molecular dynamics of hemiprotonated intercalated four-stranded i-DNA: Stable trajectories on a nanosecond scale. *Journal of the American Chemical Society*, *120*, 6147.
- Šponer, J., & Hobza, P. (1994). Bifurcated hydrogen bonds in DNA crystal structures. An ab initio quantum chemical study. *Journal of the American Chemical Society*, *116*, 709.
- Šponer, J., & Kypr, J. (1991). Different intrastrand and interstrand contributions to stacking account for roll variations at the alternating purine-pyrimidine sequences in A-DNA and A-RNA. *Journal of Molecular Biology*, *221*, 761.
- Šponer, J., & Kypr, J. (1993). Theoretical analysis of the base stacking in DNA: Choice of the force field and a comparison with the oligonucleotide crystal structures. *Journal of Biomolecular Structure & Dynamics*, *11*, 277.
- Šponer, J., Leszczynski, J., & Hobza, P. (1996a). Structures and energies of hydrogen-bonded DNA base pairs. A nonempirical study with inclusion of electron correlation. *The Journal of Physical Chemistry*, *100*, 1965.
- Šponer, J., Leszczynski, J., & Hobza, P. (1996b). On the nature of nucleic acid base stacking. Nonempirical ab initio and empirical potential characterization of 10 stacked base pairs. Comparison of stacked and H-bonded base pairs. *The Journal of Physical Chemistry*, *100*, 5590.
- Šponer, J., Leszczynski, J., Vetterl, V., & Hobza, P. (1996c). Base stacking and hydrogen bonding in protonated cytosine dimer: The role of molecular ion-dipole and induction interactions. *Journal of Biomolecular Structure & Dynamics*, *13*, 695.
- Šponer, J., Gabb, H. A., Leszczynski, J., & Hobza, P. (1997). Base-base and deoxyribose-base stacking interactions in B-DNA and Z-DNA. A quantum-chemical study. *Biophysical Journal*, *73*, 76.

- Sponer, J., Mokdad, A., Sponer, J. E., Spackova, N., Leszczynski, J., & Leontis, N. B. (2003). Unique tertiary and neighbor interactions determine conservation patterns of Cis Watson-Crick A/G base-pairs. *Journal of Molecular Biology*, *330*, 967.
- Sponer, J., Jurecka, P., & Hobza, P. (2004). Accurate interaction energies of hydrogen-bonded nucleic acid base pairs. *Journal of the American Chemical Society*, *126*, 10142.
- Sponer, J. E., Leszczynski, J., Sychrovský, V., & Šponer, J. (2005a). The sugar edge/sugar edge base pairs in RNA. Stabilities and structures from quantum chemical calculations. *Journal of Physical Chemistry B*, *109*, 18680.
- Sponer, J. E., Spackova, N., Leszczynski, J., & Šponer, J. (2005b). Principles of RNA base pairing: Structures and energies of the trans Watson-Crick/sugar edge base pairs. *Journal of Physical Chemistry B*, *109*, 11399.
- Sponer, J., Jurecka, P., Marchan, I., Luque, F. J., Orozco, M., & Hobza, P. (2006). Nature of base stacking. Reference quantum chemical stacking energies in ten unique B-DNA base pair steps. *Chemistry – A European Journal*, *12*, 2854.
- Sponer, J. E., Réblová, K., Mokdad, A., Sychrovský, V., Leszczynski, J., & Šponer, J. (2007). Leading RNA tertiary interactions: Structures, energies, and water insertion of A-minor and P-interactions. A quantum chemical view. *Journal of Physical Chemistry B*, *111*, 9153. and references cited therein.
- Sponer, J., Riley, K. E., & Hobza, P. (2008). Nature and magnitude of aromatic stacking of nucleic acid bases. *Physical Chemistry Chemical Physics*, *10*, 2595.
- Sponer, J., Zgarbova, M., Jurecka, P., Riley, K. E., Sponer, J. E., & Hobza, P. (2009). Reference quantum chemical calculations on RNA base pairs directly involving the 2-OH group of ribose. *Journal of Chemical Theory and Computation*, *5*, 1166.
- Sponer, J., Sponer, J. E., Petrov, A. I., & Leontis, N. B. (2010). Quantum chemical studies of nucleic acids can we construct a bridge to the RNA structural biology and bioinformatics communities? *Journal of Physical Chemistry B*, *114*, 15723.
- Stombaugh, J., Zirbel, C. L., Westhof, E., & Leontis, N. B. (2009). Frequency and isostericity of RNA base pairs. *Nucleic Acids Research*, *37*, 2294.
- Suzuki, M., Amano, N., Kakinuma, J., & Tateno, M. (1997). Use of a 3D structure data base for understanding sequence-dependent conformational aspects of DNA. *Journal of Molecular Biology*, *274*, 421.
- Svozil, D., Hobza, P., & Sponer, J. (2010). Comparison of intrinsic stacking energies of ten unique dinucleotide steps in A-RNA and B-DNA duplexes. Can we determine correct order of stability by quantum-chemical calculations? *The Journal of Physical Chemistry. B*, *114*, 1191.
- Szalewicz, K., & Jeziorski, B. (1998). Comment on “on the importance of the fragment relaxation energy terms in the estimation of the basis set superposition error correction to the intermolecular interaction energy”. *Journal of Chemical Physics*, *109*, 1198 [Journal of Chemical Physics, *104*, 8821 (1996)].
- Tennekes, P. P., Harju, J., Juvela, M., & Toth, L. V. (2006). HCN and HNC mapping of the protostellar core Chamaeleon-MMS1. *Astronomy & Astrophysics*, *456*, 1037.
- Vlieghe, D., Sponer, J., & Van Meervelt, L. (1999). Crystal structure of d(GGCCAATTGG) complexed with DAPI reveals novel binding mode. *Biochemistry*, *38*, 16443.
- Wang, J., Gu, J., Nguyen, M. T., Springsteen, G., & Leszczynski, J. (2013a). From formamide to purine: An energetically viable mechanistic reaction pathway. *The Journal of Physical Chemistry. B*, *117*, 2314.
- Wang, J., Gu, J., Nguyen, M. T., Springsteen, G., & Leszczynski, J. (2013b). From formamide to purine: A self-catalytic reaction pathway provides a feasible mechanism for the entire process. *The Journal of Physical Chemistry. B*, *117*, 9333.
- Wang, J., Gu, J., Nguyen, M. T., Springsteen, G. & Leszczynski, J. (2013c). From Formamide to Adenine: A Self-Catalytic Mechanism for an Abiotic Approach. *J. Phys. Chem. B*, *117*, 14039.
- Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vornrhein, C., Hartsch, T., & Ramakrishnan, V. (2000). Structure of the 30S ribosomal subunit. *Nature*, *407*, 327.
- Wing, R., Drew, H., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1980). Crystal structure analysis of a complete turn of B-DNA. *Nature*, *287*, 755.

- Yamazaki, S., Domcke, W., & Sobolewski, A. (2008). Nonradiative decay mechanisms of the biologically relevant tautomer of guanine. *The Journal of Physical Chemistry. A*, *112*, 11965.
- Yanagi, K., Prive, G. G., & Dickerson, R. E. (1991). Analysis of local helix geometry in three B-DNA decamers and eight dodecamers. *Journal of Molecular Biology*, *217*, 201.
- Yildirim, I., & Turner, D. H. (2005). RNA challenges for computational chemists. *Biochemistry*, *44*, 13225.
- Yildirim, I., Stern, H. A., Sponer, J., Spackova, N., & Turner, D. H. (2009). Effects of restrained sampling space and nonplanar amino groups on free-energy predictions for RNA with imino and sheared tandem GA base Pairs. Flanked by GC, CG, iGiC or iCIG base pairs. *Journal of Chemical Theory and Computation*, *5*, 2088.
- Zhao, Y., & Truhlar, D. G. (2008). Density functionals with broad applicability in chemistry. *Accounts of Chemical Research*, *41*, 157.
- Zhou, J., Kostko, O., Nicolas, C., Tang, X., Belau, L., de Vries, M. S., & Ahmed, M. (2009). Experimental observation of guanine tautomers with VUV photoionization. *The Journal of Physical Chemistry. A*, *113*, 4829.
- Zirbel, C. L., Sponer, J. E., Sponer, J., Stombaugh, J., & Leontis, N. B. (2009). Classification and energetics of the base-phosphate interactions in RNA. *Nucleic Acids Research*, *37*, 4898.

Metal Interactions with Nucleobases, Base Pairs, and Oligomer Sequences; Computational Approach

42

Jaroslav V. Burda, Jiří Šponer, and Filip Šebesta

Contents

Introduction	1828
Properties of Metal Adducts with Nucleobases	1829
Interactions of Bare Cations with Bases	1829
Metal Interactions in Implicit Solvent Model	1834
Interactions of Explicitly Hydrated Cations	1835
Interactions with Platinum Metal Complexes	1840
The Metal Coordination to Multiple Nucleobase Systems	1847
Interaction with Base Pairs	1847
Enhancement of Base Pairing by Pt Complexes	1856
Metal Cross-Linked DNA Structures	1858
Ions in DNA Quadruplexes	1863
Conclusions	1865
Bibliography	1866

Abstract

This chapter concerns some of the computational studies devoted to interactions of metal cations with nucleobases, nucleotides, and short oligonucleotides considered as DNA/RNA models. Our topic is fairly complex, therefore the results obtained using mainly *ab initio* and DFT methods are discussed. The first part focuses on the interactions of isolated bases with metal cations either in bare, hydrated, or ligated forms. We begin with interactions of naked cations with nucleobases in gas phase. Subsequently, solvation effects using polarizable continuum models are analyzed together with a comparison to explicitly hydrated

J.V. Burda (✉) • F. Šebesta

Faculty of Mathematics and Physics, Charles University in Prague, Prague 2, Czech Republic
e-mail: burda@karlov.mff.cuni.cz

J. Šponer

Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno, Czech Republic
e-mail: sponer@ncbr.chemi.muni.cz

© Springer Science+Business Media Dordrecht (outside the USA) 2017

J. Leszczynski et al. (eds.), *Handbook of Computational Chemistry*,

DOI 10.1007/978-3-319-27282-5_36

1827

ions. In the second part, adducts of various metals with base pairs and oligomeric models of DNA/RNA are discussed. Separate sections are devoted to complexes of promising anticancer drugs. Stacked bases and larger systems (quadruplexes) studied by semiempirical and QM/MM methods are mentioned in the last part.

Introduction

This chapter provides a brief overview of recent model calculations and simulations of metal-cation interactions with nucleic acid (NA) bases, base pairs, and short oligomeric models. From historical reasons, alkali metals and metals of alkaline earth are discussed first. All possible forms of metal cations are considered, starting with bare ions, which were mostly studied in older papers. Later, complexes with their first solvation shell or in ligated form are reviewed. Despite the simplicity and the fact that bare cations overestimate the bonding energies because of exaggerated contribution of Coulomb interaction, these models enable a clear insight into the basic bonding characteristics and other electronic properties. Many studies are focused on transition metal complexes, motivated by anticancer activity of cisplatin and other possible metallodrugs. While the electrostatic contribution dominates in the interaction of alkali metals, the coordination covalent term is important in the complexes of transition metals.

In general, the structure and function of DNA are modified by metal ions. These ions can interact with many sites in DNA (Saenger 1983; Sigel 1993), including the phosphate groups, the sugar moiety, and the NA bases. Although the metal cations usually interact with the phosphate group and, to a lesser extent, with bases, cation-base interactions are involved in many important biophysical processes, such as different stabilization of DNA triple helices (Potaman and Soyfer 1994), stabilization of quadruple helices (Hud et al. 1996), and stabilization of the ribose-base stacking in Z-DNA (Egli and Gessner 1995). It is assumed that the interaction of a divalent cation with the base can cause a significant polarization of the bases associated with stabilization of certain H-bonded base pairs and other interactions (Hud et al. 1996; Egli and Gessner 1995; Anwander et al. 1990; Santangelo et al. 2010). However, some ions do not interact with nucleic acids in a direct manner but rather contribute to unspecific loose ion atmosphere around the nucleic acids (Chu et al. 2008). While theoretical studies usually deal with binding of cations to DNA, the RNA-cation binding is much more biochemically important and diverse. Many folded RNAs contain indispensable specific structural ions, which may also be directly involved in RNA catalysis (Draper et al. 2005; Banas et al. 2009; Lilley and Eckstein 2008).

In the DNA double helix, the known sites for the cation coordination are mainly the N7 atoms of purines sometimes supported by forming the chelate structure involving also O6 atom of guanine. Some other sites, such as N3 of cytosine or thymine and N1 of adenine, are usually blocked by the hydrogen bonding. Nevertheless, several experimental as well as theoretical examples of coordination into this position exists especially in the metal bridges like cytosine-silver-cytosine (Šponer et al. 1999a) or thymine-mercury-thymine (Yamaguchi et al. 2014; Šebera

et al. 2013a) as mentioned below. It should be noted that the metal cation interactions are not restricted only to the NA bases and the cation can simultaneously interact with the phosphate group. The coordinated metal cations can interact with the NA base directly (inner-sphere coordination) or mediated by a water molecule, which links the NA base with the metal cation (outer-sphere coordination) (Sigel 1993; Egli et al. 1991). Experimental studies on 5'-monophosphates revealed the following order of macrochelate coordination involving the N7 position of purines: GMP > IMP (I – inosine) > AMP (Sigel 1993). This order was explained as a result of different basicity of the N7 sites of guanosine, inosine, and adenine. It also correlates with the dipole moments of NA bases.

Properties of Metal Adducts with Nucleobases

Interactions of Bare Cations with Bases

The first models investigated in connection with metal-nucleobases interactions employed bare cations. One of the pioneering studies, which considered the influence of water and divalent ions on the base pairing, was published in 1970 by Rozsnyai and Ladik (1970). Later Del Bene (1984) explored interactions of naked Li⁺ cation with all DNA bases. In this work, all basic features of modern quantum chemical calculations are already present – the structures were optimized at the HF/STO-3G level with single point calculations (SP) using double-zeta basis set (6-31G). As the most stable adduct, the [Li-(Gua-O6,N7)]⁺ chelate was found (with association energy of 78.4 kcal/mol – which can be compared with more recent value of 78.3 at MP2/6-31G(d,p) (Burda et al. 1996)) followed by cytosine complex [Li-(Cyt-O2,N3)]⁺ (77.1 kcal/mol). The association energy of the most stable adenine conformer [Li-(Ade-N3)]⁺ was estimated to be substantially lower – 48.4 kcal/mol (40.4 at MP2 (Burda et al. 1996)). These energies clearly demonstrate the dominant role of the monopole (cation)-dipole moment (nucleobase) electrostatic interaction. When the multipolar formula for energy evaluation between monopole and dipole is considered, $E = q_{Li} \cdot \mu_{gua} / R^2_{Li-gua}$, then ca 92 kcal/mol can be obtained for Li charge of 0.84 e, guanine dipole moment of 7.28 D, and mutual distance of 2.09 Å (supposing a collinear orientation of cation and vector of dipole moment (cf. for instance Burda and Gu 2008; Futera et al. 2009a)). In the study, both O2 and O4 thymine complexes were predicted to be more stable than the adenine adducts by about 10 kcal/mol. Closely after this keystone study, another work dealing with interactions of the A-T and G-C base pairs with Li⁺ cation appeared (Del Bene 1985). In the most stable Li⁺-A-T conformer, the Li cation was coordinated to O2 site of thymine. In the G-C pair, the N3 position of guanine was preferred by 4 kcal/mol over O6, N7 chelate structure. This A-T and G-C base pair was later explored with extended set of metal cations (Burda et al. 1997) roughly confirming all these results. We would like to address here one important point. The numbers above, as well as many other numerical values in this study, refer to gas phase interactions of metal cations with nucleic acids

Table 1 Preferred sites for metalation and protonation

Base	Metalation	Protonation
Adenine	N7 > N1	N1 > N7
Cytosine	N3	N3
Guanine	N7 > N1	N7 > N1
Thymine, uracil	O2 > O4	O4 > O2

components. Such interactions are dominated by the electrostatic effects, which are substantially (sometimes almost completely) extinct in nucleic acids as well as in typical bioinorganic experiments. This needs to be kept in mind while the results are interpreted. For more discussion of various aspects of the interplay between the gas phase interactions and real systems, see, e.g. refs. (Šponer et al. 2001; Schmidt et al. 2002; Burda et al. 2003; Petrov et al. 2005).

Metal coordination sites in natural nucleotides are: (a) oxygen atoms of phosphate groups where cations neutralize the negative charge (these sites can be considered quite unspecific for any cation), (b) hydroxyl groups of sugar moiety usually chosen by alkali metals or metals of alkaline earth, (c) nitrogen atoms of heterocyclic bases – especially N1, N3, and N7 atoms of purine and N3 of pyrimidine bases, which exhibit large affinity to cations or generally electrophiles (In DNA/RNA oligomeric sequences, only sites in minor or major groove are accessible for interactions. It means that solely the N3 and N7 sites of purine bases are available for the interactions.), and (d) oxygens of keto-groups (O2 of cytosine, O6 of guanine, and O2, O4 of thymine and uracil) as positions for binding of “hard” cations, e.g., alkali metals (and less frequently for transition metal). Preference for individual hetero atoms in all five NA bases was explored for Ru(II) cation also in study of Futera (Futera et al. 2009a) demonstrating the preference for N7 site of purine bases and (after deprotonation) N3 of pyrimidine.

Comparing cation coordination and protonation of these active sites, large similarities can be noticed in the case of guanine, cf. Table 1. On the contrary, more remarkable differences are observed for adenine, uracil, or thymine. Unfortunately, deprotonation of N3 in thymine/uracil is not considered.

Coordination of the Mg²⁺ cation to purine DNA bases was recently explored in many experimental as well as computational studies (e.g., ref. (Elmahdaoui and Tajmirriahi 1995; Sychrovsky et al. 2004)), where bonding properties of selected mono- and divalent metal cations with the N7 position of guanine and adenine were examined.

The interaction of bare monovalent (alkali metals and coinage metals) and divalent (alkaline earth and zinc group metals) cations with the N7 site of purine NA bases was investigated in gas phase using the MP2/6-31G(d,p)//HF/6-31G(d,p) level of theory (Burda et al. 1996). The stabilization energies of base...M complexes were determined according to the formula:

$$\Delta E^{\text{Stab}} = - [E^{\text{Complex}} - (E_{BSE}^{\text{base}} - E_{BSE}^{\text{metal}})] + \Delta E_{\text{deform}}^{\text{base}} \quad (1)$$

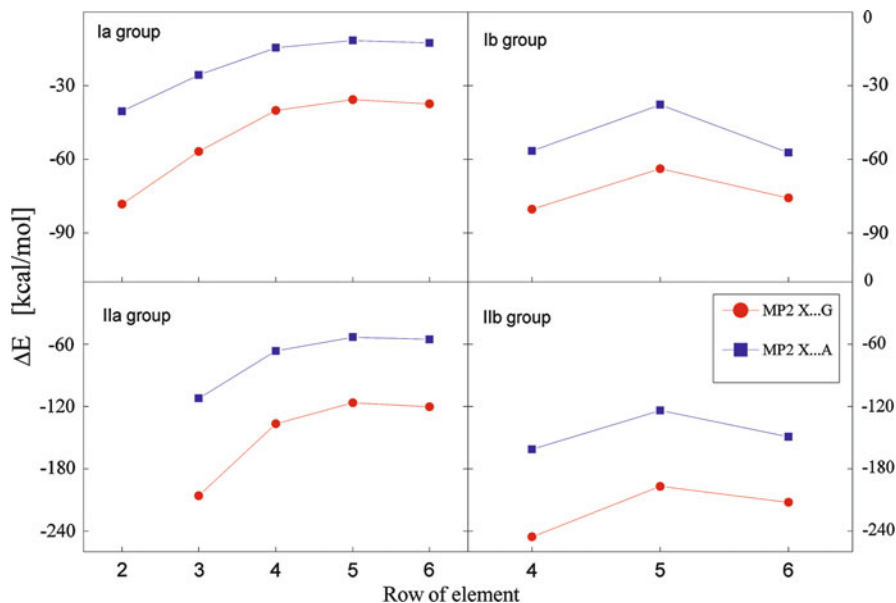


Fig. 1 Stabilization energies of metal-base complexes for Ia, Ib, IIa, and IIb metal cations interacting in N7(O6) position of the base

Here the $E_{B SSE}^{\text{base}}$ represents total energy of the base (adenine or guanine) within the basis set superposition error (BSSE) scheme of Boys Bernardi (Boys and Bernardi 1970). The results are displayed in Fig. 1 and it is evident that the values for guanine complexes are systematically nearly twice as large as those of adenine. This is due to the larger dipole moment of guanine and also its more favorable orientation close to the direction forming M-N7 bond, as it can be seen in Fig. 2a, c. The stabilization energies of the complexes with divalent ions are larger than those with monovalent ions and, as could be expected, the stabilization energies decrease with increasing atomic number of the metal ions. The only exception is revealed for the Au^+ and Hg^{2+} complexes where, due to the more pronounced relativistic effects, the respective stabilization energies are larger than the energies of the preceding cations (Ag^+ and Cd^{2+}).

In the coinage and zinc-group metal complexes, the bonding interaction is markedly stronger in comparison with coordination of the Ia and IIa metals. This is due to the presence of lower-energy vacant s-orbitals (compared with the same orbitals of the alkali metals), enabling to some extent the dative bonding into these orbitals from the occupied MO's of the bases. This leads to the increase of covalent character of the interaction that explains the basic difference between coordination of, e.g., K^+ and Cu^+ cations.

Complexes with Li^+ bare cation were computationally explored in gas phase by Russo et al. (Russo and Toscano 2001) at the DFT(B3LYP) level within several

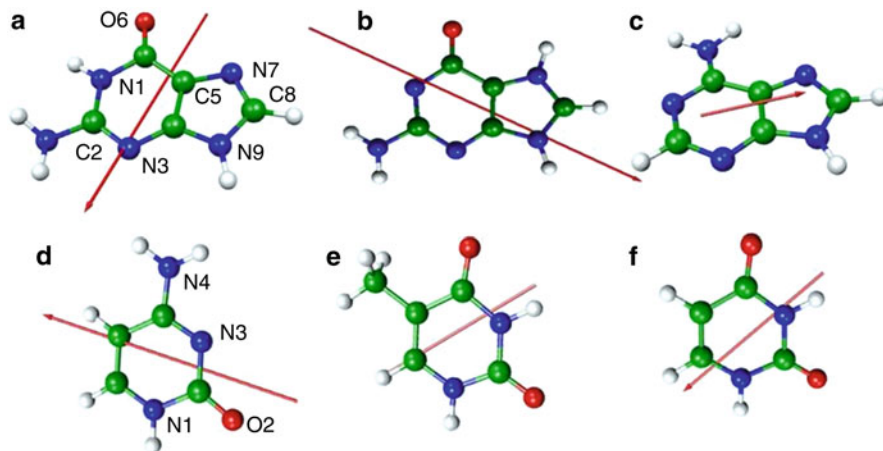
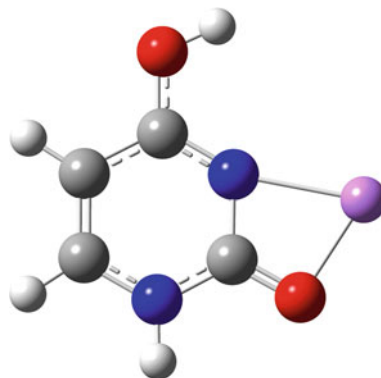


Fig. 2 Orientation and size of dipole moments in the isolated NA bases: (a) guanine – 6.98 D, (b) N7 protonated guanine tautomer – 9.16 D, (c) adenine – 2.43 D, (d) cytosine – 6.64 D, (e) thymine – 4.47 D, (f) uracil – 4.54 D determined at B3LYP/6-311++G(2df,2pd)//B3PW91/6-31G(d,p) in gas phase. Standard atom numbering is used (e.g., textbook of Saenger (1983))

Fig. 3 Structure of the $[\text{Li}-(\text{Ura-O2,N3})]^+$ cation (enol form of O2,N3-chelate)



basis sets. They found that the most stable complex of adenine with lithium is imino-tautomer with N6,N7 coordination. In this way, a five-membered chelate is formed where Li^+ -N6 bond is 1.971 Å, and Li^+ -N7 distance is 2.022 Å. The remaining explored (amino) structures, $[\text{Li}-(\text{Ade-N3,N9})]^+$, $[\text{Li}-(\text{Ade-N6,N7})]^+$, and $[\text{Li}-(\text{Ade-N3})]^+$, lie by about 2, 10, and 15 kcal/mol higher on the potential energy surface. In the guanine adducts, the most stable structure is $[\text{Li}-(\text{Gua-O6,N7})]^+$ followed by chelate $[\text{Li}-(\text{Gua-N3,N9})]^+$, its enol form, and the enol form of the (O6,N7)-chelate, which are by 11, 12, and 15 kcal/mol less stable. The most stable complex with uracil is $[\text{Li}-(\text{Ura-O4})]^+$ adduct. The chelate structures $[\text{Li}-(\text{Ura-O2,N3})]^+$ (cf. Fig. 3) and $[\text{Li}-(\text{Ura-O4,N3})]^+$ lie about 3 and 7 kcal/mol above the global minimum and the $[\text{Li}-(\text{Ura-O2})]^+$ adduct is about 5 kcal/mol above

the O4-conformer. This fact indicates the electrostatic origin of the interaction when one considers the direction of the uracil dipole moment. In the cytosine complexes, the $[\text{Li}-(\text{Cyt-O2,N3})]^+$ structure is the most stable minimum followed by enol-tautomer of $[\text{Li}-(\text{Cyt-N1,O2})]^+$ (where N1 proton is transferred to O4 site) and the enol form of $[\text{Li}-(\text{Cyt-O2,N3})]^+$. These two complexes lie about 12 and 17 kcal/mol higher on the potential energy surface. The same strategy was also applied for evaluation of Mg^{2+} and Ca^{2+} in their next paper (Russo et al. 2003). Later Marino et al. (2007, 2010) and Bagchi et al. (2012) performed DFT calculations on the interaction of naked zinc group metals with NA bases in gas phase obtaining qualitative agreement with our energies (Burda et al. 1996). Moreover pyrimidine bases are explored, too. Altogether, the preference for binding to various NA bases is $G > C > A > T > U$. We performed another study on the $\text{Ru(II)Cl(en)(arene)}$ piano-stool complex where various sites of all NA were also investigated (Futera et al. 2009a). In this study ligated Ru(II) cation, similar preferences followed for gas phase calculations: $G > A > C > T > U$. When solvated (CPCM) model was considered, the order of binding preferences changed: $G > C > T > U > A$. Nevertheless, it should be stated that the difference between binding to G and A is only about 9 kcal/mol and all three last bases (T, U, and A) are in 1 kcal/mol energy span.

The coordination of divalent metal cations with the phosphate group of various nucleotides (GMP, AMP, UMP, and CMP) was studied (probably) for the first time by Varnali (Varnali and Tozumcalgan 1995). In her study semiempirical PM3 method was used. From the results it follows that the most stable metal adducts are formed with the phosphate group of AMP closely followed by CMP for all explored metals.

The calculations of the NMR spin-spin coupling constants and the NMR shifts of the direct and water-mediated binding of divalent metal cations to guanine were performed by Sychrovský (Sychrovsky et al. 2004). The intermolecular coupling constants $^1J(X, \text{O6})$ and $^1J(X, \text{N7})$ ($X = \text{Mg}^{2+}, \text{Zn}^{2+}$) were unambiguously assigned to the specific binding motif of the hydrated cation with O6 and N7 sites of guanine. The calculated coupling constants $^1J(\text{Mg}, \text{O6})$ and $^1J(\text{Zn}, \text{O6})$ are 6.2 and -17.5 Hz for the inner-shell complex where the cation directly interacts with the guanine O6 position. For the inner-shell coordination of the cation at nitrogen N7, the calculated coupling constants $^1J(\text{Mg}, \text{N7})$ and $^1J(\text{Zn}, \text{N7})$ are 5.6 and -36.5 Hz, respectively. When the cation binding is water-mediated, the corresponding coupling constants are zero. The calculated NMR shifts $\delta(\text{N7}) = -15.3$ and -12.2 ppm upon the coordination of Mg^{2+} and Zn^{2+} ion are similar to the NMR shift of 19.6 ppm toward the high field measured by Tanaka (Tanaka et al. 2002) for the coordination of Cd^{2+} to the N7-guanine site.

The B3LYP/6-311 + G(2df,2p) level was used to explore geometry of all possible adducts originating from the interaction of Cu^{2+} cation with the most stable tautomers of NA free bases (Marino et al. 2004). Several attachment sites for both purine and pyrimidine bases have been taken into account for possible formation of both monoadducts and chelates. It was concluded that the copper ion Cu(II) has the highest affinity for the most stable tautomer of guanine base. Further, a comparison of various divalent metal cation complexes (Zn, Cu, Ni) with hypoxanthine and

uracil was performed by Matsubara (Matsubara and Hirao 2002). The B3LYP level stabilization energies of both M(II)-hypoxanthine and M(II)-uracil complexes reflect the strength of the M-N(base) interaction giving the same stability sequence $Zn > Cu > Ni$ for both bases.

Metal Interactions in Implicit Solvent Model

The calculated interaction energies of the bare cations with nucleobases have revealed a great overestimation of the electrostatic interaction in comparison with experimental samples. This overestimation is clearly due to the uncompensated charge of the bare cation in the gas-phase calculation compared to water solution where the charge of the cation is screened by solvent molecules. In this way, the electrostatic part of the metal-base interaction is substantially reduced. A role of the electrostatic contribution in the case of the Pt-base coordination will be enlightened later. The approach of implicit solvent models is very popular in many variants (COSMO (Frank Eckert 2006; Klamt and Schuurmann 1993), PCM (Mennucci et al. 1998; Miertus et al. 1981), SMD (Marenich et al. 2009), GBM (Roux et al. 1990; Mukhopadhyay et al. 2014) etc.), and there is a large number of such studies. Here only a few recent works will be mentioned for illustration.

In 2008, Ai et al. (2008) published a study on the tautomer equilibrium of adenine in presence of Zn^{2+} cation at the DFT level (B3LYP/6-311 + G**). It was found that the $[Zn-(Ade-N6,N7)]^{2+}$ imino complex is the most stable structure in accord with the gas-phase calculations of Kabeláč (Kabeláč and Hobza 2006). The latter calculations explored tautomers of all DNA bases in presence of bare Na^+ , Mg^{2+} , and Zn^{2+} cations evaluated at the RI-MP2/TZVPP level of theory.

The binding of first-row transition metal monocations (Sc^+-Cu^+) to N7 of guanine and N7 or N3 of adenine nucleobases was studied using the DFT approach with B3LYP functional. The electrostatic character of these interactions is mainly represented by metal-ligand repulsion. The M^+ -guanine binding energies are about 18–27 kcal/mol larger than those of M^+ -adenine, when the differences decrease along the row (Noguera et al. 2008).

Another study on an interaction of bare cations with metabolite of purine bases-uric acid should be mentioned (Allen et al. 2006). Geometries of the complexes of Li^+ , Na^+ , K^+ , Be^{2+} , Mg^{2+} , and Ca^{2+} metal cations with various nucleophilic sites of uric acid were optimized at the B3LYP/6-311++G(d,p) level. Single point energy calculations were performed at the MP2/6-311++G(d,p) level. It was found that cations mainly form chelate structures with a bidentate coordination. In the gas phase, the most preferred position for the interaction of Li^+ , Na^+ , and K^+ monovalent cations is between the N3 and O2 sites, while all divalent cations prefer coordination between the N7 and O6 sites of the urate. However, it was simultaneously shown that the incorporation of aqueous solution in the model has a significant impact on the relative stability of various complexes. The global minimum of urate with Mg^{2+} and Ca^{2+} cations is represented by the O2,N3-chelates

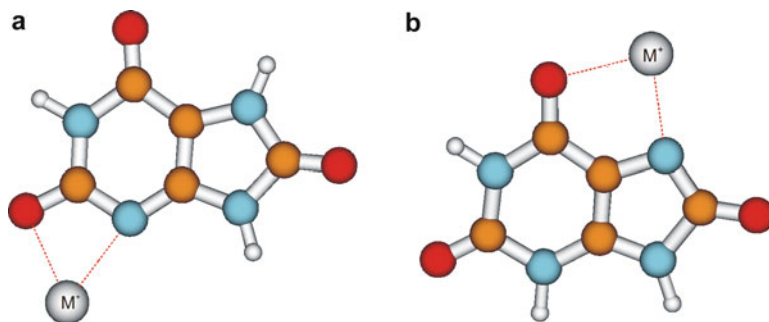


Fig. 4 The optimized structures of the metal complexes with urate: (a) O2,N3-chelate, (b) O6,N7-chelate

in analogy with monovalent cations. Moreover, the relative energy differences are very small. Especially for the Ca^{2+} structures, energies of all conformers are practically in the 2 kcal/mol range. The most stable structures are depicted in Fig. 4.

Metal cation binding to deoxyguanosine monophosphate was examined by Bouř (Andrushchenko and Bour 2009). Infrared spectra of complexes with Na^+ , Mg^{2+} , Ca^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+} , and Cd^{2+} cations were recorded and interpreted on the basis of the density functional theory. The solvation effects were simulated by PCM and cluster models (combined explicit solvent and PCM). The coordination to the guanine N7 position was considered and the obtained calculated spectra predict a blueshift of the characteristic 1578 cm^{-1} frequency, in accord with experimental data. Metal binding to the phosphate group causes significant spectral changes in the sugar-phosphate vibration region but also notable frequency shifts of the carbonyl vibrations. The Cu^{2+} and Zn^{2+} cations induce the largest changes in measured vibrational absorption, which corresponds to the computationally determined strong interaction energies in the N7-complexes. The Cu^{2+} binding to guanine was revealed to be a two-step process, which was also confirmed by the microcalorimetry titration curve.

Interactions of Explicitly Hydrated Cations

The most realistic approach to the description of metal cation interactions with nucleobases is represented by a model of the explicitly hydrated cations – usually up to hexacoordinated cations. Recently, explicit hydration has been used in connection with implicit solvation shell (combined model), e.g., in the study for hydration of Ag^+ by Martínez et al. (1997), for acid-base properties of goethite surface model (Aquino et al. 2008) or for modeling of electrified interfaces (Benedikt et al. 2013) or already mentioned study of Bouř (Andrushchenko and Bour 2009).

Table 2 Interaction energies in the complexes of solvated cation (Mw)-(N7)purine (G/tG/A)

	ΔE_{Mw-G}	ΔE_{Mw-tG}	$\Delta E_{MG\text{bare}}$	ΔE_{Mw-A}	$\Delta E_{MA\text{bare}}$
Mg^{2+}	-89.3	-84.8	-198.7	-46.0	-107.9
Ca^{2+}	-82.6	-80.1	-133.9	-33.5	-61.6
Sr^{2+}	-76.0	-	-	-28.9	-48.9
Ba^{2+}	-71.2	-71.4	-118.3	-28.1	-54.4
Zn^{2+}	-93.8	-90.1	-237.2	-53.7	-152.9
Cd^{2+}	-87.9	-107.6	-192.6	-45.9	-116.6
Hg^{2+}	-94.3	-149.2	-208.0	-55.3	-141.1

Hydrated Alkaline Earth and Zinc-Group Metal Cations

One of the first studies on this topic was published by Šponer et al. (1998a). In the study, pentaqua cation adducts (of Mg^{2+} , Ca^{2+} , Sr^{2+} , Ba^{2+} , Zn^{2+} , Cd^{2+} , and Hg^{2+}) to the N7 position of guanine were explored at the MP2/6-31G*//HF/6-31G* level. Quasi-relativistic energy-averaged Stuttgart pseudopotentials were used for the description of the third and higher-row cations. The interaction between hydrated cation and guanine is significantly reduced compared to the guanine-unsolvated cation model, cf. Table 2. The cations of IIb group together with Mg^{2+} are tightly bound to the N7 atom of guanine, whereas the O6 site is involved in H-bonding with the aqua ligands from the metal hydration shell. The metals with greater radius (Ca^{2+} , Sr^{2+} , and Ba^{2+}) prefer simultaneous coordination to the N7 and O6 atoms of the base. Also, the cation-guanine distance increases with the atomic number. The energy difference between the N7 and N7-O6 types of coordination is rather small. Relativistic effects are apparent in the case of Hg^{2+} , similarly to the complexes with the bare cation reported above (Burda et al. 1996). The Zn^{2+} and Mg^{2+} cations show a different balance between the cation-base and cation-water interactions. While the Zn^{2+} cation is bound more tightly to the base (93 kcal/mol) and its water shell is more flexible (203 kcal/mol), a different situation (with stronger metal-water binding) was found in the Mg^{2+} complex. The binding energy of the $Mg(H_2O)_5^{2+}$ cation with N7 position of guanine in GC pair is approximately 89 and binding energy of five water molecules with $Mg(GC)$ complex is about 220 kcal/mol. The different binding picture for Zn and Mg cations can be partly explained as a shift from the interaction between nucleobase and hydrated cation (Mg) toward the hydration of a metalated base (Zn).

It follows from Table 2 that despite the substantial reduction of the M-N7 interaction energy, the ratio of the energies for guanine and adenine remains approximately equal to two, similarly to the results obtained in the study of the bare cations (Burda et al. 1996). This confirms the dominant role of electrostatic term in these complexes. In platinum complexes, substantially higher coordination-covalent contribution is demonstrated in the limit of total charge going to zero (cf. in Figs. 9 and 15, below).

The tautomeric equilibrium and hydrogen bonding in nucleotide 2'-deoxyguanosine monophosphate in interaction with hydrated $[Mg(H_2O)_4]^{2+}$

cation were studied at the MP2/cc-pVDZ//B3LYP/cc-pVDZ and B3LYP/aug-cc-pVTZ//B3LYP/cc-pVDZ levels of theory by Gorb (Kosenkov et al. 2008). The Mg^{2+} ion forms two inner-shell contacts with the nucleotide, similar to small phosphorylated molecules under physiological conditions. The hydrated magnesium cation in the presence of the phosphate group can change the guanine tautomeric equilibrium in comparison to free guanine. The canonical O6-oxo form of guanine is more stable (by 6–8 kcal/mol) than the O6-hydroxo form in anticonformation. The interaction with Mg^{2+} ion is capable to suppress a spontaneous transient formation of the rare tautomer. The question, which coordination mode of hydrated Mg^{2+} cation with DNA is dominant (direct metal-N7 coordination or indirect interaction by water molecule through the polarized H-bonding), was addressed by Bandyopadhyay (Bandyopadhyay and Bhattacharyya 2003). Based on HF and DFT calculations, the authors show that both binding modes are of a similar importance.

An interesting situation was revealed when interactions of thioguanine (tG) base was explored (Šponer et al. 1999b). In accord with the Pearson HSAB principle (Parr and Pearson 1983), a stronger interaction with S6 site of thioguanine is observed for the heavier transition metal cations, as can be seen by corresponding columns ΔE_{Mw-G} and ΔE_{Mw-tG} in Table 2. While a mild weakening of the binding energy is even visible in the case of alkaline earth cations (Mg^{2+} , Ca^{2+}), a very robust adduct is formed with the Hg^{2+} cation. This strong coordination is also partially the consequence of a smaller number of water molecules in the first hydration shell. Due to this strong Hg-S coordination, two water molecules are pushed out from the first hydration shell. Optimized N7,S6-chelate structure resembles the arrangement with the bare cation, cf. Fig. 5b. However, a shorter Hg-S distance (2.50 Å) was obtained comparing the Hg-N bond (2.57 Å). A similar chelate structure was also found in the Cd^{2+} complex, despite the coordination with four water molecules, cf. Fig. 5c. A slightly longer Cd-S distance (2.67 Å)

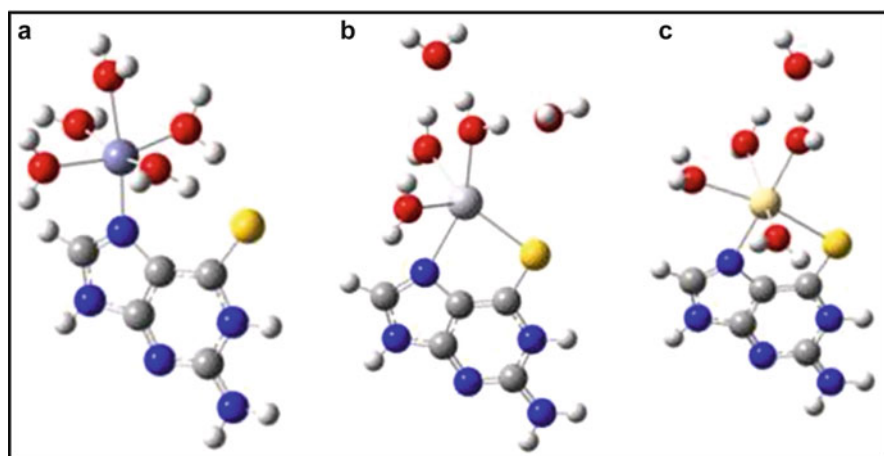


Fig. 5 Different cases of hydrated-metal coordination to thioguanine: (a) Zn, (b) Hg, and (c) Cd

than Cd-N (2.40 Å) was obtained in that structure. In the remaining complexes no direct metal-sulphur interaction was found at the HF/6-31G* optimization level. Nevertheless, using the DFT level, slightly different coordination pattern is revealed for Mg and Zn cations. The preferred structure contains six coordinated cation sphere but only with four water molecules similar to the chelate arrangement of the Cd²⁺ complex (Šponer et al. 1999b) as shown in the Fig. 5c.

Possible binding of hydrated cations to nucleobases in a cation- π manner led to the conclusion that such interactions are very unlikely in the case of nucleic acids as they are outcompeted by conventional cation binding patterns. The suggestion of existence of cation- π interactions in DNA was shown to be a case of misinterpretation of structural data (Šponer et al. 2000a).

Later extended study by Rulisek (Rulisek and Šponer 2003) explored adducts of hydrated divalent metal cations (Mg²⁺, Cu²⁺, Zn²⁺, and Cd²⁺) to the guanine N7 position interacting in outer- and inner-shell manner with an anionic phosphate group. The results obtained by the B3LYP and MP2 computational level are compared. Their calculations illustrate that the cations have a different balance of water-cation and N7-cation binding, leading to selectivity toward the N7 binding in the following order: Cu²⁺ >> Zn²⁺ = Cd²⁺ > Mg²⁺. It is consistently reflected by changes of interatomic distances upon cation substitution, changes of both absolute electronic and interaction energies. The energetics of the cation selectivity is typically on a scale of 3–10 kcal/mol. The calculations demonstrate that the inner-shell binding to the phosphate is more stable than the outer-shell binding. Nevertheless, this does not mean that such a conclusion can be generalized for real samples where the phosphate group is included in the form of sugar-phosphate backbone.

Complexes of Hydrated Copper Cations with Guanine

Interaction of hydrated Cu(I)/Cu(II) cations with guanine represents another interesting explored system. In this case also redox properties under hydration can be compared to copper complexes with water and ammonium ligands (Pavelka et al. 2008; Burda et al. 2005). One of the examined features of the copper complexes involves the number of coordinated ligands. While the Cu(I) complexes prefer two-coordinated structures, the geometry of Cu(II) complexes has usually the coordination number four or five, in accord with the small inorganic copper complexes studied previously (Pavelka and Burda 2005; Pavelka et al. 2006). Also, the affinity of various active sites of the metalated guanine for water molecules in comparison with nonmetalated (isolated) base was explored. The resulting preference for individual conformers determined at the B3LYP/6-311++G(2df,2pd) level can be seen in Table 3 and the most stable structures are displayed in Fig. 6. The affinity of individual sites of the isolated bases was explored by Poltev et al. (1996) using empirical potentials. They found that in the case of guanine hydration by a single water molecule, the global minimum structure has one water molecule between O6 and N1 sites, followed by the N1-N2 water adduct. The N9 position was not considered since this position is not accessible in nucleosides. In another study (Colominas et al. 1996), various tautomeric forms of guanine and cytosine

Table 3 The relative differences of Gibbs' free energies ΔG and stabilization energies ΔE^{Stab} and Cu-N7 and Cu-O binding energies for the selected conformers in kcal/mol

Conformer	ΔG	ΔE^{Stab}	$E(\text{Cu-N})$	$E(\text{Cu-O})$
1	0.0	-168.4	83.4	60.4
2	1.1	-167.9	81.0	59.9
3	2.6	-165.3	80.4	60.7
4	3.5	-166.5	74.4	60.6
5	3.6	-167.7	76.5	60.1

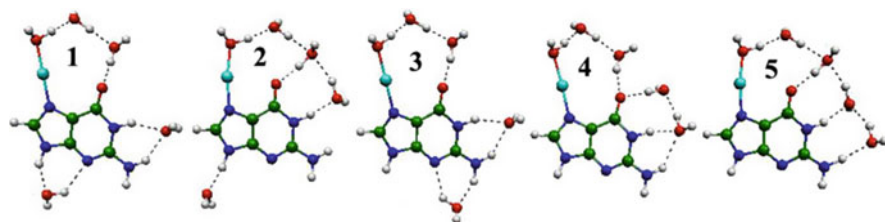


Fig. 6 The most stable conformers of $[\text{Cu}(\text{H}_2\text{O})_5(\text{N7-guanine})]^+$ complex

Table 4 Relative differences of Gibbs energies ΔG and stabilization energies ΔE^{Stab} (in kcal/mol) with respect to global minima structure. Abbreviation *c.n.* stands for the coordination number and *struct.* is used for identification of the optimized structure in Fig. 7

System	<i>c.n.</i>	<i>Struct.</i>	ΔG^{total}	$\Delta \Delta E^{\text{stab}}$
$[\text{CuG}(\text{H}_2\text{O})_4]^{2+}$	4	4 <i>b</i>	2.2	4.6
	4 ^{chel}	4 <i>j</i>	0.0	0.0
	4 ^{chel}	4 <i>k</i>	2.0	1.0
	5 ^{chel}	4 <i>g</i>	4.3	4.9
$[\text{CuG}(\text{H}_2\text{O})_5]^{2+}$	4	5 <i>a</i>	2.8	0.0
	4 ^{chel}	5 <i>w</i>	0.0	1.7
	4 ^{chel}	5 <i>y</i>	1.1	1.0
	5 ^{chel}	5 <i>h</i>	6.4	5.4

and their influence on Watson-Crick H-bonding were explored. From these results, it is clear that there is no substantial change due to the polarization effects of Cu(I) cation revealed by semiempirical methods, in comparison with the MP2 approach. Also, metal-N7 binding energies can be compared with results of our earlier studies (Burda et al. 1996, 1997) on metal coordination to the purine DNA bases. Despite the fact that in the earlier investigations only bare cations were considered at the MP2/6-31G(d)//HF/6-31G(d) level, the interaction energy is fairly similar to the analogous energies of these hydrated structures.

The adducts of the hydrated copper(II) cation with guanine were explored in the study of Pavelka (Pavelka et al. 2006). Various numbers of water molecules were considered in combination with the different coordination pattern of the Cu^{2+} cation. The most stable structures are summarized in Table 4 and displayed in Fig. 7. The full set of investigated structures can be viewed in the original paper (Pavelka et al. 2006).

From this study it follows that the penta-coordination of Cu(II) in these complexes is visibly less convenient than that in the case of the small inorganic complexes (for both aqua ligands (Burda et al. 2004) or mixed aqua-ammine ligands

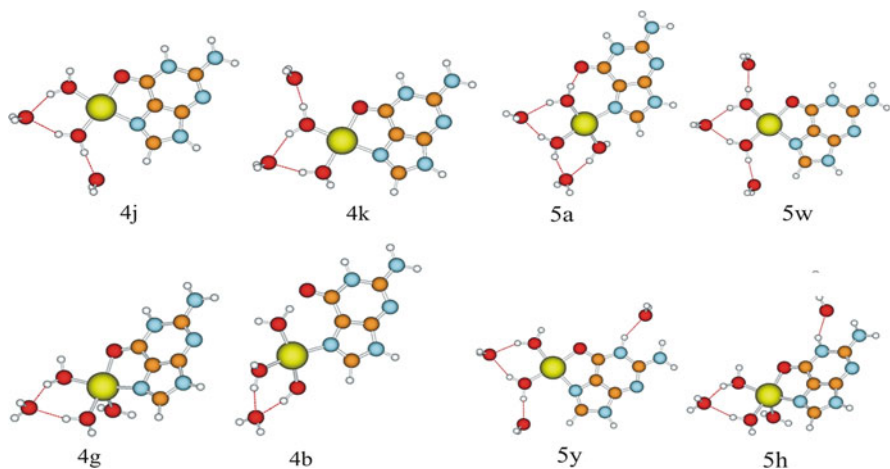


Fig. 7 The most stable structures of Cu-guanine complexes with four and five water molecules

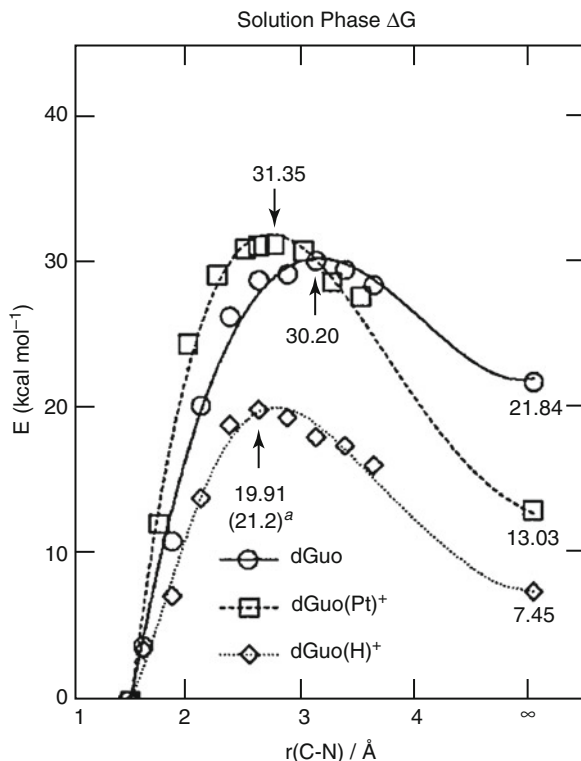
(Pavelka and Burda 2005)). The Cu(I) complexes do not create chelate structures since linear monoaqua form with the remaining water molecules in the solvation shell is substantially more stable. In the monoaqua Cu(I) and Cu(II) complexes, the strength of Cu-N7 bond ($-81/-230$ kcal/mol for Cu(I)/Cu(II)) and Cu-O(aq) bond ($-35/-46$ kcal/mol) roughly follows formal electrostatic relationship. The higher preference for N7 coordination in Cu(II) complex is related to the possibility of the higher electron transfer from more polarizable nitrogen atom.

Interactions with Platinum Metal Complexes

One of the most frequently studied metals in connection with nucleobases is platinum. Since late sixties when B. Rosenberg published his study on anticancer activity of cisplatin (Rosenberg et al. 1965, 1969), a lot of effort is devoted to this compound and its derivatives, as well as to some other transition metal complexes where similar activities can be expected. One of the early calculations was performed by Basch et al. (1986) and by Lipinski (1989). Since then a vast number of both experimental and theoretical studies can be found in literature on this topic, especially on platinum interactions with nucleobases, nucleotides, and other DNA models. One of the most important questions, which need to be solved, concerned tautomeric equilibria due to the possible point mutations. This topic was addressed in several studies considering various nucleobases. Metallodrugs are generally expected to coordinate predominately in the major groove of genomic DNA. Therefore our discussion starts with the interaction in N7 position of purine bases.

An interesting study on the difference between protonation and metalation of the N7 position of deoxyguanosine was published by Baik et al. (2002). While under

Fig. 8 Reaction free energy profile for dissociation of N9-C1' glycoside bond of dGuo, dGuo(H)⁺, and dGuo(Pt)⁺ in solvent



protonation the glycoside N9-C1' bond breaks, the Pt(II) adduct does not change the strength of the glycosidic bond substantially, as can be noticed from the Fig. 8.

The Tautomeric Equilibrium of the Metalated Nucleobases

Several platinum complexes with varying total charge were explored at B3LYP/6-31G(d) level (Burda et al. 2000): neutral trans-dichloro-ammine-platinum, +1 charged trans-diammine-chloro-platinum, and +2 charged triammine-platinum. Two tautomers of guanine were considered – keto and enol forms as well as N6-amino and syn- and anti-imino forms of adenine. In this way the role of electrostatic contribution could be elucidated. Despite the gas-phase calculations, the bonding energy and tautomeric relations in solvent can be easily estimated. In Fig. 9, the relative tautomeric stabilization energies are displayed for adenine and guanine.

From this figure it follows that the trans-imino-tautomer of adenine is better stabilized under platination, which could lead to mispairing. In the case of guanine, the regular keto form is more stable for charged complexes. Moreover, the enol preference in the electroneutral complex is caused by the additional H-bonding stabilization between platinum ligand (aqua or chloro) with hydrogen from the enol group (cf. Fig. 10a). Such stabilization may not be present in solvent due to competitive H-bonding interactions with other water molecules. This study in

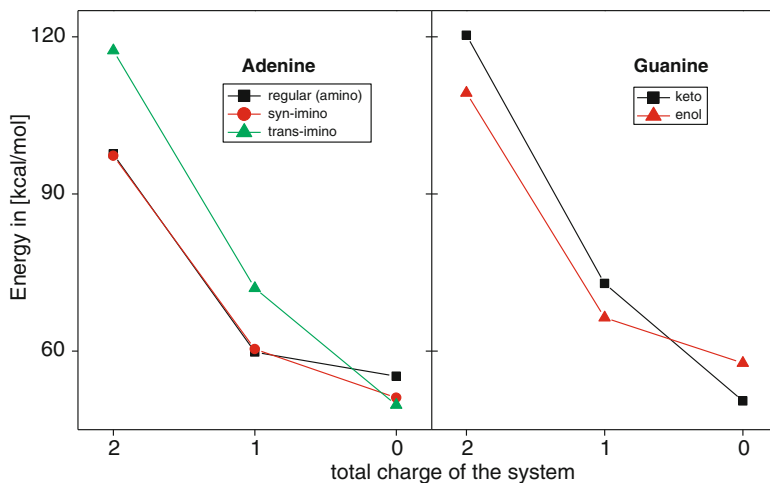


Fig. 9 Stabilization energies of the platinum-base complexes in dependence on total charge of the (a) Pt-adenine and (b) Pt-guanine complex

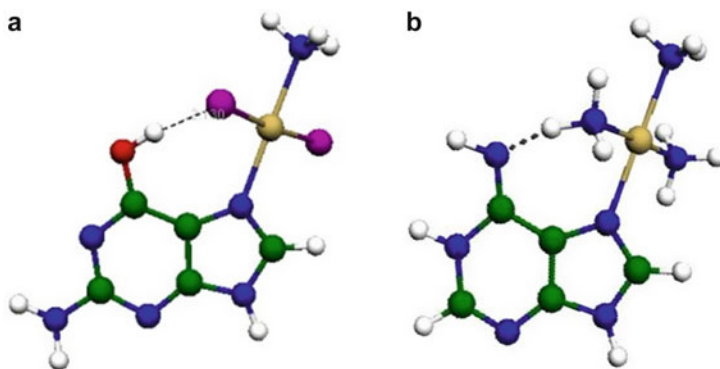


Fig. 10 Structure of platinum(II) complexes with (a) enol-tautomer of guanine and (b) trans-imino-tautomer of adenine

general suggests that the largest part of the tautomeric effects of the N7 platination is due to gas phase effects, which are assumed to be annihilated in real assays.

Metalation of the exocyclic amino group of cytosine and adenine nucleobases was explored in the case Pt(II) and Hg(II) complexes (Šponer et al. 1999c). It induces protonation of the N3 site of cytosine and N1 site of adenine. Hence, it causes a proton shift from an exocyclic to an endocyclic N atom (similar to situation in Fig. 10b). In this way the metal-assisted process can lead to the generation of rare nucleobase tautomers. The calculations demonstrate that metalation of the exocyclic amino group of nucleobases significantly increases the protonation energy of the aromatic rings of nucleobases by about 30–34 kcal/mol for the Pt(II) adduct and by about 10–14 kcal/mol for the Hg(II) adduct. This study demonstrates a

tautomeric shift that is caused by changes of the electronic structure of nucleobases and is unrelated to electrostatic effects. Thus, authors suggested a separation of purely electrostatic effects from nonelectrostatic (molecular orbital) contributions. The former ones are expected to be rather unimportant in aqueous solution or x-ray crystallography experiments, where the systems are overall strictly neutral. The nonelectrostatic contributions are assumed to remain insensitive to the environment and fully expressed under usual experimental conditions.

Rare tautomers of 1-methyluracil (MeUH) and 1-methylthymine (MeTH) in coordination with Pt(II) complexes were explored by van der Wijst et al. (2009). Comparing the calculations in gas phase and water, the influence of the solvation effects was estimated. They also showed that relative stabilization energies of the Pt(II) complexes with various tautomers of MeUH and MeTH differ from the isolated tautomers. This leads to the conclusion that some rare tautomers may become favored under metalation.

Recently, a reduction mechanism of $\text{Pt}^{\text{IV}}(\text{DACH})\text{Cl}_4$ in the presence of dGMP was explored (Šebesta and Burda 2016). The first step represents a substitution reaction of a chloro-ligand by dGMP followed by a nucleophilic attack of phosphate or sugar oxygen to the C8 position of guanine. Subsequently, the reduction reaction occurs forming the $\text{Pt}^{\text{II}}(\text{DACH})\text{Cl}_2$ complex. The whole process is completed by hydrolysis. Two different pathways for the substitution reaction were examined: direct associative and Basolo-Pearson autocatalytic mechanism. While activation barriers for direct association were too high, the barriers based on Basolo-Pearson mechanism are used for estimation of the rate constants (according to Eyring's TST) and are compared with experimental values. It was found that the rate determining step is the nucleophilic attack with a slightly faster performance in 3'-dGMP branch than in 5'-dGMP with the activation barrier of 21.1 and 20.4 kcal · mol⁻¹ comparing with experimental values of 23.8 and 23.2 kcal · mol⁻¹, respectively. The reduction reaction is connected with the electron flow from guanine to Pt complex (cf. Fig. 11).

The product of reduction reaction is a chelate structure, which dissociates within the last reaction – hydrolysis. The whole redox process (substitution, reduction, and hydrolysis reactions) is exoergic by 34 and 28 kcal · mol⁻¹ for 5'-dGMP and 3'-dGMP, respectively.

Interaction of Nucleobases with Half-Sandwich Ru(II) Complexes

Recently, computations on ruthenium complexes with nucleobases were published by Futera (Futera et al. 2009b). In his comprehensive study, interactions of piano-stool ruthenium metallodrug with nucleobases were investigated. In the same study, the reaction profile of the chloro-ligand replacement by water molecule and the role of the arene-ring size were also explored. From the point of this review, an important part of Futera's study deals with interactions of Ru(II) with various sites of all the nucleobases in vacuum and in the implicit solvent model. For the optimized structures, the bonding and stabilization energies were determined. Characteristics of the most stable conformers are summarized in Table 5. From the Table 5 it follows that in the case of purine bases, the most stable adenine structure (Ru-(Ade-N1) conformer) differs from the most stable guanine complex (Ru-(Gua-N7) conformer)

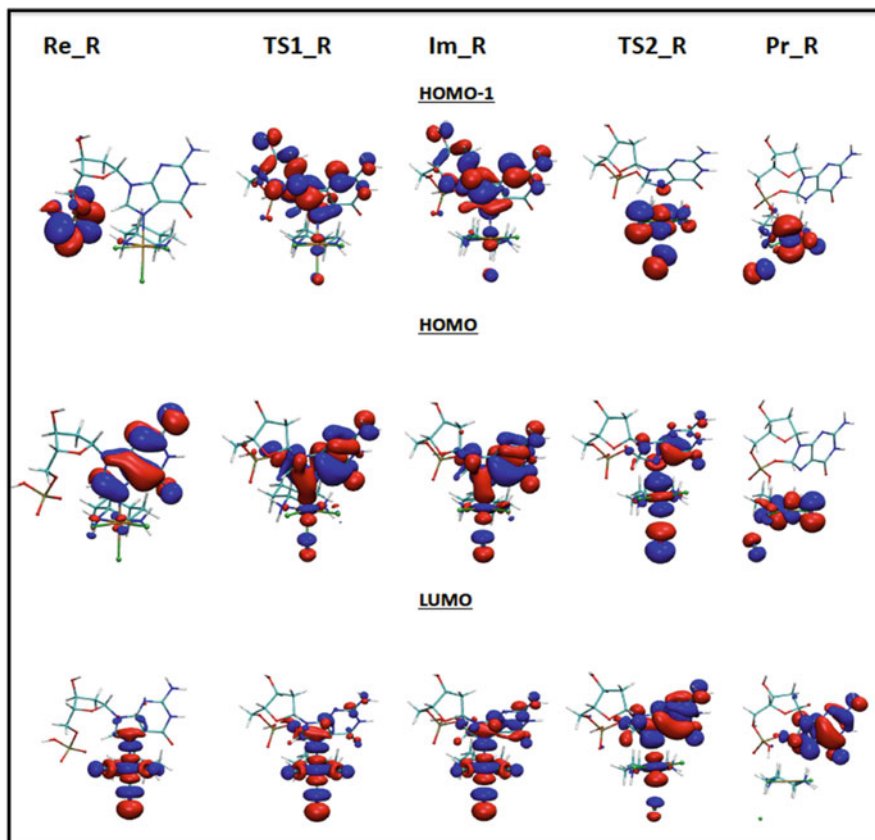


Fig. 11 Electronic flux from HOMO to LUMO of the supermolecular Tetraplatin/dGMP system

as revealed in the gas-phase calculations. Nevertheless, in the PCM model, the N7-conformer represents the most stable form of the both purine nucleobases. This change of the Ru-adenine global minimum follows from the general reduction of the electrostatic forces, which are substantially screened in PCM approaches. In gas phase, the Ru-(Ade-N1) coordination is enhanced by favorable orientation of the adenine dipole moment of about 2.70 D (determined at the MP2/6-31++G(d,p), gas-phase level) aiming in N1-C8 direction while the guanine dipole of 6.36 D points in N7-N3 direction, as can be seen in Fig. 2. Even larger dipole moment was predicted for N1-guanine tautomer with proton transferred to N7 atom ($\mu = 9.55$ D). This dipole moment correlates with the fact that the Ru-(Gua-N1) adduct exhibits the highest $\Delta E^{\text{BE}}(\text{Ru}-(\text{base}-\text{N}))$ binding energy among all the explored complexes (interestingly, in both gas phase and PCM approach). The analogous effect of decreased electrostatic enhancement is also responsible for the change in energy preference in the case of cytosine adducts. The Ru-(Cyt-O2) structure becomes

Table 5 Binding and stabilization energies (in kcal/mol)

	Ru(bz)(A-N1)		Ru(bz)(A-N3)		Ru(bz)(A-N7)		Ru(cym)(A-N7)	
Adenine	g.p.	COSMO	g.p.	COSMO	g.p.	COSMO	g.p.	COSMO
$\Delta E^{BE}(A)$	-72.7	-48.7	-69.1	-47.7	-69.7	-50.3	-65.4	-50.2
ΔE^{Stab}	431.1	389.2	426.9	387.9	426.2	390.0	441.4	396.5
Guanine	Ru(bz)(G-N1)		Ru(bz)(G-N3)		Ru(bz)(G-N7)		Ru(cym)(G-N7)	
$\Delta E^{BE}(G)$	-103.9	-59.7	-56.4	-45.3	-90.7	-56.0	-86.1	-57.2
ΔE^{Stab}	442.7	390.9	412.6	384.7	449.7	395.3	465.6	383.3
Cytosine	Ru(bz)(C-N1)		Ru(bz)(C-N3)		Ru(bz)(C-O2)			
$\Delta E^{BE}(C)$	-89.5	-56.0	-79.0	-52.8	-84.0	-45.9		
ΔE^{Stab}	441.1	392.6	438.4	392.2	444.6	389.2		
Thymine	Ru(bz)(T-N1)		Ru(bz)(T-N3)		Ru(bz)(T-O2)		Ru(bz)(T-O4)	
$\Delta E^{BE}(T)$	-65.0	-46.5	-82.8	-51.3	-53.9	-31.3	-60.0	-34.7
ΔE^{Stab}	407.4	372.3	424.9	380.3	412.5	374.0	420.2	376.2
Uracil	Ru(bz)(U-N1)		Ru(bz)(U-N3)		Ru(bz)(U-O2)		Ru(bz)(U-O4)	
$\Delta E^{BE}(U)$	-61.6	-45.9	-79.8	-51.1	-50.8	-31.7	-59.1	-34.3
ΔE^{Stab}	403.2	371.2	422.7	380.5	410.5	375.3	419.3	377.3

the least stable in implicit water model (by more than 7 kcal/mol) and the most stable adduct is Ru-(Cyt-N1) closely followed by the Ru-(Cyt-N3) complex with practically the same stabilization and total electronic energies. The thymine and uracil bases favor the N3 coordination regardless the environment. This preference is in accord with different orientation of the dipole moment of uracil and thymine in comparison with cytosine (cf. Fig. 2). Another computational study dealing with interaction of different forms of ruthenium complexes with DNA bases was published recently (Besker et al. 2007). Here similar binding energies (in comparison to Futera's work) were obtained for relevant structures in both gas phase and PCM.

Molecular mechanism for the replacement of chloro-ligand by water and subsequently of aqua-ligand by nucleobase was also explored by Futera (Futera et al. 2009c). Similarly to the platination process, the hydration reaction is mildly endoergic ($\Delta G = 2$ kcal/mol). Formation of the guanine adduct is exoergic by ca 7 kcal/mol. Nevertheless, all these reactions are faster in the case of the ruthenium complex (in equimolar concentrations) since the activation barriers are lower, e.g., the values for replacement of both adenine and guanine are about 18 kcal/mol while in the cisplatin case analogous barriers are ca 20 kcal/mol (cf. Table 8). Moreover, while "simple" mechanism was found in adenine reaction (as displayed in Fig. 12a), a two-step reaction mechanism was suggested for the guanine replacement (Fig. 12b).

In these mechanisms, a lower activation barrier for the substitution by adenine corresponds to higher rate constant ($1.7 \text{ M}\cdot\text{s}^{-1}$ vs. $0.5 \text{ M}\cdot\text{s}^{-1}$). Nevertheless, since the minimum associated with the reaction coordination ("direct reactant") of adenine is about 4 kcal/mol higher than the global minimum it means the

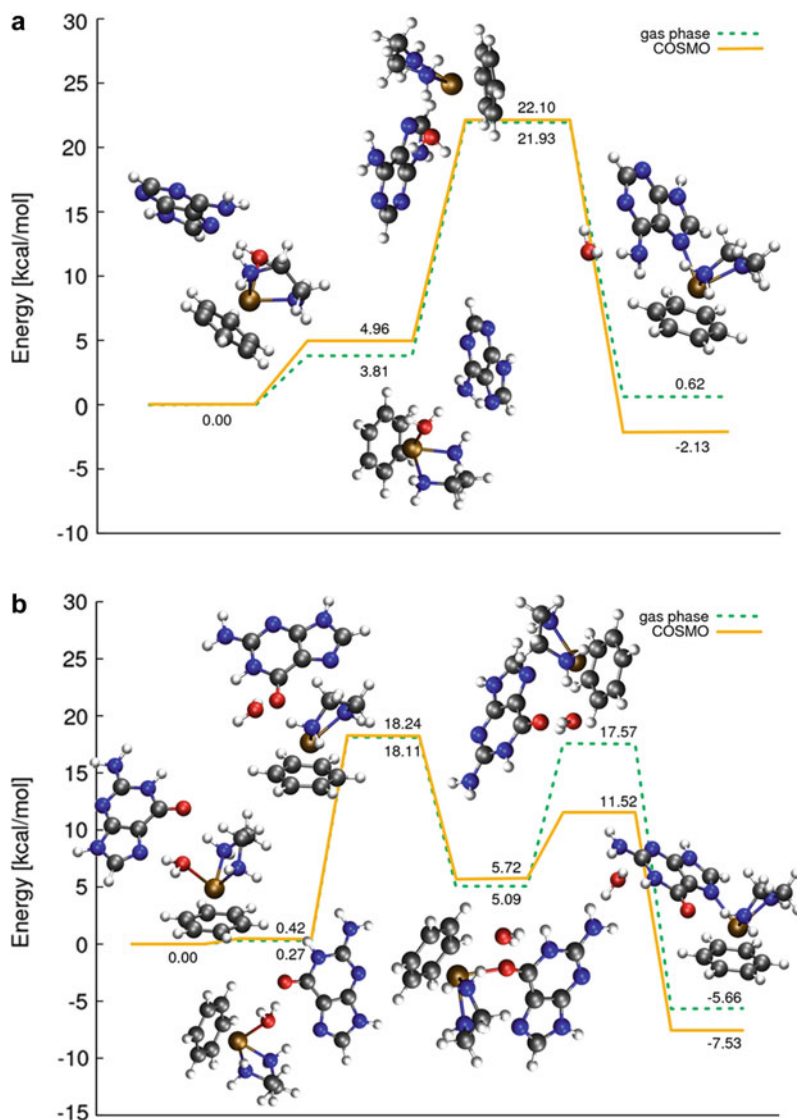


Fig. 12 (a) Reaction coordinate for replacement of water by adenine and (b) guanine

concentration of the conformer corresponding to this local minimum will be by three orders of magnitude lower (according to the Boltzmann equilibrium law). A different situation occurs for the analogous local minimum of guanine. Here the instant concentration (under assumption of equilibrium) of the local reaction minimum is comparable to the global minimum and therefore, the real reaction rate will be actually substantially higher in the reaction with guanine.

Osmium complexes were also considered as possible metallodrugs. The hydrogenation energies of various nitrogen heterocycles in the presence of osmium tetroxide were investigated and published by Deubel (2003). While hydrogenation of pyrimidine bases is exothermic, the purine C4-C5 bond does not have a tendency to hydrogenate.

Another study on tautomeric equilibria and connected proton transfer (PT) from the N3 site to O2 site of thymine was explored by Burda et al. (2016). It is shown that PT in the presence of additional water molecule reduces activation barrier from 51 kcal/mol to ca 22 kcal/mol. Nevertheless, when solvated cations ($[\text{Mg}(\text{H}_2\text{O})_6]^{2+}$, $[\text{Zn}(\text{H}_2\text{O})_6]^{2+}$, and $[\text{Hg}(\text{H}_2\text{O})_6]^{2+}$) are coordinated to the O2 and O4 site of thymine by two of their aqua ligands, then further reduction of the activation barrier to ca 6–8 kcal/mol was observed due to polarization of these ligands.

The Metal Coordination to Multiple Nucleobase Systems

Interaction with Base Pairs

Metal Cations from Ia, Ib, IIa, and IIb Groups

An influence of metal coordination in major groove on the enhancement of base pairing was explored by Burda et al. (1997). The energy decomposition for these systems requires, besides total stabilization and pair energies, also the nonadditive three-body contribution.

The studied complexes were partitioned into three subsystems: the two bases (B_1 , B_2) and a metal cation (M). The total stabilization energy (ΔE^{Stab}) is, within the BSSE counterpoise error, defined as:

$$\Delta E^{\text{Stab}} = -\{E(B_1, B_2, M) - [E(B_1, gB_2, gM) + E(gB_1, B_2, gM) + E(gB_1, gB_2, M)]\} \quad (2)$$

where $E(B_1, B_2, M)$ represents the total energy of the whole complex, and, e.g., $E(B_1, gB_2, gM)$ is a total energy of the base B_1 in the presence of the basis functions on the ghost systems B_2 and M . Alternatively, the total stabilization energy ΔE can be expressed in terms of pair stabilization energies and the three-body contribution:

$$\Delta E^{\text{Stab}} = E(B_1 - B_2) + E(B_1 - M) + E(B_2 - M) + E(3) \quad (3)$$

where each pair-stabilization energy is calculated within the BSSE scheme. For example, the $E(B_1 - B_2)$ energy can be determined from formula:

$$E(B_1 - B_2) = -\{E(B_1, B_2, gM) - [E(B_1, gB_2, gM) + E(gB_1, B_2, gM)]\} \quad (4)$$

Besides these pairwise energies, interactions of one subsystem of the complex (metal or pyrimidine) with the remaining part were also evaluated. The interaction of thymine with (metal + adenine) subsystem can be determined as:

$$E(MA - T) = -\{E(A, T, M) - [E(A, gT, M) + E(gA, T, gM)]\} \quad (5)$$

The whole systems can be regarded as a composition of a strongly bonded metal cation-purine base part plus two weak interactions: metal cation-remote pyrimidine base and Watson-Crick H-bonded base pair. The latter two terms represent only a small perturbation of the first one, and their mutual influence is basically very small. Therefore, it is not surprising that similar geometry and energy parameters for metal coordination were obtained, comparing with the analogous characteristics in ref. (Burda et al. 1996). Also, the obtained geometries match well with the results reported by Anwander et al. (1990) for complexes with Ca^{2+} , Mg^{2+} , and Zn^{2+} . Some small differences originate from a smaller basis set, which they have used: too short distances between the purine base and the metal cation, especially for the Zn^{2+} complexes (our calculations predict 1.95 Å for the Zn^{2+} GC complex, while Anwander reported 1.72 Å).

The metal cation significantly influences the geometry of the base pair. The (C)O2...H-N2(G) H-bond lengths in the GC complexes are systematically reduced, in comparison with the isolated pair. This reduction is the largest for bivalent ions (up to 0.3 Å). The central H-bond N3-H...N1 remains practically unchanged, and the third N4-H...O6 H-bond, which is the closest one to the metal cation, is significantly lengthened in comparison with the isolated GC pair. The elongation is again the largest for bivalent ions (0.65 Å in Zn^{2+} GC and Mg^{2+} GC). In the AT pair, the metal cations affect the H-bonds in a different way. The (T)O4...H-N6 H-bond, which is closer to the metal-ion coordination site, shows a substantial shortening (0.35 Å in complexes with Zn^{2+} and Mg^{2+}), while the other H-bond (N3-H...N1) is lengthened (by 0.18 Å in Zn^{2+} AT complex). The geometric rearrangements of the pair structures can be regarded as rotation around the center of the pyrimidine ring towards the metal cation in the case of the AT pair and away from it in the case of the GC pair.

Basically, the same dependencies of the stabilization energies on increasing atomic numbers of metal cations are observed in metal-purine-pyrimidine complexes like in previously published metal-purine species (Burda et al. 1996). Stabilization energies of complexes with divalent ions are larger than those of monovalent ions, and M-GC stabilization energies are larger than those for M-AT complexes. Both conclusions reflect the dominant role of the ion-dipole electrostatic contribution to the stabilization energy of these complexes.

Compared with the study of Anwander (Anwander et al. 1990), a very close agreement was obtained for complexes of Ca^{2+} with base pairs (within 5 kcal/mol). However, larger differences were found in Mg^{2+} -containing systems (≈ 20 kcal Mg^{2+} -AT and ≈ 40 kcal Mg^{2+} -GC; the values in our work describe larger stabilization – for both base pairs). Nevertheless, the use of a

minimal basis set (MBS) for zinc (all-electron calculations) nearly doubles the stability ($\Delta E^{\text{HF}}(\text{MBS}:\text{Zn}^{2+}\text{-GC}) = 448$ kcal/mol vs. our value $\Delta E^{\text{HF}}(\text{AREP}:\text{Zn}^{2+}\text{-GC}) = 254$ kcal/mol and $\Delta E^{\text{HF}}(\text{MBS}:\text{Zn}^{2+}\text{-AT}) = 328$ kcal/mol vs. $\Delta E^{\text{HF}}(\text{AREP}:\text{Zn}^{2+}\text{-AT}) = 158$ kcal/mol).

The H-bond WC interactions in the AT pair within the geometries of M-AT complexes are systematically weakened in comparison with the isolated optimized AT pair ($\Delta E^{\text{MP2}} = 12.3$ kcal/mol). This weakening, which is larger for divalent ions amounts to about 4 kcal/mol, leading to AT pairing energy of ca 8 kcal/mol. A similar weakening of H-bonds was expected to occur in the M-GC complexes (cf. Table 6). However, it was found that ΔE^{MP2} of H-bonds in the GC pair within the M-GC complexes are a little stronger (with Ca^{2+} - Ba^{2+} exceptions) than those of the H-bonds in the isolated GC pair ($\Delta E^{\text{MP2}} = -26.3$ kcal/mol). It can be shown that these changes correspond to the geometry deformation under the metal coordination. From the energy decomposition it can be concluded that the H-bond strength of GC or AT pair, calculated as a pairwise interaction energy within the optimized M-GC or M-AT complexes, is influenced by cations only marginally. However, metal cations bound to the WC base pairs dramatically (directly or indirectly) change many characteristics of the base pairing. Actually, one should not consider pairwise energies of the G-C and A-T pairs but the MG-C and MA-T H-bonding energies and these values are systematically higher – up to 24 kcal/mol for A-T and 48 kcal/mol for G-C H-bonding in the presence of Zn^{2+} cation. It amounts to nearly two times enhancement of the original base pairing energy.

Similar conclusions are also reported in other studies on the metal-nucleobases interactions. A trimer base arrangement was explored using a similar computational model. The enhancement of G.GC(rH), G.GC(H), A.AT(rH), T.AT(H) and some other base interactions including reverse Hoogsteen pairs GG(rH) and AA(rH) was proved in the presence of Li^+ and Ca^{2+} cations (Šponer et al. 1997). A study on the strength of H-bonding of WC base pairing under metalation at various active sites of bases concluded that the N3(Adenine) site available in the minor groove has higher chances for platination, in comparison with the N7-site of the base (Zhang and Huang 2007).

The results of investigation on $\text{Cu}^+/\text{Cu}^{2+}$ interaction with AT and GC pairs were published by Sodupe's group (Noguera et al. 2004, 2007). The influence of metal cations coordinated to N7 of guanine on the intermolecular proton-transfer reaction in guanine-cytosine base pair was studied with the B3LYP density functional. Gas phase metal cation interaction stabilizes the ion pair structure derived from the N1-N3 single-proton-transfer reaction, the effects being more pronounced for the divalent cation than for the monovalent one. For Cu^{2+} -GC, the reaction is largely favored due to both electrostatic and oxidative effects. Hydration of the metal cation reverts this trend due to the screening of electrostatic effects.

In the section on the interaction of hydrated metal complexes, Šponer's study (Šponer et al. 1998a) was mentioned where in addition to purine nucleobases, the GC, AT, and AA base pairs were also considered. Since the metal-purine base interaction is not substantially influenced by the remote pyrimidine base we will

Table 6 Enhancement of Watson-Crick base pairing energy under the metalation. (in kcal/mol) (energies calculated at the MP2/6-31G** level)

Metal	E(M-A)	E(A-T) ^a	E(M-AT)	E(MA-T)	E(M-A-T)	E(M-G)	E(G-C) ^a	E(M-GC)	E(MG-C)	E(M-G-C)
Cu ⁺	-54.5	-12.0	-58.6	-16.1	-70.5	-79.9	-27.2	-86.1	-33.3	-113.3
Ag ⁺	-36.0	-12.0	-40.0	-16.0	-52.0	-64.2	-27.0	-69.0	-31.8	-96.1
Au ⁺	-55.0	-12.0	-59.1	-16.1	-71.1	-75.9	-27.0	-81.8	-32.9	-108.8
Zn ²⁺	-152.9	-10.8	-165.9	-23.8	-176.7	-237.2	-26.1	-259.4	-48.2	-285.4
Cd ²⁺	-116.6	-10.9	-129.3	-23.6	-140.2	-192.6	-25.9	-211.3	-44.6	-237.2
Hg ²⁺	-141.1	-10.9	-153.8	-23.7	-164.7	-207.9	-25.9	-228.0	-45.9	-253.9
Mg ²⁺	-107.9	-10.8	-120.7	-23.5	-131.5	-198.6	-25.9	-217.8	-45.1	-243.8
Ca ²⁺	-61.6	-11.1	-73.1	-22.6	-84.2	-133.8	-25.7	-146.9	-38.8	-172.6
Si ²⁺	-48.9	-11.2	-59.8	-22.1	-71.0					
Ba ²⁺	-51.4	-11.2	-62.3	-22.1	-73.5	-118.8	-25.6	-130.4	-37.2	-156.0

focus on the base pair enhancement and some changes observed in comparison with the study on bare metal cations complexes (Burda et al. 1997). The strength of the guanine-cytosine Watson-Crick base pair is enhanced by ca. 20–30 % due to the coordination of the hydrated cation, while in the case of bare cations this enhancement was about 60–90 %.

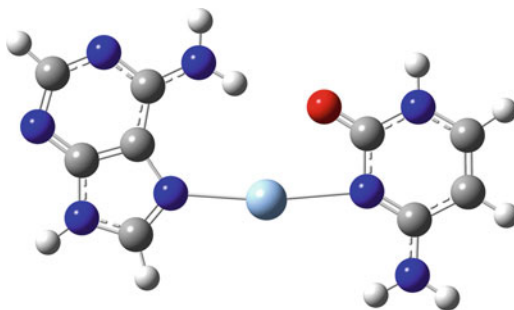
In the study of Oliva et al. (Oliva and Cavallo 2009), an impact of binding of hydrated Mg^{2+} , Mn^{2+} , Co^{2+} to the N7 position of guanine on Watson-Crick and reverse Watson-Crick base-pairs was investigated at the MP2 level. The combined model, where implicit solvent (COSMO) is considered together with an explicit first hydration shell of the cation shows that the stabilization effect of metalation on Watson-Crick geometry is rather weak – from 0.1 kcal/mol for Mg^{2+} to 0.8 kcal/mol for Co^{2+} cation. In the case of reverse Watson-Crick base pair, the stronger stabilization was determined in the range from 1.7 kcal/mol for Mg^{2+} and Mn^{2+} to 2.2 kcal/mol for Co^{2+} , which is in accord with our previous study (Šponer et al. 1998a). It was also shown that for the hexahydrated cations, the binding energies increase in the order $Mg^{2+} < Mn^{2+} < Co^{2+}$. The lowest stability of Mg^{2+} cation can be explained by missing back-donation from the aromatic ring of the guanine to empty d orbitals of metal. An impact of metalation of GC base-pair by hydrated Mg^{2+} , Ca^{2+} , Zn^{2+} , Cu^{2+} on vibrational spectra was studied using SIESTA code by Morari et al. (2012, 2014) clearly showing the effect of adduct formation.

An interesting comparison of Zn^{2+} and Mg^{2+} hydrated cations in Pu-Pu-Py triplexes was carried out (Šponer et al. 1998b) where a hydrated metal is coordinated to the N7 position of purine base attached to the Watson-Crick base pair. Using this model, metal assisted triplex stabilization was studied. It was shown that in both A.A and GG, the Hoogsteen pairing is strengthened under metalation. A substantially stronger enhancement of (MG).G pairing was revealed (19.8 and 20.4 kcal/mol for Mg^{2+} and Zn^{2+} cations, respectively), in comparison with similar adenine structures where no additional stabilization was practically detected (only about 2.7 kcal/mol for both metal cations). In another study (Šponer et al. 1999d), a more extended set of divalent metal cations (Mg^{2+} , Ca^{2+} , Sr^{2+} , Ba^{2+} , Zn^{2+} , Cd^{2+} , and Hg^{2+}) hydrated by five water molecules was explored with rAA and AT base pairs interacting with N7 site of adenine confirming the previously obtained results.

Rare tautomers of NA bases are important from the point of view of the global mutagenicity of DNA. Their formation by proton transfer between N1(G) and N3(C) in canonical Watson-Crick base pair was studied by Céron-Carrasco et al. (Ceron-Carrasco and Jacquemin 2011), extended later using ONIOM (QM/QM) method (Ceron-Carrasco et al. 2012b) for metalated B-form d(G-C,Mg-G-C,G-C) trimer. These authors explored also cisplatin cross-linked ds-DNA 5'-d(AG*G*C)-3' (Ceron-Carrasco et al. 2012a) tetramer. In the latter case, the calculations of PT energy profile were performed using ONIOM approach at M06-2X/6-311++G(d,p):AMBER(UFF) level with both crosslinked G*C base pairs and cisplatin in the QM part. LANL2DZ basis set was used for a description of Pt atom. The formation of the rare tautomer in crosslinked G*C pairs through the PT is characterized by an increase of the electronic energy by 9–12 kcal/mol if the considered DNA arrangement is close to the experimental NMR structure

Fig. 13

Cytosine- M^+ -Adenine complexes, $M = (\text{Cu}, \text{Ag}, \text{Au})$



(Wu et al. 2007). The corresponding activation barriers are only by 0.2–0.5 kcal/mol higher than reaction energy in the given case. In the former case, the hydrated Mg^{2+} cation is coordinated to nitrogen N7 of guanine. The reaction energy of 3.2 kcal/mol and the activation barrier of 8.0 kcal/mol for the considered N1(G) to N3(C) proton transfer were determined at M06-2X/6-311++G(d,p):B97-D/6-31+G(d)/PCM level. The inner part was consisted of GC base pair and hydrated Mg^{2+} cation.

An interesting idea is related to a replacement of proton in H-bonding by metal. Nowadays, these metal mediated NA base structures are highly investigated for the purpose of obtaining novel materials which might be employed as, for example, electric, fluorescent, or luminescent sensors. In this way, the noncanonical A-C base pair was examined with the coinage metal cations (Cu^+ , Ag^+ , and Au^+) used as a bridge between both bases (Šponer et al. 1999e), cf. Fig. 13. A possibility of additional water coordination to the metal was also considered. In the original paper, it is concluded that these metal bridged complexes are substantially more stable than original (protonated) base pair. The water coordination does not influence the strength of the metal bridge substantially; nevertheless, its presence has some impact on the geometry of the complexes. Comparing bonding properties of all three metal cations, it was found that Ag cation coordinates relatively weakly, which is in good accord with the previous results (Burda et al. 1996, 1997). Slightly shorter bond length $d(\text{M-N7-Adenine})$ than $d(\text{M-N3-Cytosine})$ contradicts the estimated bonding energies since ΔE^{M-A} is smaller than ΔE^{M-C} . This can be explained by two facts: (a) the electrostatic contribution to the metal coordination is much smaller in adenine case (see the size and orientation of the dipole moments of both bases in Fig. 2) and (b) the metal-cytosine interaction cannot be considered purely of the M-N3 character since the M-O2 contribution plays a nonnegligible role too.

Schreiber et al. (Schreiber and Gonzalez 2007a) explored the Ag(I) adducts with DNA base and the influence of Ag(I) coordination for adenine-cytosine mispairing. Their calculations show that in gas phase the canonical form of cytosine is stabilized upon metalation, whereas the lowest energy structure of Ag-adenine corresponds to the imino tautomer. The most stable metalated adenine-cytosine mispair is formed from the canonical cytosine and adenine tautomers. Other types of A-C pairs (e.g., reverse Wobble) were found much less stable. The same authors also performed

an interesting study dealing with the role of Ag(I) cation on electronic spectra of the A-C pairs using very accurate MS-CASPT2 approach (Schreiber and Gonzalez 2007b).

Stability of silver-mediated cytosine:cytosine or cytosine:adenine mismatches was also studied in implicit solvent in more recent studies (Megger et al. 2011a; Fortino et al. 2015). In the case of C:C mismatch, the Ag⁺ cation is coordinated to both N3 nitrogen atoms and the trans arrangement of cytosines is favorable since it enables formation of the weak hydrogen bond O2...H-N4 (2.13 Å). The presence of this hydrogen bond simultaneously causes that N3-Ag-N3 angle is about 160°. Counterpoise corrected binding energy (BE) for described base pair is 34 kcal/mol at B3LYP-D3/6-311++G(2df,2pd)/SDD/COSMO level while the gas phase BE is 43 kcal/mol according to ref. (Fortino et al. 2015).

In contrast to rather planar geometry of C-Ag-C structure, one of the pairs in DNA dimer d(C-Ag-C)₂ is characterized by rather large propeller twist of 78°. The Ag...Ag distance (3.9 Å) in the optimized dimer structure is comparable with Hg...Hg distance of 3.8–3.9 Å in (T-Hg-T)₂ dimer (Benda et al. 2011). BE of dimer (73 kcal/mol) shows that this energy is practically additive.

Considering C-Ag-A base pair, the Hoogsteen arrangement, where Ag⁺ ion is coordinated to N7(A) and N3(C) nitrogen atoms, is more favorable than the Watson-Crick type. The corresponding energy difference is 5.9 and 2.3 kcal/mol for C-Ag-A monomer and (C-Ag-A)₂ dimer, respectively.

It was found that if concentration of Ag⁺ ions is sufficient, stabilization through two Ag⁺ ions becomes preferable in the case of thymine:1,3-dideazaadenine (D) base pair (Megger et al. 2011b). In addition to usual N7(D)-Ag-N3(T) linkage, the second Ag⁺ ion is coordinated to N6(D) nitrogen and O4(T) oxygen. The optimized mutual distance between Ag atoms is 2.88 Å. The energy difference of 24 kcal/mol between the base pair with two Ag linkage atoms and the base pair with one linkage N7(D)-Ag-N3(T) plus monohydrated Ag⁺ ion was determined at the ZORA-BLYP-D/TZ2P/COSMO level.

Interestingly, possibility of the so-called M-DNA crosslink stabilization of GC base pairs by divalent zinc has been investigated by Fuentes-Cabrera et al. (2007). Further, stabilizations of the base pair of Watson-Crick (AT, AU, GC), Hoogsteen (AT, AU), and Wobble (GU) type by Ib transition metals – Cu⁺, Ag⁺, and Au⁺ cations – were explored (Marino et al. 2012). Additionally, 1-deazaadenine instead adenine was considered in the case of the Hoogsteen type base pairs. The metal cation is coordinated to N7 nitrogen of purine bases and N3 nitrogen of pyrimidine bases in these structures except for N1(G)-M-O2(U) linkage in Wobble base pair. In order to maintain one strong hydrogen bond in Watson-Crick base pairs after incorporation of the metal ion, angle N3-M-N7 is not exactly 180°, but it is in the range from 150° to 180°. The coordination of the metal cation causes an increase of binding energies (determined at the B3LYP-D3 level in gas phase) about 20–35 kcal/mol in comparison with hydrogen-bonded base pairs. Independently on the base pairs and type of binding, stabilization by Cu⁺ and Au⁺ cations is comparable, while in the case of Ag⁺ ion, the binding energy is about 10–20 kcal/mol lower in agreement with (Šponer et al. 1999e).

A detailed study of formation of (T)N3-Hg-N3(T) linkage was performed by Šebera et al. (2013b) The ds-DNA was modeled as trimer G-C, T-T, T-Hg^{II}-T, and the ONIOM (B3LYP/6-31G(d)/SDD:BP86/LANL2DZ with implicit water solvent IEF-PCM/UFF) was used. The initial step represents proton transfer from N3 to hydroxo ligand of hydrated Hg²⁺ ion and subsequent formation of the first Hg-N3(T1) bond. This step is slightly endergonic ($\Delta G_r = 1.4$ kcal/mol) and is related to the very low barrier of 2.9 kcal/mol. The reaction proceeds with the second proton transfer from N3(T2) to O2(T2) (tautomerization) or to another hydroxo ligand of Hg²⁺ ion if it is present. The activation barrier corresponding to tautomerization is 16.5 kcal/mol high, while for proton transfer to the hydroxo ligand the barrier is about 3 kcal/mol lower. The total reaction Gibb's energy is -9.5 kcal/mol and the bond length of both Hg-N3 bonds is 2.15 Å, when the deprotonated product at O2 position was considered.

In order to show changes in electronic properties after formation of Hg^{II} linkage between thymine: thymine mismatch, a charge transfer from donor Ap to acceptor G in ds-DNA tetramer sequence Ap-T-A-G was studied in the low temperature region (Kratochvílová et al. 2014). Electronic coupling $|J_{Ap,G}|$ between Ap and G bases determined using relativistic ZORA at the B3LYP/TZ2P/COSMO level shows that the value of $|J_{Ap,G}|$ increases about 4–5 meV after formation of the Hg^{II} linkage compared to T-T mismatch structure. When the structure of the metal-mediated mismatch is preserved and just Hg atom is removed, the value of $|J_{Ap,G}|$ decreases only by 1 meV. Hence, it was concluded that the structure of the thymine-thymine mismatch is more important for the charge transfer than the presence of Hg²⁺ cation.

Also differences in Raman spectra related to formation of T-Hg^{II}-T linkage were investigated both experimentally and theoretically (B3LYP-D2/6-31+G**/COSMO) (Benda et al. 2012). Spectral shifts are fairly small for both experimental and calculated spectra. However, calculated spectra provide usually slightly overestimated values compared to the experiment. An exception represents the vibrational mode corresponding to C = O stretching (1664 cm^{-1} in TpT), where the red shift of -77 cm^{-1} was observed.

The stability of dimer Tp-Hg-T · T-Hg-Tp with two Hg²⁺ ions was verified for RNA analogue U-Hg-U · U-Hg-U dimer at the spin-component-scaled MP2 level (Benda et al. 2011). The equilibrated distance $d(\text{Hg}\cdots\text{Hg})$ was determined to 3.8–3.9 Å. Simultaneously, a metallophilic attraction (dominance of a dispersion interaction over repulsive electrostatic one) between Hg atoms was estimated in the range from -1.1 to -0.7 kcal/mol according to a mutual arrangement of the metal-mediated base pairs. The trimeric systems (T-Hg-T)₃ and (U-Hg-U)₃ were studied by Marino et al. (Marino 2014) at B3LYP-D3(M06-L)/TVP level in the gas phase. The vertical binding energies between layers were 29.8 and 30.8 kcal/mol for (T-Hg-T)₃ and (U-Hg-U)₃, respectively. The binding energies of Hg²⁺ ion in T-Hg-T and U-Hg-U monomers was also determined and their values are 200 and 186 kcal/mol.

Study focused on Ag-mediated Hoogsteen-type base pairs containing cytosine and modified 1,3-dideazapurine and 1-deazapurine at exocyclic O6 position with hydroxo, methoxo, fluoro, chloro, and bromo group was published (Megger et al.

2011a). It was concluded that BE were always lower in comparison with the mediated C-Ag-G base pair except the base pair containing 1,3-dideaza-6-hydroxypurine, which is characterized by the same energy.

Interactions of Hydrated Cations with Nucleotides

Šponer et al. (2000b) studied the coordination of hydrated zinc and magnesium group divalent cations to the N7 position of purine nucleotides. They showed that the sugar-phosphate backbone provides a significant screening of the metal charge, while the backbone geometry is affected by the cation. Polarized water molecules of the cation hydration shell form very strong hydrogen bond bridges between the cation and the anionic oxygen atoms of the phosphate group. Weaker hydrogen bonds are formed between the cation hydration shell and the exocyclic purine X6 atoms. The cation binding to N7 of adenosine monophosphate forces the adenine amino group to adopt nonplanar conformation. Its nitrogen atom serves as an H-acceptor for a water molecule from the cation hydration shell. Cation binding to N7 does not lead to any major changes in the geometry of the base pairing. However, the stability of the base pairing can be increased by polarization of the purine base by the cation and by long-range electrostatic attraction between the hydrated cation and the other nucleobase. The stability of guanine-cytosine Watson-Crick base pairing is enhanced by the polarization mechanism while the stability of the adenine-thymine Watson-Crick base pair is amplified by the electrostatic effects as shown in the case of base pairs model discussed previously (Burda et al. 1997). Also, the guanine-guanine reverse-Hoogsteen base pairing is stabilized by both contributions while the adenine-adenine reverse-Hoogsteen system is not influenced by the cation. Binding of a cation to the N7 of guanine promotes transfer of its H1 proton to the N3 acceptor site of cytosine. However, the negatively charged backbone exerts a significant screening effect on this potentially mutagenic process, and the probability of such a proton transfer in DNA should be only moderately enhanced by a cation binding.

A comprehensive study dealing with a coordination of hydrated cations Zn(II) and Mg(II) to guanosine 5' monophosphate was performed by Gresh (Gresh et al. 2003) in order to obtain parameters for polarizable molecular mechanics for metal-DNA and metal-RNA simulations.

An important role of the position of Mg^{2+} ion in the active site of hepatitis delta virus (HDV) ribozyme during its self-cleaving process was described by Mlynsky et al. (2015) using the ONIOM method implemented in Gaussian 09. The reaction starts with deprotonation of O2'-H(U-1) group in the active site of HDV, which is followed by the nucleophilic attack of O2'-(U-1) on phosphorus of the scissile phosphodiester. As a consequence, the P-O5'(G1) bond is broken and O5' position is saturated by the proton from C75 (cf. Fig. 14). The protonated state at the N3 position of C75H⁺, which occurs with only low probability, was considered as a reactant (estimated pKa for this protonation is 8.8–11.2) together with $[Mg(H_2O)_nOH]^+$ used as Brønsted base. The experimentally determined pKa constant for hydrated Mg^{2+} ion is 11.4. On the basis of estimated pKa constants, the Gibbs energy correction for considered protonation state of reactant was 8.3 kcal/mol. In ONIOM calculations (BLYP/6-31 + G(d,p):AMBER(ff99bsc0χOL3), reoptimized with hybrid

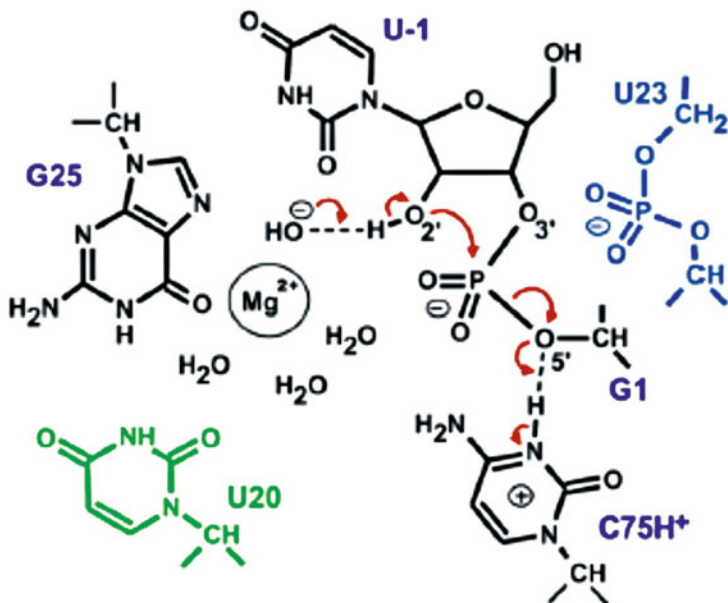


Fig. 14 Active site of HDV ribozyme with coordinated Mg^{2+} cation. Electron transfers during the self-cleavage process are depicted by *arrows*

MPW1K functional), the authors considered eight pathways with the different position of Mg^{2+} ion coordination in reactants based on previous MD simulations. The QM part contained hydrated Mg^{2+} cation, uracil U-1, G1 guanosine, cytosine C75, and guanine G25. The resulting Gibbs energy barriers are in the range from 14.2 to 28.8 kcal/mol and the lowest barrier (14.2 kcal/mol) corresponds to the initial coordination of Mg^{2+} cation to O2(U-1) and O2'(U-1).

Enhancement of Base Pairing by Pt Complexes

Several studies related to platinum metal interactions with base pairs and the influence of metal complexes on the strength of pairing should be mentioned.

Molecular structures of several Pt complexes with the Watson-Crick AT and GC base pairs were optimized using the B3LYP method. The interaction energies were analyzed using B3LYP and MP2 approaches (Burda et al. 2001). Platination causes some distortion in the H-bond arrangement of the base pairs. The pyrimidine bases rotate around their centre of mass under the influence of the charged Pt entities. This effect is quite general and was already discussed above (Burda et al. 1997; Šponer et al. 1998a, 1999b). The metal-binding affects the strength of individual H-bonds involved in the base pairing. It was concluded that ligands attached to the Pt(II) cation form rather strong intramolecular H-bonds with the X6 exocyclic site

of purine bases. The adenine amino group adopts a pyramidal-rotated geometry and its nitrogen serves as H-bond acceptor for the ammine ligands of cisplatin.

The Pt-binding has a comparable effect on the base pairing stability as binding of hydrated metals of IIa and IIb groups. In the electroneutral form, the Pt-adducts does not practically influence the base pair stability. Charged Pt adducts substantially strengthen the stability of GC base pair via polarization effects. The influence of the +2 charged Pt-adduct is even larger compared with hydrated metals of IIa and IIb groups. No such polarization effects have been revealed for the AT base pair. However, gas-phase stability of this pair is effectively enhanced by long-range electrostatic interaction between the charged metal group and thymine.

The dependence of the stabilization and Pt-N7 binding energies on the total charge of the complex is displayed in Fig. 15a. Here, the role of electrostatic contributions is clearly demonstrated for the both kinds of energies. Interestingly, in electroneutral complexes, the Pt-N7 binding energies are similar in guanine and adenine structures (≈ 50 kcal/mol corresponding to “pure” coordination covalent contribution). In the Fig. 15b, the base pairing energies are drawn showing that the geometry deformations do not basically influence these values. Nevertheless, the interaction of metalated purine base with pyrimidine base is substantially strengthened, especially in the case of charged complexes (e.g., in complexes of hydrated cisplatin). It is important to point out that the effect of base pair stabilization enhancement due to cation binding has been confirmed experimentally by Sigel and Lippert (1999).

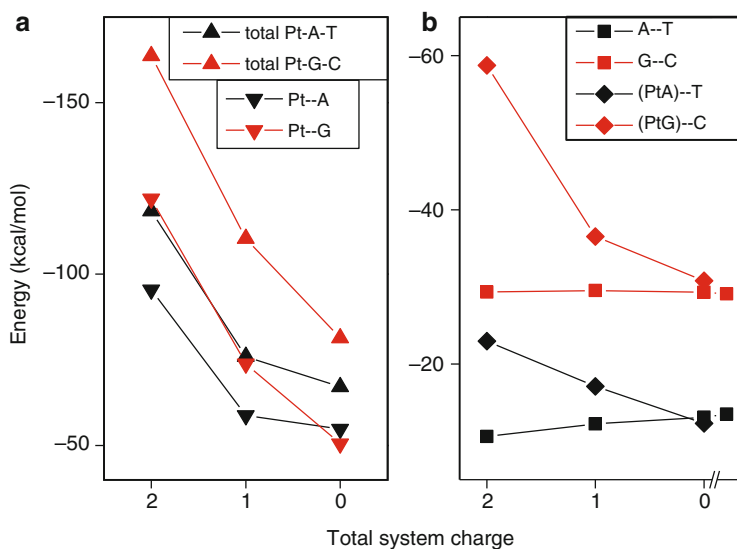


Fig. 15 (a) Stabilization and metalation Pt-N7 energies. (b) Enhancement of the Watson-Crick pairing energies for AT and GC base pairs

Table 7 Gas phase stabilization energies of the Pt-cross-linked structures (in kcal/mol)

	$\Delta E^{\text{stab}}(\text{MP2})$
Pt-a ₂ A ₂	491.5
Pt-a ₂ AG	514.6
Pt-a ₂ G ₂	528.3

Zilberberg et al. reported an influence of chelated cisplatin complex with guanine(O6,N7) on Watson-Crick base pairing (Zilberberg et al. 1997). In such chelate structures, more distinct perturbation of base pairing was revealed. However, such chelate binding pattern is unlikely to be relevant to experimental conditions.

Metal Cross-Linked DNA Structures

Platinum(II) Complexes

Cisplatin bridges between two consequent bases (1,2-GpG) are believed to be the key structure for triggering the apoptotic process. Recently, several studies on these cross-linked structures have been published. The properties of Pt-bridges were explored (Burda and Leszczynski 2003; Zeizinger et al. 2004) showing that relatively strong Pt-N7 coordination is formed. The process of the aqua ligand replacement by a nucleobase is mildly exoergic in both steps forming: (a) a monofunctional adduct and, consequently, (b) the cross-linked structure. Stabilization energies of the Pt-GG, Pt-GA, and Pt-AA bridges are collected in Table 7. The relative amount of these values correlates well with the abundance of individual structures in real samples, assuming the gas-phase calculations require some additional rescaling in correspondence to the reduced electrostatic interactions in solvent.

An influence of the sugar phosphate backbone on the strength of Pt-bridge was also examined (Zeizinger et al. 2004). An additional stabilization of the Pt cross-linked structures appeared as a result of the interaction between negatively charged phosphate group and the Pt cation.

One of the first papers studying transition states of the replacement aqua ligand by nucleobase was published by Chval (Chval and Sip 2003). His model was based on the gas-phase calculations and the estimated activation barrier is too low in comparison with the experimental value. This situation was improved in studies of Raber (Raber et al. 2005) and Baik (Baik et al. 2003). They have taken into account the hydration effects employing implicit solvent model. Especially, Raber's results are in fairly good accord with experimental data. Activation barriers of mono aqua and diaqua Pt(II) complexes are summarized in Table 8 for both guanine and adenine replacement of the first and second leaving (aqua) ligand. In Table 9, the activation barriers for second and third generation Pt(II) drugs – oxaliplatin, nedaplatin, and carboplatin – are displayed (Alberto et al. 2011), where higher barriers also occur in the case of adenine. According to the calculation of oxaliplatin, the second reaction step should be visibly more energy demanding.

Table 8 Activation barriers of the substitution reaction for the aqua ligand replacement by purine base (in kcal/mol)

1st step	Chloro-aqua	Diaqua
Guanine	21.4 ^a	19.5 ^a
	25.6 ^b	21.8 ^b
	14 ^c	17.9 ^c
Adenine		18.3 ^d
	24.0 ^a	24.8 ^a
	37.6 ^b	34.5 ^b
	14.5 ^c	14.5 ^c
2nd step		
GG hh		22.5 ^a
		23.4 ^c
GA hh		28.6 ^a

^aRef. (Raber et al. 2005)^bRef. (Baik et al. 2003)^cRef. (Chval and Sip 2003), experimental data^dRef. (Arpalahti and Lippert 1990)^eRef. (Bancroft et al. 1990)**Table 9** Activation barriers for the substitution of the aqua ligand by N7-guanine or N7-adenine in the case of oxaliplatin, nedaplatin, and carboplatin (Alberto et al. 2011) (in kcal/mol)

	Guanine	Adenine
	Monoaqua	
Oxaliplatin	10.2	17.1
Nedaplatin	24.9	26.0
Carboplatin	13.2	25.1
	Diaqua	
Oxaliplatin	21.4	24.1

A comprehensive study on formation of Pt-N7 cross-links between fully hydrated oxaliplatin – cis-[Pt(DACH)(H₂O)₂]²⁺ (DACH = cyclohexane-1R,2R-diamine) – and double-stranded pGpG·CpC dinucleotide was performed by Chval et al. (2013). Structures were optimized in the implicit solvent model at the RI-TPSS-D/MWB-60(f)/def2-SV(P)/COSMO level and the energy profile and interaction energies were subsequently obtained at the ω-B97XD/MWB-60(2f)/6-311 + G-(2d,2p)/IEFPCM/UFF and B3LYP-D/MWB-60(2f)/def2-TPZVPP/COSMO levels. The whole reaction mechanism is depicted in Fig. 16. In the study, both 5' → 3' and 3' → 5' binding directions to the DNA dimer were investigated. The reaction energies for the monoadduct formation are –3.3 kcal/mol and –1.8 kcal/mol, respectively. The subsequent cross-link formation from the monoadduct is connected with releasing of –1.0 and –8.5 kcal/mol, respectively. Formation of the first Pt-N7 bond represents the rate-limiting step and the binding to N7 (5'G) is slightly kinetically preferred over N7(3'G). The corresponding activation barrier is 21.2 kcal/mol. The following chelation reaction is more favorable in the 3' → 5' direction (activation barrier of 17.1 kcal/mol) in agreement with the experimental measurements. Formation of each Pt-N7 bond leads to an increase

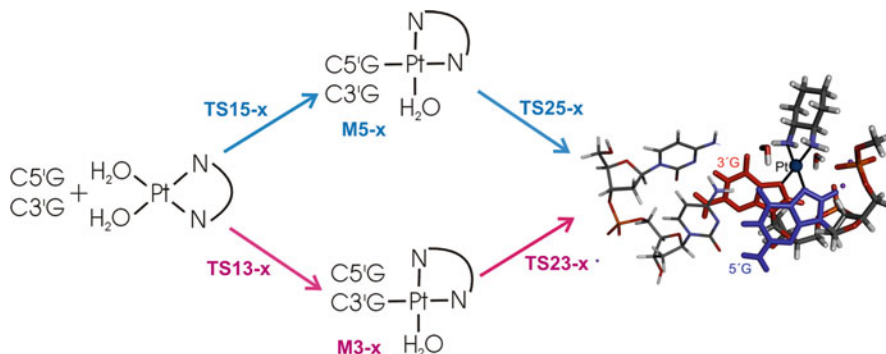


Fig. 16 Scheme of the reaction mechanism for formation of the cross-linked structure by fully hydrated oxaliplatin

of the stabilization between the ds(pGpG) and the Pt(DACH) fragments by about 15 kcal/mol in aqueous solution. Further, the solvent-phase pairing energies are enlarged by about 10 % with respect to the pGpG · CpC dinucleotide itself and do not visibly depend on the number of Pt-N7 bonds for most of the platinated structures. Simultaneously, hydrogen bond O6(G) ··· H4(C) is weakened as a consequence of interaction between amino group of Pt(DACH) and O6 (G). Nevertheless, the described weakening is compensated by strengthening of H1(G) ··· N3(C) and H2(G) ··· O2(C) H-bonds.

An arrangement of two adjacent guanosines (G), which are cross-linked by kiteplatin (cis-Pt(1,4-DACH)Cl₂), were studied by Margiotta et al. (2015) in the case of cis-Pt(1,4-DACH)d(GpG), cis-Pt(1,4-DACH)d(GGTTT), and cis-Pt(1,4-DACH)d(TGGT). An initial systematic conformation search at the LFMM (Ligand Field Molecular Mechanics) (Deeth 2001; Deeth et al. 2009) level for cis-Pt(1,4-DACH)d(GpG) was followed-up by reoptimization in gas and also aqueous phase using the semiempirical PM6-DH2 method (Korth et al. 2010; Řezáč et al. 2009). It was determined that the head to head arrangement of anticonformers of Gs is preferred over the head to tail arrangement by 6 kcal/mol (in solvent), which is in accord with experimental measurements.

Binding energies of cisplatin, oxaliplatin [(cyclohexane-1R,2R-diammine)-oxalato platinum(II)], and kiteplatin (Fig. 17) are mutually compared by Mutter et al. (2015). The strongest binding energy was determined for cisplatin if the interaction of Pt complexes with GpG · CpC dimer was considered at the DFT level (B97-D/TZVP). Binding energies of kiteplatin and oxaliplatin are comparable. This trend is independent on the fact whether the system is in the gas phase or in the water solution represented by COSMO solvent model. Nevertheless, when the GpG · CpC dimer is replaced by octamer (ds-CCTGGTCC) and QM/MM approach was used, the binding energies of cisplatin are the weakest, then follows kiteplatin and the strongest interaction shows oxaliplatin (cf. Table 10). In another study (Gkionis et al. 2013), the interactions of heptaplatin ([2-(1-Methylethyl)-1,3-dioxolane-4,5-dimethanamine-*N,N'*]-[propanedioato(2-)-*O,O'*]-platinum(II)) and lobaplatin

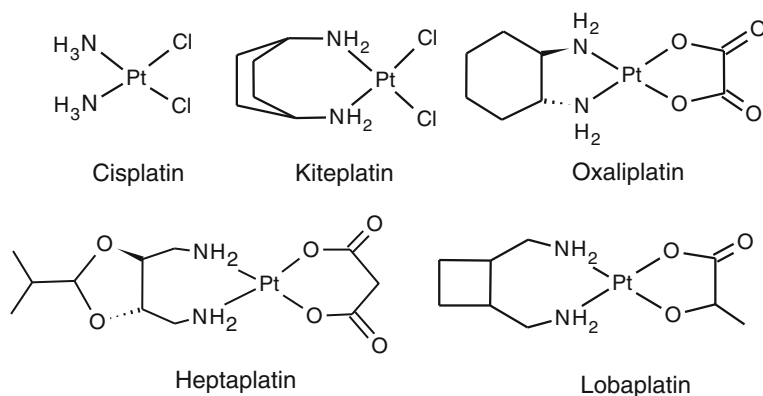


Fig. 17 Structure of the selected Pt(II) drugs

Table 10 Binding energies of cisplatin, oxaliplatin, and kiteplatin to DNA dimer GpG·CpC and octamer ds-CCTGGTCC (in kcal/mol) (Mutter et al. 2015)

	DFT BE ^{a)} Dimer	ONIOM BE ^{b)} Octamer
Cisplatin	−129.6	−245.4
Oxaliplatin	−122.6	−299.6
Kiteplatin	−122.0	−282.4

^aCounterpoise corrected binding energies determined at the B97-D/TZVP/COSMO

^bBinding energies determined the BH and H/SDD/6-31+G(d,p):AMBER/parm99

(1,2-diammino-methyl-cyclobutane-platinum(II)-lactate) (besides oxaliplatin and cisplatin) with the same octamer were investigated and it was determined that the binding energies are about 38 kcal/mol and 28 kcal/mol, resp., higher than for oxaliplatin. In the former paper, (Mutter et al. 2015) was also mentioned an interesting trend that when the DNA dimer, tetramer, octamer, and dodecamer is considered, the ONIOM binding energy of kiteplatin increases towards octamer, however it is (slightly) smaller for dodecamer. Basically it is a question why binding energy should increase so dramatically with increasing size of the system and it is pity that some deeper analysis (or energy decomposition) is not present.

Ruthenium(II) Complexes

An interesting continuation of the investigations of the “piano-stool” Ru(II) complexes with DNA base mentioned from section “Properties of Metal Adducts with Nucleobases” represents formation of Ru(II) cross-linked structures studied using hybrid ONIOM and QM/MM Umbrella sampling-MD by Futera et al. (Futera and Burda 2014; Futera et al. 2012). An interaction of piano stool complexes with DNA is to some extent similar to cisplatin. The considered representative of piano stool complexes, $[\text{Ru}^{\text{II}}(\eta^6\text{-benzene})(\text{en})\text{Cl}]^+$, must be initially activated by hydration reaction. Subsequently, the Ru(II) complex/DNA monoadduct is formed with the

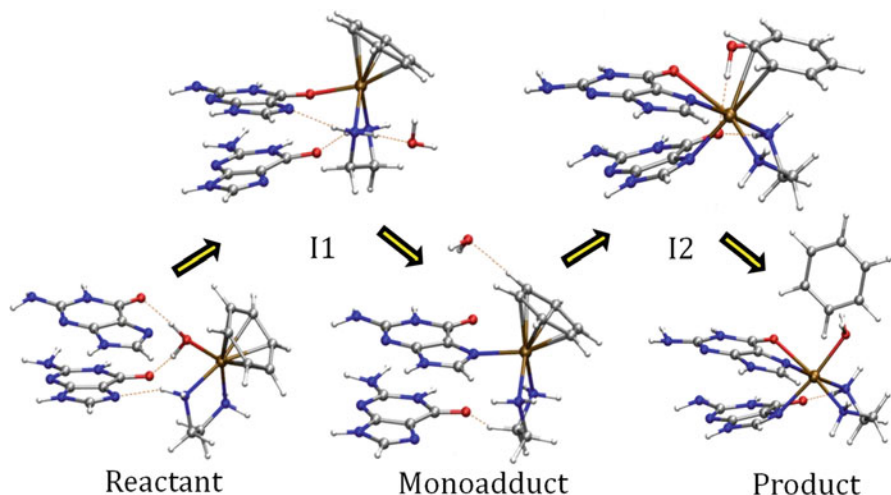


Fig. 18 Formation of Ru(II) cross-link in the presence of hydrated $[\text{Ru}^{\text{II}}(\eta^6\text{-benzene})(\text{en})\text{Cl}]^+$ complex

bond between Ru and N7 nitrogen of one of the adjacent guanines G^* . Formation of the Ru(II) cross-link is completed by binding of the complex to N7 position of the second guanine G^* (cf. Fig. 18).

The free energy profile in QM/MM MD approach was built using the umbrella sampling approach together with weighted histogram analysis (WHAM) and umbrella integration (UI) methods at the B97D/LANL2DZ:AMBER(parm03) level. Additionally, the FEP (free energy perturbation) corrections determined at B3LYP/SDD/6-31++G(d,p) level were added to increase the accuracy of the model. The ds-DNA was modeled by decamer structure with 5'-AATGG1*G2*ACCT-3' sequence. The QM core consists of the piano-stool complex and one explicit water molecule in hydration reaction and of hydrated Ru(II) complex and both adjacent guanines G^* in the remaining reactions.

For hydration reaction with smaller quantum core, the B3LYP/SDD/6-31G level was used. The Gibb's free reaction energy corresponding to this activation process is 2 kcal/mol and the activation barrier of 19.5 kcal/mol is in good agreement with experimental value (21.1 kcal/mol) (Wang et al. 2005). Formation of the monoadduct and subsequent cross-linked structure leads to decrease of total ΔG energy of the system by 3.4 kcal/mol and 9.5 kcal/mol, respectively. The most stable form of the monoadduct, i.e., with Ru-N7(G1^*) coordination, is not formed directly; the substitution goes through a structure where Ru-O6(G1^*) is formed first as an reaction intermediate. Similarly to the study (Chval et al. 2013) mentioned above, coordination of the Ru complex to guanine G represents the rate-determining step characterized with the activation barrier of 21.1 kcal/mol in the two step mechanism. Formation of the (G1^*)N7-Ru-N7(G2^*) cross-linked chelate complex is connected with release of the arene ligand, which is initially η^6 -coordinated. In the products, Ru-N7(G2^*) bond to the 3'-end guanine is substantially weaker

which is documented by bonding energies of -47 kcal/mol and -35 kcal/mol for Ru-N7(G1*) and Ru-N7(G2*), respectively. The previous ONIOM calculations at the B3LYP/6-31++G(d,p):AMBER(ff96) level (Futera et al. 2012), where the ds-DNA hexamer 5'-GCG*G*GC-3' model was used, predict higher activation barriers by about 4 kcal/mol for monoadduct formation in the two-step mechanism and by about 2 kcal/mol for direct monoadduct formation.

Ions in DNA Quadruplexes

The infrared (IR) and vibrational circular dichroism (VCD) spectra of guanosine-5'-hydrazide have been measured and analyzed on the basis of ab initio modeling (Setnicka et al. 2008). The B3LYP/6-31G(d,p) calculations predict that guanine, forming a clear solution in deuterated DMSO, is present in monomeric form in this solvent, whereas strong gelation in a phosphate buffer is due to the formation of a guanine-quartet structure. Here, the four bases are linked by hydrogen-bonded guanine moieties and stabilized by an alkali metal cation. The DFT prediction of the IR and VCD spectra are based on the nearly planar quartet structure, which is slightly distorted from the C_{4h} symmetry. The guanine bases interact via four Hoogsteen-type hydrogen bonds and a sodium cation is positioned in the middle of the guanine quartet. The obtained results are in very good agreement with the experimental spectra, indicating that calculated structure is highly probable in the gel state.

The guanine quartets were examined also by Gu and Leszczynski (2000). The normal four-stranded Hoogsteen-bonded G-quartet structures were optimized in the gas-phase with a monovalent cations obtaining the stability order $\text{Li}^+ > \text{Na}^+ > \text{K}^+$. However, after the correction on solvent effects, the stability sequence of the monovalent cation-guanine-tetrad complexes follows the opposite trend $\text{K}^+ > \text{Na}^+ > \text{Li}^+$. The preferential binding of potassium over sodium and lithium in water solutions reproduces the experimental ion selectivity of the guanine quadruplex. Moreover, the weak stabilization energy of the K^+ -G-quartet in the coplanar form corresponds with the fact that the potassium cation tends to locate between two successive quartets. These results are in accord with the study of Hud et al. (1998) on the ion selectivity of the guanine quartets in water solutions, which are governed by the relative free energies of hydration. The experimental data on the cation-oxygen distances in the sodium ion complex are 2.34 ± 0.02 Å. This value matches the HF value of 2.33 Å. The slightly shorter Na^+ -O6 distance were predicted by the DFT approach together with a significant shortening of the hydrogen bonds suggesting an overestimation of the H-bonding in the guanine quartets at the DFT level. A very comprehensive study on metal-quartet interactions has been published also by Meyer et al. (Meyer et al. 2001). It is to be noted, however, that, in principle, accurate studies of quadruplex-ion interactions would require inclusion of the whole solvated quadruplex fragment, due to the unique balance of molecular interactions in such a noncanonical structure (Šponer and Špacková 2007).

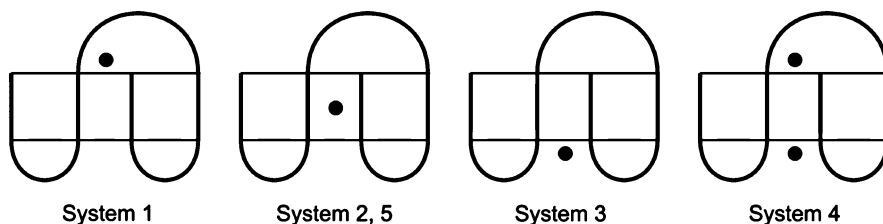


Fig. 19 An initial localization of ion(s) in 15-TBA in five studied systems

Apart from papers concerning guanine stems, duplexes consisting of all NA and isoguanine (iG) tetrads and iG pentads were investigated by Meyer et al. (Meyer et al. 2005, 2007; Meyer and Sühnel 2008) where K^+ or Na^+ ion was used for stabilization of these structures.

Seven distinct folds of two-quartet G-DNA stem which is stabilized by monovalent K^+ ion were studied using the large-scale QM calculation (Šponer et al. 2013). In contrast to the previous studies by Meyer (Meyer and Sühnel 2008), dispersion is included in the model. Optimization of the structures was performed at the RI-TPSS-D3/TZVP/COSMO level and results were compared to few MD approaches. It was shown that the relative stability of 5'-anti-anti-3' GpG steps (AA) in G-stems is underestimated in comparison with 5'-syn-anti-3' steps (SA) at the MM level. The final reevaluated stability order (due to the MM approach) gives that the AA arrangement is the most stable followed by SA step lying by 1.2 kcal/mol higher. An exception represents the terminal SA step at the 5'-end, which is by 3.2 kcal/mol more stable than the AA arrangement.

A localization of an ion in the quadruplex stem was investigated at the QM/MM MD level by Reshetnikov et al. (2011). As the model structure, the thrombin-binding aptamer (structure 15-TBA in PDB database) was chosen. The QM part was consisted of all nucleic acid bases (except T7) and K^+ or Ba^{2+} ions and was considered at the DFT level using the PW91 functional. The initial position (upper, central, lower) of ion(s) in five studied systems is depicted in Fig. 19. In systems 1–4, K^+ ion was present, while Ba^{2+} ion was considered in system 5. It was concluded that only the central binding site is stable, since the ion in the systems 1 and 3 moved to the center of the stem. Further in system 4, the lower ion was released and the upper one moved to the center.

The study of Gkionis et al. (2014) explored an ion binding in guanine stems at the DFT-D3 and the MM level (Cornell force field (Cornell et al. 1995)). Series of PES scans of binding of one ion (Li^+ , Na^+ , and K^+) to two- and three-quartet stems (without the sugar-phosphate backbone) and subsequent PES scans for binding of the second cation were performed for this purpose. In the single-ion case, the difference between the DFT-D3 and MM approach becomes acceptable, when the implicit solvent (COSMO in QM calculations and MM-PBSA (Poisson-Boltzmann surface area)) is concerned as the model. However, in the gas phase, a more attractive potential inside the channel for single-ion case was obtained at the MM level. A different situation occurs in a multi-ion case. The problem dwells in the missing polarization in the MM calculations. Symmetry adapted perturbation

theory (SAPT0) energy decomposition showed substantially weaker electrostatic repulsion between ions at QM level, where the ions can be polarized. Same results followed from the QTAIM analysis. Authors concluded that system with more than one cation cannot be even described correctly with fixed charged model, where the reduced charge on ions in MM simulation is considered.

Noncovalent interactions in two or three stacked guanine or xanthine quartets containing one or two univalent cations (Na^+ , K^+) were also investigated at DFT-D3 level by Yurenko et al. (2014). The study shows a structural and energetic compatibility between guanine and xanthine quadruplexes. The hydrogen bonding, ion coordination, and base stacking contributions to the total internal stabilization are compared. The determined contribution of hydrogen bonding is practically independent on the size of the system, and it approximately corresponds to 50% of the total noncovalent interaction. On the contrary, electrostatic repulsion between cations in trimer systems causes a slight decrease of a share of ion coordination on the total stability, which represents 34–41% of interaction energy for the duplex. A higher influence of K^+ ion on weakening of hydrogen bonds in the duplex compared to Na^+ cation is caused by stronger interaction with O6 atom given by its larger ionic radius is also mentioned. In general, the absolute values of the energy of ion binding to the considered quadruplex structures are higher for Na^+ than for K^+ . An Energy decomposition analysis further shows that the orbital contribution to ion binding is not negligible in comparison with the electrostatic term. Hence, it was concluded that ions interaction in the stem channel cannot be described purely by an electrostatic model.

A similar topic was examined also by Ida (Ida and Wu 2008). In this study, molecular dynamics was employed exploring G-quadruplex stabilized by Na and Rb cations, which were found to be tightly bound to the quadruplex structure. Moreover, in d(G(4)T(4)G(4)) sequences, the Na^+ ions are found to be located in the diagonal T-4 loop region of the G-quadruplex, which is formed by two strands of d(G4T4G4) sequence. The authors proposed that the loop Na^+ ion is located above the terminal G-quartet, coordinating to four guanine O6 atoms from the terminal G-quartet, one O2 atom from a loop thymine base, and one water molecule. The Na^+ coordination was also supported by quantum chemical calculations on ^{23}Na chemical shifts.

Conclusions

In this review, the results of some of the recent investigations on metal interactions with nucleobases, base pairs, and larger models (including base stacking or oligomer sequences) are summarized.

The studies dealing with metal cations (in naked, hydrated, and ligated forms) provide various details on their interactions with nucleobases, however such models are in many cases oversimplified. Despite the fact that some sites on the isolated nucleobase (especially in adenine) exhibits higher affinity to metal cations, in DNA helix not all of them are available for interactions since they are involved in H-bonding (purine N1 site) or in the glycosidic bond (N9 site).

All forms of metal cations enhance the Watson-Crick base pairing interaction if their positive charge is not fully compensated. The different mechanism for the AT and GC adducts was revealed. While in AT the direct electrostatic link between remote thymine (negatively charged O4) and metal cation exists, in the GC pair the nonadditive three-body term is important since the positively charged $\text{NH}_2(\text{C})$ exogroup is in the proximity of the metal cation.

Clearly, metalation at the N7 position leads into many new, exceptional properties of the studied systems that are dependent on characteristics of involved metals. Some of them are discussed in details in various parts of this text.

Acknowledgments The study was supported by projects Grant Agency of the Czech Republic (GAČR) No. 16-06240S. The authors thank the Meta-Centers in Prague (Charles University and Czech Technical University), Brno (Masaryk University), Pilsen (University of West Bohemia) for the generous support of the computational resources.

Bibliography

- Ai, H. Q., Yang, A. B., & Li, Y. G. (2008). Theoretical study on the interactions between Zn^{2+} and adenine isomers in aqueous phase. *Acta Physico-Chimica Sinica*, 24(6), 1047–1052.
- Alberto, M. E., Butera, V., & Russo, N. (2011). Which one among the Pt-containing anticancer drugs more easily forms monoadducts with G and A DNA bases? A comparative study among oxaliplatin, nedaplatin, and carboplatin. *Inorganic Chemistry*, 50(15), 6965–6971. doi:10.1021/ic200148n.
- Allen, R. N., Shukla, M. K., Burda, J. V., & Leszczynski, J. (2006). Theoretical study of interaction of urate with Li^+ , Na^+ , K^+ , Be^{2+} , Mg^{2+} , and Ca^{2+} metal cations. *Journal of Physical Chemistry A*, 110, 6139–6144.
- Andrushchenko, V., & Bour, P. (2009). Infrared absorption detection of metal ion-deoxyguanosine monophosphate binding: Experimental and theoretical study. *Journal of Physical Chemistry B*, 113(1), 283–291. doi:10.1021/jp8058678.
- Anwander, E. H. S., Probst, M. M., & Rode, B. M. (1990). The influence of Li^+ , Na^+ , Mg^{2+} , Ca^{2+} , and Zn^{2+} ions on the hydrogen bonds of the Watson-Crick base pair. *Biopolymers*, 29, 757–769.
- Aquino, A. J. A., Tunega, D., Haberhauer, G., Gerzabek, M. H., & Lischka, H. (2008). Acid–base properties of a goethite surface model: A theoretical view. *Geochimica et Cosmochimica Acta*, 72, 3587–3602.
- Araplahti, J., & Lippert, B. (1990). Coordination of aquated cis-platinum(II) diamines to purine nucleosides. Kinetics of complex formation. *Inorganic Chemistry*, 29, 104–109.
- Bagchi, S., Mandal, D., Ghosh, D., & Das, A. K. (2012). Density functional theory study of interaction, bonding and affinity of group IIb transition metal cations with nucleic acid bases. *Chemical Physics*, 400, 108–117. doi:10.1016/j.chemphys.2012.03.003.
- Baik, M.-H., Friesner, R. A., & Lippard, S. J. (2002). Theoretical study on the stability of N-glycosyl bonds: Why does N7-platination not promote depurination? *Journal of the American Chemical Society*, 124, 4495–4503.
- Baik, M. H., Friesner, R. A., & Lippard, S. J. (2003). Theoretical study of cisplatin binding to purine bases: Why does cisplatin prefer guanine over adenine? *Journal of the American Chemical Society*, 125(46), 14082–14092.
- Banas, P., Jurecka, P., Walter, N. G., Šponer, J., & Otyepka, M. (2009). Theoretical studies of RNA catalysis: Hybrid QM/MM methods and their comparison with MD and QM methods. *Methods*, 49, 202–216.

- Bancroft, D. P., Lepre, C. A., & Lippard, S. J. (1990). Platinum-195 NMR kinetic and mechanistic studies of cis- and trans-diamminedichloroplatinum(II) binding to DNA. *Journal of the American Chemical Society*, *112*, 6860–6867.
- Bandyopadhyay, D., & Bhattacharyya, D. (2003). Different modes of interaction between hydrated magnesium ion and DNA functional groups: Database analysis and ab initio studies. *Journal of Biomolecular Structure & Dynamics*, *21*(3), 447–458.
- Basch, H., Krauss, M., Stevens, W. J., & Cohen, D. (1986). Binding of $\text{Pt}(\text{NH}_3)_3^{2+}$ to nucleic acid bases. *Inorganic Chemistry*, *25*, 684–688.
- Benda, L., Straka, M., Tanaka, Y., & Sychrovsky, V. (2011). On the role of mercury in the non-covalent stabilisation of consecutive U-HgII-U metal-mediated nucleic acid base pairs: Metallophilic attraction enters the world of nucleic acids. *Physical Chemistry Chemical Physics*, *13*(1), 100–103. doi:10.1039/C0CP01534B.
- Benda, L., Straka, M., Sychrovský, V., Bouř, P., & Tanaka, Y. (2012). Detection of mercury–TpT dinucleotide binding by Raman spectra: A computational study. *The Journal of Physical Chemistry. A*, *116*(32), 8313–8320. doi:10.1021/jp3045077.
- Benedikt, U., Schneider, W. B., & Auer, A. A. (2013). Modelling electrified interfaces in quantum chemistry: Constant charge vs. constant potential. *Physical Chemistry Chemical Physics*, *15*, 2712–2724.
- Besker, N., Coletti, C., Marrone, A., & Re, N. (2007). Binding of antitumor ruthenium complexes to DNA and proteins: A theoretical approach. *Journal of Physical Chemistry B*, *111*(33), 9955–9964. doi:10.1021/jp072182q.
- Boys, S. F., & Bernardi, F. (1970). The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics*, *19*, 553–566.
- Burda, J. V., & Gu, J. (2008). A computational study on DNA bases interactions with dinuclear tetraacetato-diaqua-dirhodium(II, II) complex. *Journal of Inorganic Biochemistry*, *102*, 53–62.
- Burda, J. V., & Leszczynski, J. (2003). How strong can the bend be on a DNA helix from cisplatin? DFT and MP2 quantum chemical calculations of cisplatin-bridged DNA purine bases. *Inorganic Chemistry*, *42*(22), 7162–7172.
- Burda, J. V., Šponer, J., & Hobza, P. (1996). Ab Initio study of the interaction of guanine and adenine with various mono- and bivalent metal cations (Li^+ , Na^+ , K^+ , Rb^+ , Cs^+ ; Cu^+ , Ag^+ , Au^+ ; Mg^{2+} , Ca^{2+} , Sr^{2+} , Ba^{2+} ; Zn^{2+} , Cd^{2+} , and Hg^{2+}). *Journal of Physical Chemistry*, *100*(17), 7250–7255.
- Burda, J. V., Šponer, J., Leszczynski, J., & Hobza, P. (1997). Interaction of DNA base pairs with various metal cations (Mg^{2+} , Ca^{2+} , Sr^{2+} , Ba^{2+} , Cu^+ , Ag^+ , Au^+ , Zn^{2+} , Cd^{2+} , and Hg^{2+}): Nonempirical ab initio calculations on structures, energies, and nonadditivity of the interaction. *Journal of Physical Chemistry B*, *101*(46), 9670–9677.
- Burda, J. V., Šponer, J., & Leszczynski, J. (2000). The interactions of square platinum(II) complexes with guanine and adenine: A quantum-chemical ab initio study of metalated tautomeric forms. *Journal of Biological Inorganic Chemistry*, *5*(2), 178–188.
- Burda, J. V., Šponer, J., & Leszczynski, J. (2001). The influence of square planar platinum complexes on DNA base pairing. An ab initio DFT study. *Physical Chemistry Chemical Physics*, *3*(19), 4404–4411.
- Burda, J. V., Šponer, J., Hrabáková, J., Zeizinger, M., & Leszczynski, J. (2003). The influence of N-7 guanine modifications on the strength of Watson-Crick base pairing and guanine N-1 acidity: Comparison of gas-phase and condensed-phase trends. *Journal of Physical Chemistry B*, *107*(22), 5349–5356.
- Burda, J. V., Pavelka, M., & Simanek, M. (2004). Theoretical model of copper Cu(I)/Cu(II) hydration. DFT and ab initio quantum chemical study. *Journal of Molecular Structure (THEOCHEM)*, *683*(1–3), 183–193. doi:10.1016/j.theochem.2004.06.013.
- Burda, J. V., Shukla, M. K., & Leszczynski, J. (2005). Theoretical model of the aqua-copper $[\text{Cu}(\text{H}_2\text{O})_5]^+$ cation interactions with guanine. *Journal of Molecular Modeling*, *11*, 362–369.
- Burda, J. V., Murray, J. S., Gutierrez-Oliva, S., Politzer, P., & Toro-Labbe, A. (2015). *Journal of Chemical Theory and Computation*. Submitted.

- Ceron-Carrasco, J. P., & Jacquemin, D. (2011). Influence of Mg²⁺ on the guanine-cytosine tautomeric equilibrium: Simulations of the induced intermolecular proton transfer. *Chemphyschem*, *12*(14), 2615–2623. doi:10.1002/cphc.201100264.
- Ceron-Carrasco, J. P., Jacquemin, D., & Cauet, E. (2012a). Cisplatin cytotoxicity: A theoretical study of induced mutations. *Physical Chemistry Chemical Physics*, *14*(36), 12457–12464. doi:10.1039/C2CP40515F.
- Ceron-Carrasco, J. P., Requena, A., & Jacquemin, D. (2012b). Impact of DFT functionals on the predicted magnesium-DNA interaction: An ONIOM study. *Theoretical Chemistry Accounts*, *131*(3). doi:10.1007/s00214-012-1188-9.
- Chu, V. B., Bai, Y., Lipfert, J., Herschlag, D., & Doniach, S. (2008). A repulsive field: Advances in the electrostatics of the ion atmosphere. *Current Opinion in Chemical Biology*, *12*, 619–625.
- Chval, Z., & Sip, M. (2003). Transition states of cisplatin binding to guanine and adenine: Ab initio reactivity study. *Collection of Czechoslovak Chemical Communications*, *68*, 1105–1118.
- Chval, Z., Kabelac, M., & Burda, J. V. (2013). Mechanism of the cis- Pt(1R,2R-DACH)(H₂O)(2) (2+) intrastrand binding to the double-stranded (pGpG)center dot(CpC) dinucleotide in aqueous solution: A computational DFT study. *Inorganic Chemistry*, *52*(10), 5801–5813. doi:10.1021/ic302654s.
- Colominas, C., Luque, F. J., & Orozco, M. (1996). Tautomerism and protonation of guanine and cytosine. Implications in the formation of hydrogen-bonded complexes. *Journal of the American Chemical Society*, *118*(29), 6811–6821.
- Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, J., Kenneth, M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., James Caldwell, J. W., & Kollman, P. A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, *117*, 5179–5197.
- Deeth, R. J. (2001). The ligand field molecular mechanics model and the stereoelectronic effects of d and s electrons. *Coordination Chemistry Reviews*, *212*(1), 11–34. doi:10.1016/S0010-8545(00)00354-4.
- Deeth, R. J., Anastasi, A., Diedrich, C., & Randell, K. (2009). Molecular modelling for transition metal complexes: Dealing with d-electron effects. *Coordination Chemistry Reviews*, *253*(5–6), 795–816. doi:10.1016/j.ccr.2008.06.018.
- Del Bene, J. E. (1984). Molecular orbital study of the Li⁺ complexes of the DNA bases. *Journal of Physical Chemistry*, *88*, 5927–5931.
- Del Bene, J. E. (1985). Molecular orbital theory of the hydrogen bond. *Journal of Molecular Structure*, *124*, 201–212.
- Deubel, D. V. (2003). Reactivity of osmium tetroxide towards nitrogen heterocycles: Implications for the molecular recognition of DNA mismatch. *Angewandte Chemie, International Edition*, *42*(17), 1974–1977. doi:10.1002/anie.200250462.
- Draper, D. E., Grilley, D., & Soto, A. M. (2005). Ions and RNA folding. *Annual Review of Biophysics and Biomolecular Structure*, *34*, 221–243.
- Egli, M., & Gessner, R. V. (1995). Stereoelectronic effects of deoxyribose O4' on DNA conformation. *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 180–184.
- Egli, M., Williams, L. D., Fredericks, C. A., & Rich, A. (1991). DNA-nogalamycin interactions. *Biochemistry*, *30*, 1364.
- Elmahdaoui, L., & Tajmirrahi, H. (1995). A comparative-study of Atp and Gtp complexation with trivalent Al, Ga and Fe cations – Determination of cation-binding site and nucleotide conformation by ftir difference spectroscopy. *Journal of Biomolecular Structure & Dynamics*, *13*(1), 69–86.
- Fortino, M., Marino, T., & Russo, N. (2015). Theoretical study of silver-ion-mediated base pairs: The case of C–Ag–C and C–Ag–A systems. *The Journal of Physical Chemistry. A*, *119*(21), 5153–5157. doi:10.1021/jp5096739.
- Frank Eckert, A. K. (2006). Accurate prediction of basicity in aqueous solution with COSMO-RS. *Journal of Computational Chemistry*, *27*(1), 11–19.

- Fuentes-Cabrera, M., Sumpster, B. G., & Šponer, J. E. (2007). Theoretical study on the structure, stability, and electronic properties of the guanine–Zn–cytosine base pair in M-DNA. *Journal of Physical Chemistry B*, *111*, 870–875.
- Futera, Z., & Burda, J. V. (2014). Reaction mechanism of Ru(II) piano-stool complexes: Umbrella sampling QM/MM MD study. *Journal of Computational Chemistry*, *35*, 1446–1456.
- Futera, Z., Klenko, J., Šponer, J. E., Šponer, J., & Burda, J. V. (2009a). Interactions of the “Piano-stool” [Ruthenium(II)(g6-arene)(en)Cl] complexes with water and nucleobases; ab initio and DFT study. *Journal of Computational Chemistry*, *30*, 1758–1770.
- Futera, Z., Klenko, J., Šponer, J. E., Šponer, J., & Burda, J. V. (2009b). Interactions of the “piano-stool” [ruthenium(II)(arene)(en)Cl] complexes with water and nucleobases; ab initio and DFT study. *Journal of Computational Chemistry*, *30*, 1758.
- Futera, Z., Klenko, J., Šponer, J. E., Šponer, J., & Burda, J. V. (2009c). Interactions of the “Piano-stool” ruthenium(II)(eta(6)-arene)(en)Cl (+) complexes with water and nucleobases; A initio and DFT study. *Journal of Computational Chemistry*, *30*(12), 1758–1770. doi:10.1002/jcc.21179.
- Futera, Z., Platts, J. A., & Burda, J. V. (2012). Binding of piano-stool Ru(II) complexes to DNA; QM/MM study. *Journal of Computational Chemistry*, *33*, 2092–2101.
- Gkionis, K., Mutter, S. T., & Platts, J. A. (2013). QM/MM description of platinum-DNA interactions: Comparison of binding and DNA distortion of five drugs. *RSC Advances*, *3*(12), 4066–4073. doi:10.1039/C3RA23041D.
- Gkionis, K., Kruse, H., Platts, J. A., Mládek, A., Koča, J., & Šponer, J. (2014). Ion binding to quadruplex DNA stems. Comparison of MM and QM descriptions reveals sizable polarization effects not included in contemporary simulations. *Journal of Chemical Theory and Computation*, *10*(3), 1326–1340. doi:10.1021/ct4009969.
- Gresh, N., Šponer, J. E., Špačková, N., Leszczynski, J., & Šponer, J. (2003). Theoretical study of binding of hydrated cations Zn(II) and Mg(II) to guanosine 5' monophosphate. Towards polarizable molecular mechanics for DNA and RNA. *The Journal of Physical Chemistry B*, *107*, 8669–8681.
- Gu, J. D., & Leszczynski, J. (2000). A remarkable alteration in the bonding pattern: An HF and DFT study of the interactions between the metal cations and the Hoogsteen hydrogen-bonded G-tetrad. *Journal of Physical Chemistry A*, *104*(26), 6308–6313.
- Hud, N. V., Smith, F. W., Anet, F. A. L., & Feigon, J. (1996). The selectivity for K⁺ versus Na⁺ in DNA quadruplexes is dominated by relative free energies of hydration: A thermodynamic analysis by ¹H NMR. *Biochemistry*, *35*(48), 15383–15390.
- Hud, N. V., Schultze, P., & Feigon, J. (1998). Ammonium ion as an NMR probe for monovalent cation coordination sites of DNA quadruplexes. *Journal of the American Chemical Society*, *120*, 6403–6408.
- Ida, R., & Wu, G. (2008). Direct NMR detection of alkali metal ions bound to G-quadruplex DNA. *Journal of the American Chemical Society*, *130*, 3590–3594.
- Kabeláč, M., & Hobza, P. (2006). Na⁺, Mg²⁺, and Zn²⁺ binding to all tautomers of adenine, cytosine, and thymine and the eight most stable keto/enol tautomers of guanine: A correlated ab initio quantum chemical study. *Journal of Physical Chemistry B*, *110*, 14515–14523.
- Klamt, A., & Schuurmann, G. (1993). Cosmo – A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *Journal of the Chemical Society, Perkin Transactions*, *2*(5), 799–805.
- Korth, M., Pitoňák, M., Řezáč, J., & Hobza, P. (2010). A transferable H-bonding correction for semiempirical quantum-chemical methods. *Journal of Chemical Theory and Computation*, *6*(1), 344–352. doi:10.1021/ct900541n.
- Kosenkov, D., Gorb, L., Shishkin, O. V., Šponer, J., & Leszczynski, J. (2008). Tautomeric equilibrium, stability, and hydrogen bonding in 2'-deoxyguanosine monophosphate complexed with Mg²⁺. *Journal of Physical Chemistry B*, *112*, 150–158.
- Kratochvílová, I., Golan, M., Vala, M., Špérová, M., Weiter, M., Páv, O., Šebera, J., Rosenberg, I., Sychrovský, V., Tanaka, Y., & Bickelhaupt, F. M. (2014). Theoretical and experimental study

- of charge transfer through DNA: Impact of mercury mediated T-Hg-T base pair. *The Journal of Physical Chemistry B*, 118(20), 5374–5381. doi:[10.1021/jp501986a](https://doi.org/10.1021/jp501986a).
- Lilley, D. M. J., & Eckstein, F. (2008). *Ribozymes and RNA catalysis*. Cambridge: Royal Society of Chemistry.
- Lipinski, J. (1989). Electronic structure of platinum(II) antitumor complexes and their interactions with nucleic acid bases. Part II. *Journal of Molecular Structure (THEOCHEM)*, 201, 295–305.
- Marenich, A. V., Cramer, C. J., & Truhlar, D. G. (2009). Universal solvation model based on solute electron density and a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *Journal of Physical Chemistry B*, 113, 6378–6396.
- Margiotta, N., Petruzzella, E., Platts, J. A., Mutter, S. T., Deeth, R. J., Rinaldo, R., Papadia, P., Marzilli, P. A., Marzilli, L. G., Hoeschele, J. D., & Natile, G. (2015). DNA fragment conformations in adducts with Kiteplatin. *Dalton Transactions*, 44(8), 3544–3556. doi:[10.1039/C4DT01796J](https://doi.org/10.1039/C4DT01796J).
- Marino, T. (2014). DFT investigation of the mismatched base pairs (T-Hg-T)₃, (U-Hg-U)₃, d(T-Hg-T)₂, and d(U-Hg-U)₂. *Journal of Molecular Modeling*, 20(6), 1–5. doi:[10.1007/s00894-014-2303-8](https://doi.org/10.1007/s00894-014-2303-8).
- Marino, T., Toscano, M., Russo, N., & Grand, A. (2004). Gas-phase interaction between DNA and RNA bases and copper(II) ion: A density functional study. *International Journal of Quantum Chemistry*, 98(4), 347–354.
- Marino, T., Mazzuca, D., Toscano, M., Russo, N., & Grand, A. (2007). Gas phase interaction of zinc ion with purine and pyrimidine DNA and RNA bases. *International Journal of Quantum Chemistry*, 107, 311–320.
- Marino, T., Mazzuca, D., Russo, N., Toscano, M., & Grand, A. (2010). On the interaction of rubidium and cesium mono-, strontium and barium bi-cations with DNA and RNA bases. A theoretical study. *International Journal of Quantum Chemistry*, 110(1), 138–147. doi:[10.1002/qua.22076](https://doi.org/10.1002/qua.22076).
- Marino, T., Russo, N., Toscano, M., & Pavelka, M. (2012). Theoretical investigation on DNA/RNA base pairs mediated by copper, silver, and gold cations. *Dalton Transactions*, 41(6), 1816–1823. doi:[10.1039/C1DT11028D](https://doi.org/10.1039/C1DT11028D).
- Martínez, J. M., Pappalardo, R. R., & Marcos, E. S. (1997). Study of the Ag⁺ hydration by means of a semicontinuum quantum-chemical solvation model. *Journal of Physical Chemistry A*, 101, 4444–4448.
- Matsubara, T., & Hirao, K. (2002). Density functional study of the binding of the cyclen-coordinated M(II) (M = Zn, Cu, Ni) complexes to the DNA base. Why is Zn better to bind? *Journal of Molecular Structure (THEOCHEM)*, 581, 203–213.
- Megger, D. A., Fonseca Guerra, C., Bickelhaupt, F. M., & Müller, J. (2011a). Silver(I)-mediated Hoogsteen-type base pairs. *Journal of Inorganic Biochemistry*, 105(11), 1398–1404. doi:[10.1016/j.jinorgbio.2011.07.005](https://doi.org/10.1016/j.jinorgbio.2011.07.005).
- Megger, D. A., Fonseca Guerra, C., Hoffmann, J., Brutschy, B., Bickelhaupt, F. M., & Müller, J. (2011b). Contiguous metal-mediated base pairs comprising two AgI ions. *Chemistry - A European Journal*, 17(23), 6533–6544. doi:[10.1002/chem.201002944](https://doi.org/10.1002/chem.201002944).
- Mennucci, B., Cammi, R., & Tomasi, J. (1998). Excited states and solvatochromic shifts within a nonequilibrium solvation approach: A new formulation of the integral equation formalism method at the self-consistent field, configuration interaction, and multiconfiguration self-consistent field level. *Journal of Chemical Physics*, 109(7), 2798–2807.
- Meyer, M., & Sühnel, J. (2008). Density functional study of adenine tetrads with N6–H6···N3 hydrogen bonds. *The Journal of Physical Chemistry. A*, 112(18), 4336–4341. doi:[10.1021/jp710242k](https://doi.org/10.1021/jp710242k).
- Meyer, M., Steinke, T., Brandl, M., & Sühnel, J. (2001). Density functional study of guanine and uracil quartets and of guanine quartet/metal ion complexes. *Journal of Computational Chemistry*, 22(1), 109–124.
- Meyer, M., Hocquet, A., & Sühnel, J. (2005). Interaction of sodium and potassium ions with sandwiched cytosine-, guanine-, thymine-, and uracil-base tetrads. *Journal of Computational Chemistry*, 26(4), 352–364. doi:[10.1002/jcc.20176](https://doi.org/10.1002/jcc.20176).

- Meyer, M., Steinke, T., & Sühnel, J. (2007). Density functional study of isoguanine tetrad and pentad sandwich complexes with alkali metal ions. *Journal of Molecular Modeling*, *13*(2), 335–345. doi:10.1007/s00894-006-0148-5.
- Miertus, S., Scrocco, E., & Tomasi, J. (1981). Electrostatic interaction of a solute with a continuum. A direct utilization of AB initio molecular potentials for the prevision of solvent effects. *Chemical Physics*, *55*(1), 117–129.
- Mlynsky, V., Walter, N. G., Šponer, J., Otyepka, M., & Banas, P. (2015). The role of an active site Mg^{2+} in HDV ribozyme self-cleavage: Insights from QM/MM calculations. *Physical Chemistry Chemical Physics*, *17*(1), 670–679. doi:10.1039/C4CP03857F.
- Morari, C., Bogdan, D., & Muntean, C. M. (2012). Binding effects of Mn^{2+} and Zn^{2+} ions on the vibrational properties of guanine-cytosine base pairs in the Watson-Crick and Hoogsteen configurations. *Journal of Molecular Modeling*, *18*(11), 4781–4786. doi:10.1007/s00894-012-1480-6.
- Morari, C., Muntean, C. M., Tripon, C., Buimaga-Iarinca, L., & Calborean, A. (2014). DFT investigation of the vibrational properties of GC Watson-Crick and Hoogsteen base pairs in the presence of Mg^{2+} , Ca^{2+} , and Cu^{2+} ions. *Journal of Molecular Modeling*, *20*(4). doi:10.1007/s00894-014-2220-x.
- Mukhopadhyay, A., Aguilar, B. H., Tolokh, I. S., & Onufriev, A. V. (2014). Introducing charge hydration asymmetry into the generalized born model. *Journal of Chemical Theory and Computation*, *10*, 1788–1794.
- Mutter, S., Margiotta, N., Papadia, P., & Platts, J. (2015). Computational evidence for structural consequences of kateplatin damage on DNA. *JBIC, Journal of Biological Inorganic Chemistry*, *20*(1), 35–48. doi:10.1007/s00775-014-1207-5.
- Noguera, M., Bertran, J., & Sodupe, M. (2004). A quantum chemical study of Cu^{2+} interacting with guanine-cytosine base pair. Electrostatic and oxidative effects on intermolecular proton-transfer processes. *Journal of Physical Chemistry A*, *108*(2), 333–341. doi:10.1021/jp036573q.
- Noguera, M., Branchadell, V., Constantino, E., Rios-Font, R., Sodupe, M., & Rodríguez-Santiago, L. (2007). On the bonding of first-row transition metal cations to guanine and adenine nucleobases. *Journal of Physical Chemistry A*, *111*(39), 9823–9829. doi:10.1021/jp073858k.
- Noguera, M., Bertran, J., & Sodupe, M. (2008). $Cu^{2+}/+$ cation coordination to adenine-thymine base pair. Effects on intermolecular proton-transfer processes. *Journal of Physical Chemistry B*, *112*(15), 4817–4825. doi:10.1021/jp711982g.
- Oliva, R., & Cavallo, L. (2009). Frequency and effect of the binding of Mg^{2+} , Mn^{2+} , and Co^{2+} ions on the guanine base in Watson-Crick and reverse Watson-Crick base pairs. *Journal of Physical Chemistry B*, *113*(47), 15670–15678. doi:10.1021/jp906847p.
- Parr, R. G., & Pearson, R. G. (1983). Absolute hardness: Companion parameter to absolute electronegativity. *Journal of the American Chemical Society*, *105*(26), 7512–7516.
- Pavelka, M., & Burda, J. V. (2005). Theoretical description of copper $Cu(I)/Cu(II)$ complexes in mixed ammine-aqua environment. DFT and ab initio quantum chemical study. *Chemical Physics*, *312*, 193–204. doi:10.1016/j.chemphys.2004.11.034.
- Pavelka, M., Šimánek, M., Šponer, J., & Burda, J. V. (2006). Copper cation interactions with biologically essential types of ligands: A computational DFT study. *Journal of Physical Chemistry A*, *110*, 4795–4809.
- Pavelka, M., Shukla, M. K., Leszczynski, J., & Burda, J. V. (2008). Theoretical study of hydrated copper(II) interactions with guanine: A computational density functional theory study. *Journal of Physical Chemistry A*, *112*(2), 256–267. doi:10.1021/jp074891+.
- Petrov, A. S., Lamm, G., & Pack, G. R. (2005). Calculation of the binding free energy for magnesium – RNA interactions. *Biopolymers*, *77*(3), 137–154.
- Poltve, V. I., Malenkov, G. G., Gonzales, E. J., Teplukhin, A. V., Rein, R., Shibata, M., & Miller, J. H. (1996). Modeling DNA hydration: Comparison of calculated and experimental hydration properties of nucleic acid bases. *Journal of Biomolecular Structure and Dynamics*, *13*(4), 717–725.
- Potaman, V. N., & Soyfer, V. N. (1994). Divalent metal cations upon coordination to the N7 of purines differentially stabilize the PyPuPu DNA triplex due to unequal Hoogsteen-type hydrogen bond enhancement. *Journal of Biomolecular Structure & Dynamics*, *11*, 1035–1040.

- Raber, J., Zhu, C., & Eriksson, L. A. (2005). Theoretical study of cisplatin binding to DNA: The importance of initial complex stabilisation. *Journal of Physical Chemistry*, *109*, 11006–11015.
- Reshetnikov, R. V., Šponer, J., Rassokhina, O. I., Kopylov, A. M., Tsvetkov, P. O., Makarov, A. A., & Golovin, A. V. (2011). Cation binding to 15-TBA quadruplex DNA is a multiple-pathway cation-dependent process. *Nucleic Acids Research*, *39*(22), 9789–9802. doi:10.1093/nar/gkr639.
- Řezáč, J., Fanfrlík, J., Salahub, D., & Hobza, P. (2009). Semiempirical quantum chemical PM6 method augmented by dispersion and H-bonding correction terms reliably describes carious types of noncovalent complexes. *Journal of Chemical Theory and Computation*, *5*(7), 1749–1760. doi:10.1021/ct9000922.
- Rosenberg, B., Van Camp, L., & Krigas, T. (1965). Inhibition of cell division in *Escherichia coli* by electrolysis products from a platinum electrode. *Nature*, *205*(4972), 698–699.
- Rosenberg, B., Van Camp, L., Trosko, J. L., & Mansour, V. H. (1969). Platinum drugs: a new class of potent antitumor agents. *Nature*, *222*, 385–391.
- Roux, B., Yu, H. A., & Karplus, M. (1990). Molecular basis for the born model of ion solvation. *Journal of Physical Chemistry*, *94*, 4683–4688.
- Rozsnyai, F., & Ladik, J. (1970). Calculation of effects of hydration and divalent metal ions on DNA base pairs. *Bulletin of the American Physical Society*, *15*(3), 325.
- Rulisek, L., & Šponer, J. (2003). Outer-shell and inner-shell coordination of phosphate group to hydrated metal ions (Mg^{2+} , Cu^{2+} , Zn^{2+} , Cd^{2+}) in the presence and absence of nucleobase. The role of nonelectrostatic effects. *Journal of Physical Chemistry B*, *107*(8), 1913–1923. doi:10.1021/jp027058f.
- Russo, N., & Toscano, M., A. G. (2001). Lithium affinity for DNA and RNA nucleobases. The role of theoretical information in the elucidation of the mass spectrometry data. *Journal of Physical Chemistry B*, *105*, 4735–4741.
- Russo, N., Toscano, M., & Grand, A. (2003). Gas-phase absolute Ca^{2+} and Mg^{2+} affinity for nucleic acid bases. A theoretical determination. *Journal of Physical Chemistry A*, *107*(51), 11533–11538. doi:10.1021/jp0358681.
- Saenger, W. (1983). *Principles of nucleic acid structure*. New York: Springer.
- Santangelo, M. G., Antoni, P. M., Spingler, B., & Jeschke, G. (2010). Can copper(II) mediate Hoogsteen base-pairing in a left-handed DNA duplex? A pulse EPR study. *ChemPhysChem*, *11*, 599–606.
- Schmidt, K. S., Reedijk, J., Weisz, K., Basilio Janke, E. M., Šponer, J. E., Šponer, J., & Lippert, B. (2002). Loss of Hoogsteen pairing ability upon N1 adenine platinum binding. *Inorganic Chemistry*, *41*, 2855–2863.
- Schreiber, M., & Gonzalez, L. (2007a). Structure and bonding of Ag(I)-DNA base complexes and Ag(I)-adenine-cytosine mismatches: An ab initio study. *Journal of Computational Chemistry*, *28*, 2299–2308. doi:10.1002/jcc.20743.
- Schreiber, M., & Gonzalez, L. (2007b). The role of Ag(I) ions in the electronic spectroscopy of adenine-cytosine mismatches: A MS-CASPT2 theoretical study. *Journal of Photochemistry and Photobiology A: Chemistry*, *190*(2–3), 301–309. doi:10.1016/j.jphotochem.2007.01.035.
- Šebera, J., Burda, J. V., Straka, M., Ono, A., Kojima, C., Tanaka, Y., & Sychrovský, V. (2013a). Formation of a thymine-Hg II-thymine metal-mediated DNA base pair: Proposal and theoretical calculation of the reaction pathway. *Chemistry - A European Journal*, *19*, 9884–9894.
- Šebera, J., Burda, J., Straka, M., Ono, A., Kojima, C., Tanaka, Y., & Sychrovský, V. (2013b). Formation of a thymine-HgII-thymine metal-mediated DNA base pair: Proposal and theoretical calculation of the reaction pathway. *Chemistry - A European Journal*, *19*(30), 9884–9894. doi:10.1002/chem.201300460.
- Šebesta, F., & Burda, J. V. (2016). Reduction process of tetraplatin in the presence of deoxyguanosine monophosphate (dGMP): A computational DFT study. *Chemistry - A European Journal*, *22*, 1037–1047.
- Setnicka, V., Novy, J., Bohm, S., Sreenivasachary, N., Urbanova, M., & Volka, K. (2008). Molecular structure of guanine-quartet supramolecular assemblies in a gel-state based on a DFT calculation of infrared and vibrational circular dichroism spectra. *Langmuir*, *24*(14), 7520–7527. doi:10.1021/la800611h.

- Sigel, H. (1993). Interactions of metal ions with nucleotides and nucleic acids and their constituents. *Chemical Society Reviews*, 22, 255–267.
- Sigel, R. K. O., & Lippert, B. (1999). Pt^{II} coordination to guanine-N7: Enhancement of the stability of the Watson–Crick base pair with cytosine. *Chemical Communications*, 2167.
- Šponer, J., & Špacková, N. (2007). Molecular dynamics simulations and their application to four-stranded DNA. *Methods*, 43, 278–284.
- Šponer, J., Burda, J. V., Mejzlik, P., Leszczynski, J., & Hobza, P. (1997). Hydrogen-bonded trimers of DNA bases and their interaction with metal cations: Ab initio quantum-chemical and empirical potential study. *Journal of Biomolecular Structure & Dynamics*, 14(5), 613–628.
- Šponer, J., Burda, J. V., Sabat, M., Leszczynski, J., & Hobza, P. (1998a). Interaction between the guanine-cytosine Watson-Crick DNA base pair and hydrated group IIa (Mg²⁺, Ca²⁺, Sr²⁺, Ba²⁺) and group IIb (Zn²⁺, Cd²⁺, Hg²⁺) metal cations. *Journal of Physical Chemistry A*, 102(29), 5951–5957.
- Šponer, J., Sabat, M., Burda, J. V., Doody, A. M., Leszczynski, J., & Hobza, P. (1998b). Stabilization of the purine center dot purine center dot pyrimidine DNA base triplets by divalent metal cations. *Journal of Biomolecular Structure & Dynamics*, 16(1), 139–143.
- Šponer, J., Sabat, M., Burda, J., Leszczynski, J., Hobza, P., & Lippert, B. (1999a). Metal ions in non-complementary DNA base pairs: An ab initio study of Cu(I), Ag(I), and Au(I) complexes with cytosine-adenine base pair. *Journal of Biological Inorganic Chemistry*, 4, 537–545.
- Šponer, J., Burda, J. V., Leszczynski, J., & Hobza, P. (1999b). Interactions of hydrated IIa and IIb group metal cations with thio-guanine-cytosine DNA base pair: Ab initio and density functional theory investigation of polarization effects, differences among cations, and flexibility of the cation hydration shell. *Journal of Biomolecular Structure & Dynamics*, 17(1), 61–77.
- Šponer, J., Šponer, J. E., Gorb, L., Leszczynski, J., & Lippert, B. (1999c). Metal-stabilized rare tautomers and mismatches of DNA bases: N6-metalated adenine and N4-metalated cytosine, theoretical and experimental views. *Journal of Physical Chemistry A*, 103, 11406–11413.
- Šponer, J., Sabat, M., Burda, J. V., Leszczynski, J., & Hobza, P. (1999d). Interaction of the adenine-thymine Watson-Crick and adenine-adenine reverse-Hoogsteen DNA base pairs with hydrated group IIa (Mg²⁺, Ca²⁺, Sr²⁺, Ba²⁺) and IIb (Zn²⁺, Cd²⁺, Hg²⁺) metal cations: Absence of the base pair stabilization by metal-induced polarization effects. *Journal of Physical Chemistry B*, 103(13), 2528–2534.
- Šponer, J., Sabat, M., Burda, J. V., Leszczynski, J., Hobza, P., & Lippert, B. (1999e). Metal ions in non-complementary DNA base pairs: An ab initio study of Cu(I), Ag(I), and Au(I) complexes with the cytosine-adenine base pair. *Journal of Biological Inorganic Chemistry*, 4(5), 537–545.
- Šponer, J., Šponer, J. E., & Leszczynski, J. (2000a). Cation – Pi and amino-acceptor interactions between hydrated metal cations and DNA bases. A quantum-chemical view. *Journal of Biomolecular Structure and Dynamics*, 17(6), 1087–1096.
- Šponer, J., Sabat, M., Gorb, L., Leszczynski, J., Lippert, B., & Hobza, P. (2000b). The effect of metal binding to the N7 site of purine nucleotides on their structure, energy, and involvement in base pairing. *Journal of Physical Chemistry B*, 104(31), 7535–7544. doi:10.1021/jp001711m.
- Šponer, J. E., Leszczynski, J., Glahe, F., Lippert, B., & Šponer, J. (2001). Protonation of platinated adenine nucleobases. Gas phase vs condensed phase picture. *Inorganic Chemistry*, 40, 3269–3278.
- Šponer, J., Mládek, A., Špačková, N., Cang, X., Cheatham, T. E., & Grimme, S. (2013). Relative stability of different DNA guanine quadruplex stem topologies derived using large-scale quantum-chemical computations. *Journal of the American Chemical Society*, 135(26), 9785–9796. doi:10.1021/ja402525c.
- Sychrovsky, V., Šponer, J., & Hobza, P. (2004). Theoretical calculation of the NMR spin-spin coupling constants and the NMR shifts allow distinguishability between the specific direct and the water-mediated binding of a divalent metal cation to guanine. *Journal of the American Chemical Society*, 126(2), 663–672.
- Tanaka, Y., Kojima, C., Morita, E. H., Kasai, Y., Yamasaki, K., Ono, A., Kainosho, M., & Taira, K. (2002). Identification of the metal ion binding site on an RNA Motif from hammerhead ribozymes using ¹⁵N NMR spectroscopy. *Journal of the American Chemical Society*, 124, 4595–4601.

- van der Wijst, T., Guerra, C. F., Swart, M., Bickelhaupt, F. M., & Lippert, B. (2009). Rare tautomers of 1-methyluracil and 1-methylthymine: Tuning relative stabilities through coordination to Pt-II complexes. *Chemistry - A European Journal*, *15*(1), 209–218. doi:[10.1002/chem.200801476](https://doi.org/10.1002/chem.200801476).
- Varnali, T., & Tozumcalgan, D. (1995). Interaction of divalent metal-cations and nucleotides – A computational study. *Structural Chemistry*, *6*(4–5), 343–348.
- Wang, F. Y., Habtemariam, A., van der Geer, E. P. L., Fernandez, R., Melchart, M., Deeth, R. J., Aird, R., Guichard, S., Fabbiani, F. P. A., Lozano-Casal, P., Oswald, I. D. H., Jodrell, D. I., Parsons, S., & Sadler, P. J. (2005). Controlling ligand substitution reactions of organometallic complexes: Tuning cancer cell cytotoxicity. *PNAS*, *102*(51), 18269–18274.
- Wu, Y. B., Bhattacharyya, D., King, C. L., Baskerville-Abraham, I., Huh, S. H., Boysen, G., Swenberg, J. A., Temple, B., Campbell, S. L., & Chaney, S. G. (2007). Solution structures of a DNA dodecamer duplex with and without a cisplatin 1,2-d(GG) intrastrand cross-link: Comparison with the same DNA duplex containing an oxaliplatin 1,2-d(GG) intrastrand cross-link. *Biochemistry*, *46*(22), 6477–6487.
- Yamaguchi, H., Šebera, J., Kondo, J., Oda, S., Komuro, T., Kawamura, T., Daraku, T., Kondo, Y., Okamoto, I., Ono, A., Burda, J. V., Kojima, C., Sychrovský, V., & Tanaka, Y. (2014). The structure of metallo-DNA with consecutive T-HgII-T base-pairs explains positive reaction entropy for the metallo-base-pair formation. *Nucleic Acids Research*, *42*, 4094–4099.
- Yurenko, Y. P., Novotny, J., Sklenar, V., & Marek, R. (2014). Exploring non-covalent interactions in guanine- and xanthine-based model DNA quadruplex structures: A comprehensive quantum chemical approach. *Physical Chemistry Chemical Physics*, *16*(5), 2072–2084. doi:[10.1039/C3CP53875C](https://doi.org/10.1039/C3CP53875C).
- Zeizinger, M., Burda, J. V., & Leszczynski, J. (2004). The influence of a sugar-phosphate backbone on the cisplatin-bridged BpB models of DNA purine bases. Quantum chemical calculations of Pt(II) bonding characteristics. *Physical Chemistry Chemical Physics*, *6*(10), 3585–3590.
- Zhang, Y., & Huang, K. X. (2007). On the interactions of hydrated metal cations (Mg²⁺, Mn²⁺, Ni²⁺, Zn²⁺) with guanine-cytosine Watson-Crick and guanine-guanine reverse-Hoogsteen DNA base pairs. *Journal of Molecular Structure (THEOCHEM)*, *812*(1–3), 51–62. doi:[10.1016/j.theochem.2007.02.009](https://doi.org/10.1016/j.theochem.2007.02.009).
- Zilberberg, I. L., Avdeev, V. I., & Zhidomirov, G. M. (1997). Effect of cisplatin binding on guanine in nucleic acid: An ab initio study. *Journal of Molecular Structure (THEOCHEM)*, *418*, 73.

M. Alaraby Salem, Melis Gedik, and Alex Brown

Contents

Introduction	1876
Theory	1877
The Two-Level Model	1878
Computational Implementation	1879
Applications	1881
Fluorescent Proteins	1881
Nucleic Acid Base Analogues	1886
Conclusions	1889
Bibliography	1889

Abstract

Two-photon absorption (TPA) leads to higher-energy excited electronic states via the simultaneous absorption of two photons. In TPA, the absorption is directly proportional to the square of incident light intensity, and thus lasers are required for excitation. The advantages of TPA microscopy include better focus and less out-of-focus bleaching, together with absorption at longer wavelengths than in one-photon absorption, which leads to deeper penetration in scattering media, such as tissues. However, TPA probes are usually associated with less sensitivity, and thus designing TPA fluorophores with large absorption probability is an important area of research. TPA of biological molecules like fluorescent proteins and nucleic acids is of particular interest. These molecules are experimentally produced through utilization of the naturally present transcription mechanism in the cell and thus pose less cell toxicity. In this chapter, we review the theory of TPA highlighting the computational approaches used to study biological

M.A. Salem • M. Gedik • A. Brown (✉)

Department of Chemistry, University of Alberta, Edmonton, AB, Canada

e-mail: msalem@ualberta.ca; mgedik@ualberta.ca; alex.brown@ualberta.ca

molecules. We discuss the computational methods available for exploring TPA and recent computational studies on the TPA of fluorescent proteins and nucleic acid base analogues. The chapter concludes by highlighting possible research avenues and unanswered questions.

Introduction

Two-photon absorption (TPA) is defined as the simultaneous absorption of two photons, of the same or different energies, leading to excitation to a higher electronic state. Although this phenomenon was predicted theoretically in 1931 by Goppert-Mayer, experimental evidence was not available until 1961 (Kaiser and Garrett 1961) after the development of lasers. Since TPA is a third-order nonlinear process, where the absorption is directly proportional to the square of incident light intensity, intense light sources, i.e., lasers, are required to observe significant TPA. The quadratic dependence on light intensity provides better focus and less out-of-focus bleaching and thus deeper penetration in scattering media, such as tissues, as compared to one-photon absorption (OPA) (Denk et al. 1990; Drobizhev et al. 2011). However, there is a drawback as TPA probes are usually associated with less sensitivity. This drove the desire to design TPA fluorophores with large TPA probabilities; see, for example, the reviews by Pawlicki et al. (2009), Ciuciu et al. (2014), and Morales et al. (2013) for more recent work. In TPA, the transition to the excited electronic state is achieved by two photons, and thus the energy of incident photons is approximately half of those used for OPA in the same system. Therefore, the excited state leading to fluorescence can be accessed at longer wavelengths, desirable in biological media where the absorption in the IR or near-IR region is required to overcome the competitive absorption of other naturally present pigments. Moreover, TPA follows different quantum mechanical selection rules from OPA, and therefore both spectra can be used in a complementary fashion (Drobizhev et al. 2011).

The focus of this chapter is the TPA of biological molecules which is gaining more interest as exemplified by the increasing number of recent studies cited here. Optimizing the photophysical TPA properties of proteins and other molecules that could be naturally present in biological systems is of paramount advantage. These molecules are produced through hijacking the transcription mechanism naturally present in the cell, and thus they are usually more benign to the complex cell environment than synthetically introduced dyes or quantum dots that may have other interactions and/or cell toxicity. Moreover, no membrane penetration is needed in the case of endogenously produced biological molecules. For more information on the advantages of using biological molecules, see, for instance, Campbell and Davidson (2010) about fluorescent proteins and the references therein.

In this chapter, we review the theory of TPA highlighting the computational approaches used to study biological molecules. We then discuss the computational studies on the TPA of fluorescent proteins which are of central importance in modern biological imaging and the TPA of nucleic acid base analogues which provide a

newer imaging tool that appears to be very promising. We conclude each section with open questions that should drive more computational (and experimental) work.

Theory

The theory of multiphoton absorption has been reviewed by Cronstrand et al. (2005). Here, we present a brief description of the theory of TPA. The theoretical description of TPA can be understood by considering the interaction of an external electric field, E , with a dielectric material. The induced polarization, P , within the dielectric material is given by

$$P = \varepsilon_0 \chi^{(1)} E + \varepsilon_0 \chi^{(2)} E^2 + \varepsilon_0 \chi^{(3)} E^3 + \dots, \quad (1)$$

where ε_0 is the permittivity of free space, and $\chi^{(n)}$ is n^{th} -order electric susceptibility: $\chi^{(1)}$ represents the linear susceptibility and the quantities $\chi^{(2)}, \chi^{(3)}, \dots$ denote the nonlinear susceptibilities. The two-photon absorption probability (or the cross section, σ^{TPA}) is related to the imaginary part of the third-order susceptibility, as

$$\sigma^{TPA} = \frac{24\pi^2 \hbar \omega^2}{c^2} \text{Im}(\chi^{(3)}), \quad (2)$$

where ω is the photon energy, c is the speed of light, and \hbar is the (reduced) Planck constant.

Susceptibilities are closely related to (hyper)polarizabilities which can be, in principle, computed by quantum mechanical methods (Cronstrand et al. 2005). However, a great simplification can be achieved when the photon energy is such that near-resonant conditions are achieved (Cronstrand et al. 2005). The two-photon transition matrix can then be defined as

$$S_{\alpha\beta} = \sum_n \left[\frac{\langle 0 | \mu_\alpha | n \rangle \langle n | \mu_\beta | f \rangle}{\omega_n - \omega} + \frac{\langle 0 | \mu_\beta | n \rangle \langle n | \mu_\alpha | f \rangle}{\omega_n - \omega} \right] \quad (3)$$

where μ_α (or μ_β) refers to the dipole moment operator in one Cartesian direction ($\alpha, \beta = x, y$, and z), ω_n is the energy spacing from the ground state, $|0\rangle$, to the intermediate state $|n\rangle$, and $|f\rangle$ is the final excited state.

Although this sum-over-states expression (SOS) has been used to compute TPA (see, e.g., Yang et al. 2005), it is generally not computationally efficient due to its slow convergence. On the other hand, response theory offers a more computationally efficient framework through the residue analysis of (hyper)polarizabilities. The two-photon transition matrix elements, $S_{\alpha\beta}$, and the transition dipole moments between excited states can be deduced from the single and double residues of the quadratic response function, respectively (Olsen and Jørgensen 1985).

For linearly polarized light, the transition moment for TPA is

$$\delta^{TPA} = \frac{1}{15} \sum_{\alpha\beta} [S_{\alpha\alpha} S_{\beta\beta}^* + 2S_{\alpha\beta} S_{\alpha\beta}^*]. \quad (4)$$

It can then be shown that the transition moment is related to the previous definition in Eq. 2 by

$$\sigma^{TPA} = \frac{24\pi^2 \hbar \omega^2}{c^2} \text{Im}(\chi^{(3)}) = \frac{4\pi^3 a_0^5 \alpha \omega^2}{c} \delta^{TPA} \quad (5)$$

where α is the fine structure constant and a_0 is the Bohr radius. The two-photon absorption cross section can be related to experimental measurements by including a term for lifetime broadening, $\frac{1}{\pi\Gamma}$, derived from a Lorentzian function:

$$\sigma^{TPA} = \frac{4\pi^2 a_0^5 \alpha \omega^2}{c} \frac{\delta^{TPA}}{\Gamma}, \quad (6)$$

where $\omega = \frac{\omega_f}{2}$ is the photon energy and Γ is the broadening factor. In previous computations, Γ has been chosen as 0.1 eV (Cronstrand et al. 2005; Nifosi and Luo 2007a; Salem and Brown 2014) or 0.5 eV (Friese et al. 2012; Silva et al. 2009) for comparison with experiment. The TPA cross section, σ^{TPA} , is conveniently reported in units of $10^{-50} \text{ cm}^4 \text{ s molecule}^{-1} \text{ photon}^{-1}$ that is referred to as one GM (which stands for ‘‘Goepfert-Mayer’’). Care should be taken when values of computed cross sections are compared in microscopic units due to variations caused by different forms of Eq. 6. In a recent comprehensive study, Beerepoot et al. (2015) discuss the reasons behind these variations and give recommendations for the details that should be provided when TPA cross sections are reported: (i) excitation energies and TPA strengths in atomic units, (ii) the formulae for the conversion to macroscopic units (Eq. 6) and that of the transition moment (Eq. 4), (iii) the proper choice of the conversion to match the experiments (here, we use the full definition for δ^{TPA} and scale Eq. 6 for comparison with single-beam experiments), and (iv) the type of lineshape function (Lorentzian or Gaussian) and the broadening factor. The interested reader is invited to examine the details of the discussion provided by Beerepoot et al. (2015).

The Two-Level Model

The SOS expression (Eq. 3) can be truncated to include only the initial and final (two) states, the so-called two-level model. More generally, the most significant transitions in an excitation scheme can be considered in an N -level model. Such

truncations can be used to relate TPA cross sections to transition dipole moments and permanent dipole moments of the ground state and excited states (Meath and Jagatap 2011, 2013; Meath et al. 2006).

For the first excited state ($f = 1$), and setting $\omega_0 = 0$, the SOS expression can be written within the two-level model as

$$S_{\alpha\beta} = \frac{2}{\omega_1} [\mu_{\beta 01} (\mu_{\alpha 11} - \mu_{\alpha 00}) + \mu_{\alpha 01} (\mu_{\beta 11} - \mu_{\beta 00})]. \quad (7)$$

For the same Cartesian direction ($\alpha = \beta$), it can be expressed as

$$S_{\alpha\alpha} = \frac{4}{\omega_1} \mu_{\alpha 01} (\mu_{\alpha 11} - \mu_{\alpha 00}). \quad (8)$$

From this expression and Eqs. (3) and (4), it is clear that (within the two-level model) the TPA cross section is proportional to the squares of the transition dipole moment and the difference between the permanent dipole moments of the excited and ground states. Hence, structural or environmental effects that increase the extent of charge delocalization upon excitation (while having a large transition dipole moment) would increase the TPA of a system (see section “Fluorescent Proteins”).

Computational Implementation

In this section, key computational implementations of TPA theory within electronic structure programs that can be used for studying biological molecules are highlighted. The interested reader is directed to see the original references for a more in-depth presentation of the computational details. In principle, any method that can provide excited-state properties, i.e., transition dipole moments between and permanent dipole moments for excited states, can be used to compute TPA cross sections within the SOS formalism (with the concomitant slow convergence mentioned previously). In practice, however, TDDFT (Runge and Gross 1984) with quadratic response (Safek et al. 2003; Frediani et al. 2005; Tretiak and Chernyak 2003) has been used extensively to compute TPA cross sections for large molecules because of its relatively low computational cost and, within known limitations, good accuracy (Nifosì and Luo 2007a; Yuan et al. 2013; Terenziani et al. 2008; Beerepoot et al. 2014; Nifosì and Luo 2007b; Vivas et al. 2010; Kamarchik and Krylov 2011; Steindal et al. 2012; Nayyar and Tretiak 2013). In standard response theory, each excited state within the frequency range of interest has to be addressed separately to obtain an absorption spectrum. This complicates the problem in large molecules, as they have high excited-state densities. Damped response theory has been formulated to compute TPA spectra in such cases (Kristensen et al. 2011).

Until recently, higher-level *ab initio* methods were limited to small molecules. A benchmark in 2006 evaluated the effect of electron correlation in TPA computation (for small molecules) using a hierarchy of coupled cluster (CC) models ranging from CC2 to CC3 (Koch et al. 1997; Christiansen et al. 1995a), and the results were compared to those from TDDFT (Paterson et al. 2006). Both Pople- and Dunning-style basis sets were tested. The results highlight the importance of diffuse basis functions especially in larger molecules with the Dunning series performing significantly better than Pople basis sets. These early results showed that TPA cross sections computed via TDDFT with the CAM-B3LYP functional (Yanai et al. 2004) and a modest basis set were in good agreement with the CC3 results (to within 15%). Another study (Friese et al. 2012) on the yellow fluorescent protein (YFP) (Wachter et al. 1998) showed that TPA results of CAM-B3LYP are in qualitative agreement with those from resolution-of-identity (RI) CC2 (Christiansen et al. 1995b; Hättig et al. 1998). A recent benchmark compared the TPA cross sections of fluorescent protein (FP) chromophores as computed by TDDFT for four functionals (B3LYP (Becke 1988), PBE0 (Adamo 1999), CAM-B3LYP, and LC-BLYP (Iikura et al. 2001; Tawada et al. 2004)) and full CC2 to averaged experimental values (see section “Fluorescent Proteins”) (Salem and Brown 2014). Results showed that the B3LYP (Becke 1993) functional with the modest 6-31+G(d,p) basis set can be used to semiquantitatively compare the TPA for the lowest energy excitation of fluorescent protein chromophores. Generally in the response theory formulation, time-dependent perturbation theory is applied to an approximate state, and properties are derived from the residues of the relevant response functions (*vide supra*). An alternative formulation employing the expectation value approach has been developed (Nanda and Krylov 2015) for computing TPA with the equation-of-motion coupled cluster model with single and double substitutions (EOM-CCSD) (Stanton and Bartlett 1993; Bartlett 2010). The method utilizes the RI and Cholesky decomposition (CD) techniques (for further information regarding these approximations, please see Nanda and Krylov (2015) and the references therein). In this approach, expressions for exact states are first derived, and then approximate wave functions are used for evaluation of matrix elements. Other methods that have been applied to large biological molecules, e.g., proteins, are reviewed in section “Understanding TPA of FPs via Computational Studies.”

A number of the computational approaches developed for TPA evaluation are now available in distributed *ab initio* electronic structure software. The DALTON (Aidas et al. 2014) program has flexibility in TPA computation by having both TDDFT and the CC hierarchy of methods with quadratic response available. The GAMESS-US (Schmidt et al. 1993) program can also be utilized for TPA calculations based on TDDFT. Other software includes Q-Chem with recent TPA implementation for the algebraic diagrammatic construction (ADC) method (Shao et al. 2015; Krylov and Gill 2013) and EOM-CCSD (Nanda and Krylov 2015). The implementation of RI-CC2 in Turbomole (2015) (not yet in the latest publicly distributed version, i.e., Version 7.0) also opens the field for the computation of TPA for relatively large biological systems (Wachter et al. 1998).

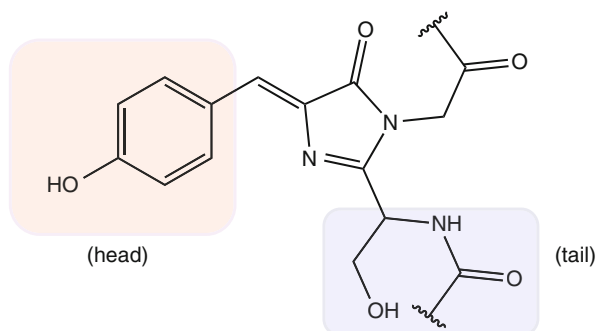
Applications

Fluorescent Proteins

Fluorescent proteins (FPs) are a family of homologues of the *Aequorea victoria* green fluorescent protein (*avGFP*) that was discovered in the 1960s (Shimomura et al. 1962). Later, the cloning of the GFP gene (Prasher et al. 1992) and its expression in other organisms while maintaining its fluorescent properties (Chalfie et al. 1994; Inouye and Tsuji 1994) paved the avenue to a new era of microscopy and cell biology. FPs are capable of forming a fluorescent chromophore by a posttranslational modification of three amino acids (Ormo et al. 1996; Yang et al. 1996). Changing these precursory amino acids yields different chromophores (see Fig. 1). The chromophore structure provides a coarse tuning of the color of the fluorescence associated with a given FP. Fine tuning is provided by the rest of the protein residues which form a barrel shape encapsulating the chromophore, while changes in the close-by residues of the chromophore can also strongly influence its spectral properties, as further discussed below. Among the full palette of FPs that is now available spanning red to blue FPs (Nifosì and Tozzini 2012; Shaner et al. 2005), red-shifted FPs have been of great interest due to their lower cell toxicity and better tissue penetration for light absorption and emission at these wavelengths (Weissleder 2001). For more information about FPs, the interested reader is directed to any of the available reviews (see, for instance, the book chapter by Campbell and Davidson (2010) and the references therein). The majority of photophysical studies on FPs focused on their OPA with a recent growing interest in TPA properties. Both OPA and TPA properties of FPs have been reviewed by Nifosì and Tozzini (2012).

Experimentally, there is a known general difficulty in measuring absolute TPA cross sections (Oulianov et al. 2001). Measurements in biological systems like FPs are more challenging, and, therefore, the reported measurements of TPA cross sections for a single FP may differ by up to two orders of magnitude, e.g., the reported measurements of the low-energy peak of EGFP (Zhang et al. 1996) vary from 1.5 GM (Hosoi et al. 2008) to 180 GM (Kawano et al. 2008). In 2011, Drobizhev et al. comprehensively cited the available experimental measurements

Fig. 1 GFP chromophore highlighting the head and tail positions that vary between FP chromophores



of TPA cross sections of FPs and discussed the sources of experimental error that could lead to such discrepancies (Drobizhev et al. 2011). Nevertheless, while there are discrepancies in the quantitative determination of the TPA cross sections, the experimentally determined TPA spectra of FPs exhibit common qualitative features: (i) The TPA spectrum generally shows two regions of absorption, one at (nearly) double the wavelength of the OPA peak and an additional (usually stronger) absorption band at a shorter wavelength. (ii) The long-wavelength (low-energy) TPA peak is blueshifted with respect to the corresponding OPA peak in the FPs with anionic chromophores (Drobizhev et al. 2011). In the rest of this section, we discuss the levels of complexity and the various computational approaches and benchmarks associated with TPA studies of FPs highlighting how computations improved our understanding of TPA in FPs and some of the unanswered questions.

Levels of Complexity and Computational Approaches

In general, three levels of complexity influence the TPA of a given FP: (i) The intrinsic nature of the chromophore which is directly related to the extent of π -conjugation and the involved transition and permanent dipole moments (see below) (Nifosi and Luo 2007a; Salem and Brown 2014). (ii) The effect of the protein shell on the geometry of the chromophore (Salem and Brown 2014, 2015; List et al. 2012). This includes consideration of both the change of chromophore conformation due to nearby residues and the conformational flexibility permitted by the hydrogen-bond network of interactions (including with water in the protein cavity) and the steric hindrance inside the protein pocket. (iii) The electric field of the protein surrounding the chromophore, which can greatly influence the TPA cross section (Drobizhev et al. 2011; List et al. 2012).

The first level of complexity is addressed by studying the isolated chromophore. The GFP chromophore is shown in Fig. 1, where we highlight the two sites that vary across FP chromophores. The “head” arises from the R-group of the amino acid Tyr66 which is the second amino acid in the tripeptide that matures into the chromophore. Point mutation of Tyr66 changes the head of the resulting chromophore. The “tail” of the chromophore arises from the amino acid Ser65. An additional step occurs in the maturation of the red FP family resulting in an acylimine moiety which is thought to increase the π -conjugation in these systems (Gross et al. 2000). Further cyclization in the tail region can also take place in some chromophores, such as in mKO (Kikuchi et al. 2008) and mOrange (Shaner et al. 2004; Shu et al. 2006). When studying the bare chromophore, the connections to the protein are broken and capped with hydrogens or methyl groups. In proteins with no additional maturation step (i.e., GFP-like), the tail is usually removed, and the imidazolinone ring is capped. Methyl capping of the GFP chromophore yields the thoroughly studied HBDI model (Nifosi and Tozzini 2012). Salem and Brown showed that methyl capping of FP chromophores with modified tails, e.g., RFP or mBlueberry (Ai et al. 2007), can be necessary to obtain an optimized structure resembling the one found in the protein (Salem and Brown 2014). Even when only the isolated chromophore is considered, its size renders TDDFT as the method of choice for the majority of TPA studies (see section “Computational Implementation”). For TDDFT studies of FP chromophores, it has been shown

that B3LYP and PBE0 functionals yield S_0 to S_1 excitation energies that are very close to the experimental values in the FP, while both energies and TPA cross sections are systematically larger with CAM-B3LYP and LC-BLYP (Salem and Brown 2014). TPA cross sections computed with the four functionals (with the 6-31+G(d,p) basis set) were significantly lower than the corresponding CC2 values using the same basis set (Salem and Brown 2014). The most sophisticated ab initio methods considered at this level of complexity, i.e., for isolated chromophores, were RI-CC2 (Friese et al. 2012), full CC2 (Salem and Brown 2014), and EOM-CCSD (Nanda and Krylov 2015). At this level of complexity, the absolute values of the TPA cross sections for a given protein determined experimentally cannot be reproduced quantitatively by any available computational method. Therefore, computational studies on the isolated chromophores are useful for initial screening of modified chromophores (i.e., to rank their intrinsic TPA cross section) or to study qualitatively the TPA photophysics for a given chromophore.

To address the second level of complexity, the chromophore needs to be studied in its native conformation dictated by the surrounding protein shell. This can be readily accomplished if the experimental three-dimensional structure is available for the studied FP, i.e., a PDB file is available. To go beyond the single-structure static picture, molecular dynamics (MD) simulations can be utilized to sample accessible chromophore conformations in its protein environment. To date, only one study systematically looked at the variations of TPA with chromophore conformations (see section “Understanding TPA of FPs via Computational Studies”) (Salem and Brown 2015). Only a small number of studies have accounted for the third level of complexity by including the electronic (nonmechanical) effect of the protein environment on its TPA. The polarizable embedding (Olsen et al. 2010) (PE)-DFT/MM scheme allows for mutual polarization between the chromophore in its ground state and the protein (ground-state polarizable embedding). More importantly, PE-DFT accounts for the response of the protein electronic degrees of freedom to the excitation of the chromophore through a response formalism (List et al. 2012). The two available studies that utilized PE-DFT are discussed in the following section.

Understanding TPA of FPs via Computational Studies

Although studying a bare chromophore is the simplest approximation, it provides a very good starting point for understanding the TPA properties of a given protein. Through this first level of complexity, Nifosì and Luo explained the origin of the high-energy absorption peak in the TPA spectrum that has no corresponding OPA peak (Nifosì and Luo 2007b). They showed that this peak can be attributed to excitations to higher electronic states (S_0 to S_n). Further, they used TDDFT at the B3LYP/6-31G+(d) level of theory in the gas phase, to compute TPA for a series of FP chromophore models and predicted the presence of such high-energy peaks in other chromophores as well (Nifosì and Luo 2007a). These predicted absorptions were later validated via experiments (Drobizhev et al. 2011). Computations at this first level of complexity also helped to understand the blueshift of the low-energy TPA peak in anionic FP chromophores. It was first suggested that this shift is due to a bright TPA state that is dark under OPA (Hosoi et al. 2008).

No computation, however, found evidence for such a state (Epifanovsky et al. 2009; Polyakov et al. 2010). Kamarchik and Krylov computed the non-Condon effects in the TPA spectrum of an anionic HBDI model (see Fig. 1 and the discussion in section “Levels of Complexity and Computational Approaches”). Their results showed that an increase in the transition moment associated with certain vibrational modes of the S_1 electronic state causes a preferential vibronic transition. Ai et al. reached a similar conclusion via a CASSCF computation (Ai et al. 2013).

Very recently, Salem and Brown screened 22 hypothetical chromophores derived from noncanonical amino acids (ncAAs) for their (S_0 to S_1) TPA cross sections (Salem and Brown 2015). A chromophore with a nitro group (that replaces the hydroxyl group in Fig. 1) exhibited an exceptionally large TPA cross section relative to other ncAA-derived and natural chromophores as computed at the B3LYP/6-31+G(d,p) level of theory in PCM. Using a two-level model (see section “The Two-Level Model”), they attributed this large intrinsic TPA to both a large transition dipole moment (μ_{01}) and a large difference between permanent dipoles of the excited and ground states ($\Delta\mu$). They accounted for the conformational dependence of the TPA (the second level of complexity) through a MD simulation of a modified EGFP. Results showed that the TPA dependence on conformation is directly related to the change in $\Delta\mu$ in the direction of the C–C=C bond in the plane of the chromophore.

Chakarbarti and Ruud showed that systems with supermolecular π -stacking have enhanced TPA cross sections associated with intermolecular charge-transfer (ICT) transitions (Chakarbarti and Ruud 2009a, b). Interestingly, YFP, a mutant of GFP, has π -stacking between the chromophore and a close-by Tyr residue. Beerepoot et al. studied the TPA of the YFP π -stacking systems (chromophore and close-by residue) via RI-CC2/aug-cc-pVDZ and CAM-B3LYP with both aug-cc-pVDZ and a smaller 6-31+G(d) basis set (Beerepoot et al. 2014). Their results show ICT transitions with significantly enhanced TPA. They further determined the dependence of (part of) the TPA spectrum on the separation between the π -aromatic systems (across which ICT occurs). This work partially accounts for the third level of complexity, that is, the local environment of the chromophore. In a very recent study, Drobizhev et al. evaluated local electric fields in a series of 26 GFP homologues that share the same anionic structure. Their study elegantly compared experimental measurements of OPA and TPA spectra and the deduced total and long-range fields to findings from MD simulations to justify their method of measurements (Drobizhev et al. 2015).

To date, only two QM/MM studies fully accounted for the third level of complexity mentioned in section “Levels of Complexity and Computational Approaches.” The first study by Steindal et al. (at the PE-CAM-B3LYP/6-31+G(d) level of theory) showed that embedding the GFP chromophore in the FP matrix significantly enhances its TPA as compared to the isolated chromophore model (Steindal et al. 2012). On the technical side, they mention the importance of including PE to reproduce experimental spectra. Although they theoretically accounted for (nearly) all

the factors that should affect the TPA of GFP, they failed to reproduce quantitatively the experimentally measured intensities of the TPA spectrum. The need to include PE to correctly account for the protein matrix and the TPA enhancement it produces are confirmed in the second study by List et al. that examined DsRED, a red FP, also at the PE-CAM-B3LYP/6-31+G(d) level of theory. Further, List et al. studied the physical mechanisms leading to the enhancement of TPA cross sections by the electrostatic effects of the protein (List et al. 2012). To do so, they determined the TPA cross sections for the isolated chromophore at the geometry optimized in vacuum and at its conformation in the native protein. By comparing the results of the three computations, they distinguished the effect at different levels of complexity. To account for the importance of PE, they computed the TPA and other parameters with the PE turned off (which they refer to as NPE and is the same as electrostatic embedding) and with PE frozen at the ground-state value only (FPE). Through a plane-projected two-level model that accounts for the angle between dipoles, they noticed that the increased TPA cross section is related to an enlarged $\Delta\mu$ (the difference between permanent dipoles of the ground and excited states) and its closer alignment with μ_{01} (the transition dipole moment). According to their results, the ground-state dipole is actually larger than the excited-state one, so increasing the former and decreasing the latter increases $\Delta\mu$. Computation of the ground- and excited-state permanent dipoles using different levels of theory showed that the ground-state polarization (FPE) causes enhancement of both permanent dipoles with the ground state being slightly more enhanced. Further inclusion of full PE cancels the enhancement of the excited-state dipole which increases $\Delta\mu$ and also decreases the angle between $\Delta\mu$ and μ_{01} leading to the TPA enhancement. Evaluation of the individual contribution of each amino acid showed that the largest contribution to TPA enhancement comes from the positively charged Lys163 which is also mainly responsible of the blueshift of the OPA spectrum. They finally suggested a mutation that should improve the TPA intensity and at the same time redshift the OPA spectrum.

Although the absolute TPA cross sections computed by List et al. (2012) match the experiment (perhaps somewhat fortuitously), there are still open questions. Clearly, more PE-DFT/MM studies need to be carried out before there is confidence that we can successfully reproduce experimental TPA cross sections, as Steindal et al. (2012) failed to reproduce them although they used the same method that should account for the full environmental effects. The conclusions of List et al. regarding the direction of $\Delta\mu$ mediation by the protein (List et al. 2012) contradict those by Drobizhev et al. (2009). As pointed out by Beerepoot et al. (2015), there is a need to evaluate the inherent error associated with CAM-B3LYP through a benchmark against a higher-level, more accurate method. More experiments need to be undertaken to test the tuning suggestions presented by some of the computational studies mentioned above. Finally, FPs for the promising chromophores made from noncanonical amino acids (as computed by Salem and Brown 2015) need to be experimentally synthesized and tested.

Nucleic Acid Base Analogues

The basic building blocks of the nucleic acids DNA and RNA are the five nitrogenous bases: adenine, cytosine, guanine, thymine, and uracil. The ultrashort excited-state lifetimes of these bases grant them photostability which is crucial for the protection of the nucleic acid to photo-initiated processes. As a result of their non-emissive nature, the natural nucleobases cannot be utilized as fluorescent probes. Modifications to the structure of the nucleobases can yield molecules with different photophysical characteristics such as enhancement of their fluorescence. In the design of fluorescent base analogues, it is desirable that analogues exhibit sensitivity to their local environment while having minimal impact on the overall nucleic acid structure. The utilization of these “isomorphic” analogues is advantageous as they offer minimal perturbation to the helix, whereas bulky dyes can disrupt interactions necessary for nucleic acid functionality. Experiments utilizing fluorescent DNA/RNA nucleotide analogues have provided an effective way to study conformational dynamics of nucleic acids over a range of timescales (Alemán et al. 2014; Song and Mandal 2014; Ha 2004; Xavier Michalet and Jäger 2006). For example, studies in the millisecond to second timescale provide information about protein-DNA interactions, while fluorescence anisotropy can provide structural characteristics on the picosecond to nanosecond timescale (Katilius and Woodbury 2006). Fluorescence microscopy at the single-molecule level has emerged as a powerful technique to investigate conformational dynamics of DNA (Pitchaiya et al. 2014). Challenges such as photobleaching arise with excitation in the UV range for many fluorescent base analogues. Investigating the photophysical properties and understanding how they are related to their structure are crucial for the design of more effective analogues. We refer the reader to recent, comprehensive discussions of modified nucleobases by Sinkeldam et al. (2010) and Matsika (2015). The most common challenges in the design of fluorescent nucleobase analogues for biological applications are the following (and these echo the general challenges outlined in the introduction): (i) The absorption cross sections of nucleobase analogues are much lower compared to the extrinsic dyes, and hence their brightness is insufficient for accurate imaging. (ii) Most analogues absorbing in the UV region can be prone to photobleaching. (iii) The penetration of light in biological tissues is poor at these wavelengths due to scattering processes. An alternative approach is to use multiphoton excitation in the near-IR range. This change in absorption wavelengths enables the analogues to be utilized as probes for applications requiring deep tissue penetration and can overcome the problems arising from photobleaching.

While several experiments have investigated the two-photon absorption properties of nucleobase analogues, corresponding computational studies are far more limited; to our knowledge there is only one. Nearly a decade ago, the two-photon cross sections of the guanine analogue 6-methylisoxanthopterin (6-MI) and the adenine analogue 4-amino-6-methyl-8-(2'-deoxy- β -D-ribofuranosyl)-7(8H)-pteridone (6-MAP) were measured to be 0.8 GM and 3.4 GM, respectively (Katilius and Woodbury 2006; Stanley et al. 2005). Recently, Lane and Magennis determined the

σ_{TPA} of 2-aminopurine (2-AP), as well as tricyclic cytosine (tC), as 0.2 and 1.5 GM, respectively (Lane and Magennis 2012). Although these analogues demonstrate resistance to photobleaching, their brightness is not sufficient for single-molecule detection. Recently, a set of isomorphous analogues synthesized by Tor et al. have been investigated for two-photon induced fluorescence (Lane et al. 2014). The structures of these modified nucleosides are depicted in Fig. 2.

Computational Investigations of Nucleoside Analogues

As with the study of fluorescent proteins, we can investigate the properties of nucleic acid analogues in varying levels of complexity. The simplest approach (i.e., first level of complexity) is to determine the excitation properties of the isolated nucleobase which is the chromophoric moiety in the molecule. We can build upon that by including the ribose or deoxyribose (nucleoside), and, furthermore, we can investigate the analogues as a nucleotide or sequence of nucleotides (oligonucleotide). Of course in all of these cases, one would ideally want to include the effects of solvent, although often this can be accomplished through the use of polarizable continuum models (PCM). In some cases, the first level of complexity is sufficient for predicting the desired properties, and inclusion of the sugar moiety (and phosphate group) does not alter the findings. An extensive analysis of the one-photon excitation properties of the thiophene analogues (thA, thU, thG, and thC) by Gedik and Brown (2013) demonstrated this effect. For TPA, it is expected (yet remains to be confirmed) that the sugar group should have a more drastic effect on the cross sections since the orientation/geometry of the ribose will alter the permanent dipole moments (ground and excited state) of the molecule thus affecting $\Delta\mu$. Another factor that needs to be considered in the determination of TPA is the possible role of tautomerization. For example, it has been established experimentally that the tC analogue exists as a mixture of amino and imino tautomers (Lane and Magennis 2012). In order to obtain accurate theoretical data, both tautomers would need to be examined.

To our knowledge, the only theoretical investigation of the two-photon excitation properties of nucleic acid analogues was performed by Samanta and Pati (2015). In this study, the TPA cross sections of the set of analogues depicted in Fig. 2 were determined at the B3LYP/6-31++G(d,p) level of theory. The computations were performed both in the gas phase and in solvent (implicit solvation with PCM). The OPA parameters have been determined by Samanta et al. in a previous study (Samanta et al. 2012). A summary of their findings is presented in Table 1.

The results from Table 1 reveal that small alterations in the structure of the nucleosides can greatly impact their photophysical properties. The analogues display a very large variation in their TPA cross sections ranging from approximately 4 to 73 GM (computational determination) or 0.17 to 7.6 GM (experimental measurement) in water. From the computational results, the thiophene and furan analogues display the lowest cross sections, while the quinazoline analogue shows the highest σ_{TPA} in water. This trend is not observed in the experimental findings. Both experiment and theory ascertain the $S_0 \rightarrow S_1$ transition as both OPA and TPA

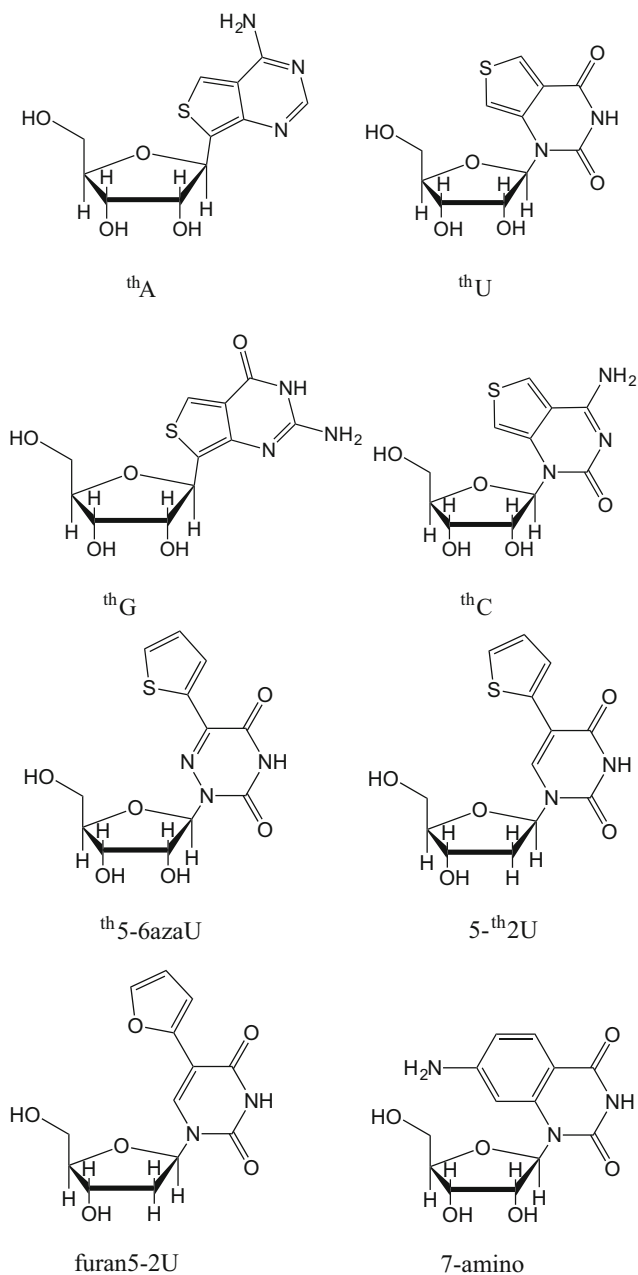


Fig. 2 Structures of the modified nucleosides as designed by Lane et al. (2014).

thA:thieno[3,4-d]adenosine, thU:thieno[3,4-d]uridine, thG:thieno[3,4-d]guanosine, thC:thieno[3,4-d]cytidine, th5-6azaU:5-(thiophen-2-yl)-6-aza-uridine, 5-th2U:5-(thiophen-2-yl)-2'-deoxyuridine, **furan5-2U**:5-(furan-2-yl)-2'-deoxyuridine, **7-amino**:7-amino-1-ribose-quinazoline-2,4(1H,3H)-dione

Table 1 Two-photon absorption wavelengths and cross sections in GM of the modified uridine analogues. Experimental findings from Lane et al. (2014) and computations at the B3LYP/6-31+G(d,p) level of theory in the gas phase and PCM water by Samanta and Pati (2015)

Analogue	Gas		Water		Experimental	
	E(nm)	σ_{TPA}	E(nm)	σ_{TPA}	E(nm)	σ_{TPA}
th A	681	7.45	711	3.98	NR	NR
th C	678	4.78	656	11.32	NR	NR
th G	642	5.54	714	16.90	NR	NR
th U	595	3.74	618	12.82	690	0.17
th 5-6azaU	713	7.08	751	31.30	690	3.8
					740	0.81
5- th 2U	670	4.55	681	18.23	690	7.6
					740	0.33
Furan5-2U	662	11.36	661	11.70	690	2.1
					740	0.18
7-Amino	563	3.72	605	72.65	690	1.8
					740	0.34

NR = not reported

active for all of the analogues. It is evident that more theoretical studies need to be performed (i) to test other functionals and/or higher-level ab initio methods, (ii) to elucidate the effect of small structural changes on the TPA cross sections, and (iii) to examine the effects of conformational flexibility, e.g., the orientation of the sugar group mentioned previously. Further investigation of two-photon absorption of modified nucleobases by computational approaches is crucial for shedding light on the structure-photophysics interplay, which will ultimately aid in the design of better analogues.

Conclusions

In this chapter, we presented a brief discussion of the theory of TPA. With a focus on TPA in biological molecules, we reviewed the available computational methods and software packages. We then presented the applications of TPA utilizing fluorescent proteins and emissive nucleic acid base analogues. In each case, we gave a brief introduction to the topic followed by a review of the computational studies carried out on these two classes of molecules. We concluded each section with highlights of areas that require further computational investigation. What is clear is that there is a bright future for the computational study of TPA with applications to biological imaging.

Bibliography

- Adamo, C., & Barone, V. (1999). Toward reliable density functional methods without adjustable parameters: The PBE0 model. *Journal of Chemical Physics*, *110*, 6158.
- Ai, H. W., Shaner, N. C., Cheng, Z., Tsien, R. Y., & Campbell, R. E. (2007). *Biochemistry*, *46*, 5904.

- Ai, Y., Tian, G., & Luo, Y. (2013). *Molecular Physics*, *111*, 1316.
- Aidas, K., Angeli, C., Bak, K. L., Bakken, V., Bast, R., Boman, L., Christiansen, O., Cimiraglia, R., Coriani, S., Dahle, P., Dalskov, E. K., Ekström, U., Enevoldsen, T., Eriksen, J. J., Ettenhuber, P., Fernández, B., Ferrighi, L., Fliegl, H., Frediani, L., Hald, K., Halkier, A., Hättig, C., Heiberg, H., Helgaker, T., Hennum, A. C., Hettner, H., Hjertens, E., Hst, S., Hyvik, I.-M., Iozzi, M. F., Jansik, B., Jensen, H. J. Aa., Jonsson, D., Jørgensen, P., Kauczor, J., Kirpekar, S., Kjrgaard, T., Klopper, W., Knecht, S., Kobayashi, R., Koch, H., Kongsted, J., Krapp, A., Kristensen, K., Ligabue, A., Lutns, O. B., Melo, J. I., Mikkelsen, K. V., Myhre, R. H., Neiss, C., Nielsen, C. B., Norman, P., Olsen, J., Olsen, J. M. H., Osted, A., Packer, M. J., Pawłowski, F., Pedersen, T. B., Provasi, P. F., Reine, S., Rinkevicius, Z., Ruden, T. A., Ruud, K., Rybkin, V., Salek, P., Samson, C. C. M., Snchez de Mers, A., Saue, T., Sauer, S. P. A., Schimmelpennig, B., Sneskov, K., Steindal, A. H., Sylvester-Hvid, K. O., Taylor, P. R., Teale, A. M., Tellgren, E. I., Tew, D. P., Thorvaldsen, A. J., Thgersen, L., Vahtras, O., Watson, M. A., Wilson, D. J. D., Ziolkowski, M., & Ågren, H. (2014). The Dalton quantum chemistry program system. *WIREs Computational Molecular Science*, *4*, 269–284.
- Alemán, E. A., de Silva, C., Patrick, E. M., Musier-Forsyth, K., & Rueda, D. (2014). *Journal of Physical Chemistry Letters*, *5*, 777.
- Bartlett, R. J. (2010). *Molecular Physics*, *108*, 2905.
- Becke, A. D. (1988). *Physical Review A*, *38*, 3098.
- Becke, A. D. (1993). *Journal of Chemical Physics*, *98*, 5648.
- Beerepoot, M. T. P., Friese, D. H., & Ruud, K. (2014). *Physical Chemistry Chemical Physics*, *16*, 5958.
- Beerepoot, M. T. P., Friese, D. H., List, N. H., Kongsted, J., & Ruud, K. (2015). *Physical Chemistry Chemical Physics*, *17*, 19306.
- Campbell, R. E., & Davidson, M. W. (2010). In *Molecular Imaging with Reporter Genes* (pp. 3–40). Cambridge University Press.
- Chakrabarti, S., & Ruud, K. (2009a). *Physical Chemistry Chemical Physics*, *11*, 2592.
- Chakrabarti, S., & Ruud, K. (2009b). *Journal of Physical Chemistry A*, *113*, 5485.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., & Prasher, D. (1994). *Science*, *263*, 802.
- Christiansen, O., Koch, H., & Jørgensen, P. (1995a). *Journal of Chemical Physics*, *103*, 7429.
- Christiansen, O., Koch, H., & Jørgensen, P. (1995b). *Chemical Physics Letters*, *243*, 409.
- Ciuciu, A. I., Firmansyah, D., Hugues, V., Blanchard-Desce, M., Gryko, D. T., & Flamigni, L. (2014). *Journal of Materials Chemistry C*, *2*, 4552.
- Cronstrand, P., Luo, Y., & Ågren, H. (2005). *Advances in Quantum Chemistry*, *50*, 1.
- Denk, W., Strickler, J., & Webb, W. (1990). *Science*, *248*, 73.
- Drobizhev, M., Tillo, S., Makarov, N. S., Hughes, T. E., & Rebane, A. (2009). *Journal of Physical Chemistry B*, *113*, 12860.
- Drobizhev, M., Makarov, N. S., Tillo, S. E., Hughes, T. E., & Rebane, A. (2011). *Nature Methods*, *8*, 393.
- Drobizhev, M., Callis, P. R., Nifosì, R., Wicks, G., Stoltzfus, C., Barnett, L., Hughes, T. E., Sullivan, P., & Rebane, A. (2015). *Science Reports*, *5*, 13223.
- Epifanovsky, E., Polyakov, I., Grigorenko, B., Nemukhin, A., & Krylov, A. I. (2009). *Journal of Chemical Theory and Computation*, *5*, 1895.
- Frediani, L., Rinkevicius, Z., & Ågren, H. (2005). *Journal of Chemical Physics*, *122*, 244104.
- Friese, D. H., Hättig, C., & Ruud, K. (2012). *Physical Chemistry Chemical Physics*, *14*, 1175.
- Gedik, M., & Brown, A. (2013). *Journal of Photochemistry and Photobiology A: Chemistry A*, *259*, 25.
- Goppert-Mayer, M. (1931). *Annals of Physics*, *9*, 273.
- Gross, L. A., Baird, G. S., Hoffman, R. C., Baldrige, K. K., & Tsien, R. Y. (2000). *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 11990.
- Ha, T. (2004). *Biochemistry*, *43*, 4055.
- Hättig, C., Christiansen, O., & Jørgensen, P. (1998). *Journal of Chemical Physics*, *108*, 8331.

- Hosoi, H., Yamaguchi, S., Mizuno, H., Miyawaki, A., & Tahara, T. (2008). *Journal of Physical Chemistry B*, *112*, 2761.
- Iikura, H., Tsuneda, T., Yanai, T., & Hirao, K. (2001). *Journal of Chemical Physics*, *115*, 3540.
- Inoué, S., & Tsuji, F. I. (1994). *FEBS Letters*, *341*, 277.
- Kaiser, W., & Garrett, C. G. B. (1961). *Physical Review Letters*, *7*, 229.
- Kamarchik, E., & Krylov, A. I. (2011). *Journal of Physical Chemistry Letters*, *2*, 488.
- Katilius, E., & Woodbury, N. W. (2006). *Journal of Biomedical Optics*, *11*, 044004.
- Kawano, H., Kogure, T., Abe, Y., Mizuno, H., & Miyawaki, A. (2008). *Nature Methods*, *5*, 373.
- Kikuchi, A., Fukumura, E., Karasawa, S., Mizuno, H., Miyawaki, A., & Shiro, Y. (2008). *Biochemistry*, *47*, 11573.
- Koch, H., Christiansen, O., Jørgensen, P., Sanchez de Merás, A. M., & Helgaker, T. (1997). *Journal of Chemical Physics*, *106*, 1808.
- Kristensen, K., Kauczor, J., Thorvaldsen, A. J., Jørgensen, P., Kjærgaard, T., & Rizzo, A. (2011). *Journal of Chemical Physics*, *134*, 214104.
- Krylov, A. I., & Gill, P. M. (2013). *WIREs Computational Molecular Science*, *3*, 317.
- Lane, R. S. K., & Magennis, S. W. (2012). *RSC Advances*, *2*, 11397.
- Lane, R. S. K., Jones, R., Sinkeldam, R. W., Tor, Y., & Magennis, S. W. (2014). *ChemPhysChem*, *15*, 867.
- Lee, C., Yang, W., & Parr, R. G. (1988). *Physical Review B*, *37*, 785.
- List, N. H., Olsen, J. M. H., Jensen, H. J. A., Steindal, A. H., & Kongsted, J. (2012). *Journal of Physical Chemistry Letters*, *3*, 3513.
- Matsika, S. (2015). *Topics in Current Chemistry*, *355*, 209.
- Meath, W. J., & Jagatap, B. N. (2011). *Journal of Physics B*, *44*, 205401.
- Meath, W. J., & Jagatap, B. N. (2013). *Journal of Chemical Physics*, *139*, 144104.
- Meath, W. J., Jagatap, B. N., & Kondo, A. E. (2006). *Journal of Physics B*, *39*, S605.
- Morales, A. R., Frazer, A., Woodward, A. W., Ahn-White, H. Y., Fonari, A., Tongwa, P., Timofeeva, T., & Belfield, K. D. (2013). *Journal of Organic Chemistry*, *78*, 1014.
- Nanda, K. D., & Krylov, A. I. (2015). *Journal of Chemical Physics*, *142*, 064118.
- Nayyar, I. H., & Tretiak, S. (2013). *Journal of Physical Chemistry C*, *117*, 18170.
- Nifosì, R., & Luo, Y. (2007a). *Journal of Physical Chemistry B*, *111*, 14043.
- Nifosì, R., & Luo, Y. (2007b). *Journal of Physical Chemistry B*, *111*, 505.
- Nifosì, R., & Tozzini, V. (2012). In G. Jung (Ed.), *Fluorescent proteins I* (Springer series on fluorescence, Vol. 11, pp. 3–40). Berlin/Heidelberg: Springer.
- Olsen, J., & Jørgensen, P. (1985). *Journal of Chemical Physics*, *82*, 3235.
- Olsen, J. M., Aidas, K., & Kongsted, J. (2010). *Journal of Chemical Theory and Computation*, *6*, 3721.
- Ormo, M., Cubitt, A. B., Kallio, K., Gross, L. A., Tsien, R. Y., & Remington, S. (1996). *Science*, *273*, 1392.
- Oulianov, D., Tomov, I., Dvornikov, A., & Rentzepis, P. (2001). *Optics Communication*, *191*, 235.
- Paterson, M. J., Christiansen, O., Pawłowski, F., Jørgensen, P., Hättig, C., Helgaker, T., & Sałek, P. (2006). *Journal of Chemical Physics*, *124*, 054322.
- Pawlicki, M., Collins, H., Denning, R., & Anderson, H. (2009). *Angewandte Chemie International Edition*, *48*, 3244.
- Pitchiaya, S., Heinicke, L. A., Custer, T. C., & Walter, N. G. (2014). *Chemistry Reviews*, *114*, 3224.
- Polyakov, I. V., Grigorenko, B. L., Epifanovsky, E. M., Krylov, A. I., & Nemukhin, A. V. (2010). *Journal of Chemical Theory and Computation*, *6*, 2377.
- Prasher, D. C., Eckenrode, V. K., Ward, W. W., Prendergast, F. G., & Cormier, M. J. (1992). *Gene*, *111*, 229.
- Runge, E., & Gross, E. K. U. (1984). *Physical Review Letters*, *52*, 997.
- Sašek, P., Vahtras, O., Guo, J., Luo, Y., Helgaker, T., & Ågren, H. (2003). *Chemical Physical Letters*, *374*, 446.
- Salem, M. A., & Brown, A. (2014). *Journal of Chemical Theory and Computation*, *10*, 3260.

- Salem, M. A., & Brown, A. (2015). *Physical Chemistry Chemical Physics*, 17, 25563.
- Samanta, P. K., & Pati, S. K. (2015). *Physical Chemistry Chemical Physics*, 17, 10053.
- Samanta, P. K., Manna, A. K., & Pati, S. K. (2012). *Journal of Physical Chemistry B*, 116, 7618.
- Schmidt, M. W., Baldrige, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S. J., Windus, T. L., Dupuis, M., & Montgomery, J. A. (1993). *Journal of Computational Chemistry*, 14, 1347.
- Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E., & Tsien, R. Y. (2004). *Nature Biotechnology*, 22, 1567.
- Shaner, N. C., Steinbach, P. A., & Tsien, R. Y. (2005). *Nature Methods*, 2, 905.
- Shao, Y., Gan, Z., Epifanovsky, E., Gilbert, A. T., Wormit, M., Kussmann, J., Lange, A. W., Behn, A., Deng, J., Feng, X., Ghosh, D., Goldey, M., Horn, P. R., Jacobson, L. D., Kaliman, I., Khaliullin, R. Z., Kuš, T., Landau, A., Liu, J., Proynov, E. I., Rhee, Y. M., Richard, R. M., Rohrdanz, M. A., Steele, R. P., Sundstrom, E. J., Woodcock, H. L., III, Zimmerman, P. M., Zuev, D., Albrecht, B., Alguire, E., Austin, B., Beran, G. J. O., Bernard, Y. A., Berquist, E., Brandhorst, K., Bravaya, K. B., Brown, S. T., Casanova, D., Chang, C. M., Chen, Y., Chien, S. H., Closser, K. D., Crittenden, D. L., Diedenhofen, M., DiStasio, R. A., Jr., Do, H., Dutoi, A. D., Edgar, R. G., Fatehi, S., Fusti-Molnar, L., Ghysels, A., Golubeva-Zadorozhnaya, A., Gomes, J., Hanson-Heine, M.W., Harbach, P. H., Hauser, A. W., Hohenstein, E. G., Holden, Z. C., Jagau, T. C., Ji, H., Kaduk, B., Khistyayev, K., Kim, J., Kim, J., King, R. A., Klunzinger, P., Kosenkov, D., Kowalczyk, T., Krauter, C. M., Lao, K. U., Laurent, A. D., Lawler, K. V., Levchenko, S. V., Lin, C. Y., Liu, F., Livshits, E., Lochan, R. C., Luenser, A., Manohar, P., Manzer, S. F., Mao, S. P., Mardirossian, N., Marenich, A. V., Maurer, S. A., Mayhall, N. J., Neuscammen, E., Oana, C. M., Olivares-Amaya, R., O'Neill, D. P., Parkhill, J. A., Perrine, T. M., Peverati, R., Prociuk, A., Rehn, D. R., Rosta, E., Russ, N. J., Sharada, S. M., Sharma, S., Small, D. W., Sodt, A., Stein, T., Stück, D., Su, Y. C., Thom, A. J., Tsuchimochi, T., Vanovschi, V., Vogt, L., Vydrov, O., Wang, T., Watson, M. A., Wenzel, J., White, A., Williams, C. F., Yang, J., Yeganeh, S., Yost, S. R., You, Z. Q., Zhang, I. Y., Zhang, X., Zhao, Y., Brooks, B. R., Chan, G. K., Chipman, D. M., Cramer, C. J., Goddard, W. A., III, Gordon, M. S., Hehre, W. J., Klamt, A., Schaefer, H. F., III, Schmidt, M. W., Sherrill, C. D., Truhlar, D. G., Warshel, A., Xu, X., A. Aspuru-Guzik, Baer, R., Bell, A. T., Besley, N. A., Chai, J. D., Dreuw, A., Dunietz, B. D., Furlani, T. R., Gwaltney, S. R., Hsu, C. P., Jung, Y., Kong, J., Lambrecht, D. S., Liang, W., Ochsenfeld, C., Rassolov, V. A., Slipchenko, L. V., Subotnik, J. E., Voorhis, T. V., Herbert, J. M., Krylov, A. I., Gill, P. M., & Head-Gordon, M. (2015). *Molecular Physics*, 113(2), 184.
- Shimomura, O., Johnson, F. H., & Saiga, Y. (1962). *Journal of Cellular and Comparative Physiology*, 59, 223.
- Shu, X., Shaner, N. C., Yarbrough, C. A., Tsien, R. Y., & Remington, S. J. (2006). *Biochemistry*, 45, 9639.
- Silva, D. L., Krawczyk, P., Bartkowiak, W., & Mendonça, C. R. (2009). *Journal of Chemical Physics*, 131, 244516.
- Sinkeldam, R. W., Greco, N. J., & Tor, Y. (2010). *Chemistry Reviews*, 110, 2579.
- Song, G., & Mandal, M. (2014). In *RNA Nanotechnology* (pp. 153–184). Pan Stanford.
- Stanley, R. Y. J., Zhanjia, Y., & Hawkins, M. E. (2005). *Journal of Physical Chemistry B*, 109, 3690.
- Stanton, J. F., & Bartlett, R. J. (1993). *Journal of Chemical Physics*, 98, 7029.
- Steindal, A. H., Olsen, J. M. H., Ruud, K., Frediani, L., & Kongsted, J. (2012). *Physical Chemistry Chemical Physics*, 14, 5440.
- Tawada, Y., Tsuneda, T., Yanagisawa, S., Yanai, T., & Hirao, K. (2004). *Journal of Chemical Physics*, 120, 8425.
- Terenziani, F., Katan, C., Badaeva, E., Tretiak, S., & Blanchard-Desce, M. (2008). *Advanced Materials*, 20, 4641.
- Tretiak, S., & Chernyak, V. (2003). *Journal of Chemical Physics*, 119, 8809.
- TURBOMOLE V7.0 2015, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.

- Vivas, M. G., Silva, D. L., Misoguti, L., Zaleśny, R., Bartkowiak, W., & Mendonca, C. R. (2010). *Journal of Physical Chemistry A*, *114*, 3466.
- Wachter, R. M., Elsliger, M. A., Kallio, K., Hanson, G. T., & Remington, S. J. (1998). *Structure*, *6*, 1267.
- Weissleder, R. (2001). *Nature Biotechnology*, *19*, 316.
- Xavier Michalet, S. W., & Jäger, M. (2006). *Chemistry Reviews*, *106*, 1785.
- Yanai, T., Tew, D. P., & Handy, N. C. (2004). A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chemical Physics Letters*, *393*, 51.
- Yang, F., Moss, L. G., & Phillips, G. (1996). *Nature Biotechnology*, *14*, 1246.
- Yang, G., Qin, C., Su, Z., & Dong, S. (2005). *Journal of Molecular Structure: THEOCHEM*, *726*, 61.
- Yuan, L., Lin, W., Chen, H., Zhu, S., & He, L. (2013). *Angewandte Chemie International Edition*, *52*, 10018.
- Zhang, G., Gurtu, V., & Kain, S. R. (1996). *Biochemical and Biophysical Research Communications*, *227*, 707.

Consequences of Electron Attachment to Modified Nucleosides Incorporated into DNA

44

Lidia Chomicz-Mańka, Paweł Wityk, Łukasz Golon, Magdalena Zdrowowicz, Justyna Wiczek, Kinga Westphal, Michał Żyndul, Samanta Makurat, and Janusz Rak

Contents

Introduction	1896
Where Do the Electrons Come From?	1897
Electron Attachment and Anion Radical Stability	1899
Electron Impact on Native DNA	1901
Why Use Modified Nucleosides of High Electron Affinity?	1903
Dehalogenation of Modified Nucleobases	1904
What Next with the Nucleobase Radicals Produced?	1907
Quantum Chemistry as a Tool to Search for New Therapeutics	1909
Summary	1912
Bibliography	1913

Abstract

Radiotherapy is still one of the most common modalities employed in anticancer treatment. However, hypoxia, a hallmark of solid tumor cells, makes cancer significantly more resistant than the adjacent healthy tissues toward ionizing radiation. In order to overcome this detrimental situation, radiosensitizers, i.e., substances that lower the lethal dose of ionizing radiation, should be used in clinics concomitantly with radiotherapy. In the current review, one class of radiosensitizers – modified nucleosides – are discussed in terms of their mode of action. As radiosensitizers, these modified nucleosides must employ a reaction with hydrated electrons that leaves behind the reactive radicals of nucleobases in

L. Chomicz-Mańka (✉) • P. Wityk • Ł. Golon • M. Zdrowowicz • J. Wiczek • K. Westphal • M. Żyndul • S. Makurat • J. Rak
Biological Sensitizers Laboratory, Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland
e-mail: lidia.chomicz@ug.edu.pl; pawel.wityk@phdstud.ug.edu.pl;
lukasz.golon@phdstud.ug.edu.pl; magdalena.zdrowowicz@phdstud.ug.edu.pl;
justyna.wiczek@phdstud.ug.edu.pl; kinga.westphal@phdstud.ug.edu.pl;
michal.zyndul@phdstud.ug.edu.pl; samantamakurat@gmail.com; janusz.rak@ug.edu.pl

DNA. The latter species, in turn, lead to serious DNA damage and the electron affinity of nucleosides and their ease to undergo dissociative electron attachment decide their efficiency. The mechanistic details of the process beginning from electron attachment to a nucleobase and ending with DNA damage are analyzed with the help of molecular modeling at the quantum chemistry (QM) level. First, the mechanism of action of ionizing radiation on biological matter, with particular emphasis on the formation of solvated electrons ubiquitously created under hypoxic condition, is discussed. Then, electron attachment to the modified nucleosides, to their halogen derivatives in particular, coupled to radical processes resulting in DNA damage are described in terms of the results of QM calculations. Finally, the perspectives of computer-based design of electrophilic nucleoside radiosensitizers are shortly discussed.

Introduction

The average life expectancy continuously rises due to the enormous progress in medicine observed in the last decades. Thus, the aging population is one of the reasons that increase the contribution of cancer to the variety of diseases that concern contemporary society. Indeed, as indicated in the review article published by Jemal et al. in 2011 (Jemal et al. 2011): “Cancer is the leading cause of death in economically developed countries and the second leading cause of death in developing countries.” It is therefore quite probable that, in the following years, we will witness a spectacular increase of the global burden of cancer.

Radiotherapy belongs to the most common anticancer therapies (Gaba 2011). Over 50 % of cancer patients receive radiotherapy at certain stages of their treatment. It is, however, worth emphasizing that most solid tumor cells are hypoxic, which makes them resistant toward the action of ionizing radiation (Daşu and Denekamp 1998; Oronsky et al. 2011; Williams et al. 2001). As a result, the tissues adjacent to tumor tissues that are exposed to the effects of ionizing radiation may develop secondary tumors – the most dangerous side effect of radiotherapy. This situation calls for the employment of **radiosensitizers**, i.e., chemical molecules that, when introduced into cancer cells, enable the lethal dose of ionizing radiation to be significantly reduced. Hence, the employment of radiosensitizers should make radiotherapy much more effective and safer toward healthy, nonhypoxic tissues.

The above mentioned oxygen effect – cancer cells with a decreased level of oxygen are two- to threefold less sensitive to high-energy radiation than normal cells (Oronsky et al. 2011) – provides the basis for “Trojan horse” radiotherapy. Developing such a therapy, one might employ the fact that in hypoxic cells, the DNA damaging activity of hydroxyl radicals, a common product of water radiolysis and the main cytotoxic factor under oxic conditions, is highly impaired. On the other hand, under hypoxia, the gray (Gy) of the absorbed radiation produces solvated electrons to an amount equal to the number of hydroxyl radicals (von Sonntag 2006). However, in water, the former species are not reactive toward DNA. It thus seems that a way to efficient radiotherapy under hypoxia is to “activate” solvated

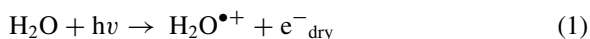
electrons so that they become a DNA damaging factor. Modified nucleosides could be a solid foundation for such an anticancer therapy. After transfer to the cellular cytoplasm, the modified nucleosides must be phosphorylated by kinases to form the appropriate triphosphates (TPs). Then, if the TPs are accepted by DNA polymerases, the modified nucleosides will be introduced into the cellular DNA during replication or repair. 5-Bromouracil or 5-iodouracil, which as valence anions dissociate in an almost barrier-less manner, producing the reactive uracil-5-yl radical, are exemplar nucleosides for Trojan horse radiotherapy (Rak et al. 2015). Indeed, DNA labeled with electron affinic nucleosides that are prone to dissociative electron attachment should be, in contrast to unmodified DNA, sensitive to solvated electrons.

The remaining part of the chapter is focused on the computational approach to DNA radiosensitization with modified electron affinic nucleosides. In the next section, the formation of hydrated electrons, as they are formed during radiotherapy, is shortly characterized. Then, the process of electron attachment to a molecule, in terms of vertical and adiabatic electron affinity, as well as the stability of the formed anion in terms of detachment energy, are defined. The subsequent section is devoted to electron attachment to native DNA. Next, the Trojan horse strategy based on modified nucleosides is described in detail and the particular role of halogenated nucleosides in the development of such an approach is highlighted. Finally, the possible direction for the computational design of future, nucleoside-based radiosensitizers is briefly discussed.

Where Do the Electrons Come From?

Since cells are mostly composed of water, the main effect of their exposure to ionizing radiation (e.g., in radiotherapy treatment) is **water radiolysis**. Radiolysis of bulk water leads to the formation of a few primary species that are formed along the radiation track: genotoxic hydroxyl radicals (OH^\bullet), **hydrated electrons** (e^-_{sol}), hydrogen atoms (H^\bullet), and less harmful species such as hydrogen peroxide (H_2O_2), H_3O^+ , H_2 , and O_2 (von Sonntag 2006). The products of water radiolysis are generated in the following sequence of events:

- (i) High-energy photon (e.g., γ -rays or X-rays; photon energy ~ 1 MeV) hits a water molecule leading to its ionization ($\text{H}_2\text{O}^{\bullet+}$) or excitation (H_2O^*). “Dry” high-energy electrons are released within 1 fs (Eq. 1),



- (ii) During the physicochemical stage (fs-ps), the high-energy electron may ionize the surrounding environment – along the radiation track – creating more free electrons and producing more reactive species (OH^\bullet , H^\bullet , H_2O_2).
- (iii) Finally, high-energy electrons become thermalized, which takes place after 1 ps after their formation (Fig. 1). In contrast to UV radiation, one can observe isolated spurs (places where the concentration of radiolysis products are very

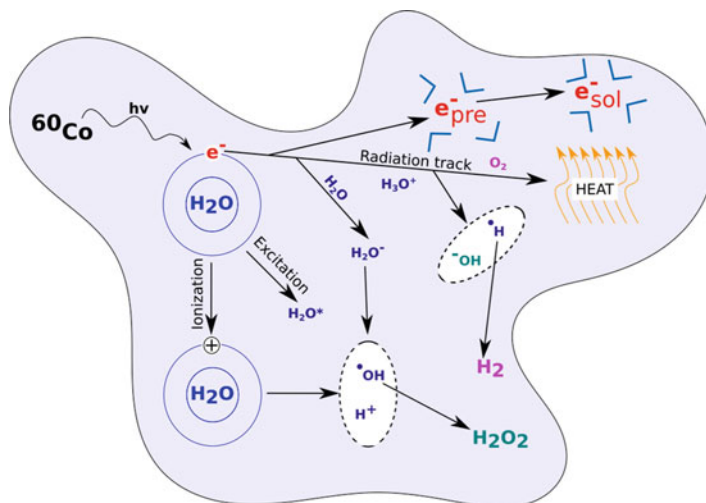
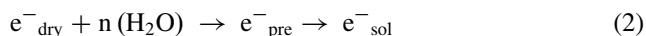


Fig. 1 Schematic representation of bulk water radiolysis induced by γ -rays

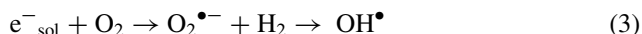
high – along the radiation track) during the irradiation of bulk water. After the physicochemical stage, the chemical stage (ps– μ s) occurs and in this time the products of water radiolysis expand from spurs to solute via diffusion. It seems important to present some of the physical properties of hydrated electron (e^-_{sol}). It is the smallest possible anion with a maximum absorbance at 720 nm, whose lifetime is on the order of microseconds in liquid water and in the order of days in ammonia (Abel et al. 2012).

It is worth emphasizing that e^-_{dry} becomes surrounded by weakly bound water molecules on the fs time scale (from -1.5 to -1 eV, Abel et al. 2012) creating **prehydrated electrons** (lifetime < 1 ps) as the precursor of hydrated (solvated) electron (Eq. 2).



Solvated electrons are bound to water in about 1 ps (Le Caer 2011) with binding energies of 3.4 eV for internally solvated electrons and 1.6 eV for surface-type electrons (Abel et al. 2012). For many years, there has been an ongoing debate about what solvated electron looks like. Kumar et al. proposed recently the minimal model of a cavity surrounded by four water molecules, inside which the electron is localized (Kumar et al. 2015). This interaction disrupts the hydrogen bonding pattern inside the liquid water. Moreover, the calculated properties of this suggested model are in good agreement with experimental data. However, this representation is proposed only in pure water, while in more biological solutions, clusters of electrons with, e.g., sodium ion $(H_2O)_n^-Na^+$ are observed, manifesting a huge impact on the electron stability and properties (Abel et al. 2012).

The most abundant and reactive products of water radiolysis are hydroxyl radicals (OH^\bullet) and solvated electrons (e^-_{sol}), which are formed almost in equal molecular yields under **hypoxic conditions** (von Sonntag 2006). However, we should rather consider the situation in the presence of molecular oxygen, as biological processes occur in an oxygenated environment. Free electrons react with oxygen leading to an increased production of hydroxyl radicals through the intermediate, superoxide, radical state ($\text{O}_2^{\bullet-}$) – the precursor of highly reactive OH^\bullet (Eq. 3) (Bielski and Richter 1977; Yamaguchi et al. 2005).



Electron Attachment and Anion Radical Stability

Once the secondary electrons are produced in the cell via irradiation-induced water radiolysis, they are willing to react with biomolecules found nearby. The willingness of the electrons to react with the other moieties is bound with those moieties through the **electron affinity** (EA).

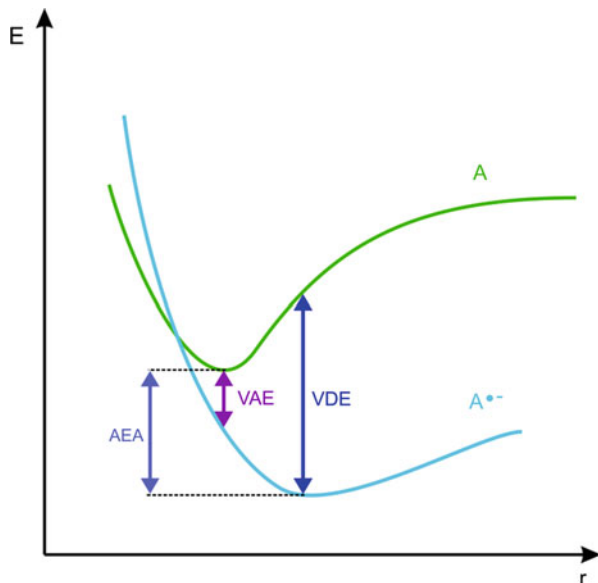
EA can be defined as the difference in energy (or enthalpy or Gibbs free energy) between the moiety in its neutral state (A) and its electron adduct (anion radical form, $\text{A}^{\bullet-}$). Similarly, if one is trying to attach an electron to a cationic molecule B^+ , EA will be then defined as the difference in energies between this B^+ and its electron adduct – the neutral radical B^\bullet . And in the case of a neutral radical (C^\bullet), EA will be defined as the difference in energy of C^\bullet and the corresponding electron adduct, which will be then an anion, C^- . Having in mind these specific cases, we discuss the electron attachment phenomena, based on the first example, which is the pair of the neutral A and the electron adduct $\text{A}^{\bullet-}$ moieties.

The higher electron affinity means, obviously, a higher probability to catch an electron from the surrounding environment. Moreover, it is worth noting that EA is not exactly the same as **ionization potential** (IP). While EA is bound with an attaching electron, IP is bound with a detaching electron from the same neutral moiety. We consider then the following pairs: A and $\text{A}^{\bullet-}$ for EA and A and $\text{A}^{\bullet+}$ for IP. Talking about the irradiation effects on modified DNA, we will be interested in EA rather than IP values.

Three more specific terms are related to the general EA term. These are **adiabatic electron affinity** (AEA), **vertical attachment energy** (VAE, also sometimes named VEA – vertical electron affinity), and **vertical detachment energy** (VDE). AEA is the difference in energy of A and $\text{A}^{\bullet-}$ in their fully relaxed geometries, being the minimum of their potential energy surface (Fig. 2, Eq. 4). The higher (positive) AEA value, the more adiabatically stable the produced anion radical $\text{A}^{\bullet-}$.

If we are dealing with rapid electron transfer processes, when the molecular geometries have insufficient time to adjust to the changed electronic characteristics, we should discuss vertical electron stability and the two terms describing it: VEA and VDE. If the excess electron attaches to the neutral A, the electron adduct $\text{A}^{\bullet-}$

Fig. 2 How to calculate AEA, VEA, and VDE values



has no time for geometrical relaxation (thus, $A^{\bullet-}$ is still in an optimal geometry of neutral A), VEA is the difference in energy of A and $A^{\bullet-}$ (as it was for AEA), but both A and $A^{\bullet-}$ are in the relaxed A geometry (Eq. 5, Fig. 2). In parallel, for the fast detaching electron from the anion radical $A^{\bullet-}$ molecule, we allow no time for geometry relaxation of the produced A (thus, A will be still in the optimum $A^{\bullet-}$ geometry). VDE is defined then as the difference in energy of A and $A^{\bullet-}$ (as it was for AEA and VEA), but both A and $A^{\bullet-}$ are in the relaxed $A^{\bullet-}$ geometry (Eq. 6, Fig. 2).

$$\text{AEA} = E(\text{A}) - E(\text{A}^{\bullet-}) \quad (4)$$

$$\text{VEA} = E(\text{A}) - E(\text{A}^{\bullet-}, \text{ in the A geometry}) \quad (5)$$

$$\text{VDE} = E(\text{A, in the A}^{\bullet-} \text{ geometry}) - E(\text{A}^{\bullet-}) \quad (6)$$

Sometimes it may happen, for the considered molecules, that one obtains positive AEA values and negative VAE values (Li et al. 2002a). Special attention should be given then to the chosen calculation methodology level.

The last point on which we would like to draw attention in this section concerning the description of electron attachment in numbers is bound with the electron-adducts' stability. Namely, one should distinguish the **electronic stability** (both adiabatic – bound with high AEA value, and vertical – bound with a positive VDE value) from the **thermodynamic stability**. The difference is that even if $A^{\bullet-}$, produced via electron attachment, is electronically stable (does not detach its

additional electron spontaneously), it may be thermodynamically unstable if it is prone to undergo further degradation processes (like heterolytic dissociation into an anion and radical). If the kinetic barriers for all of the possible $A^{\bullet-}$ degradation processes are too high to be overcome under ambient conditions, one can conclude that $A^{\bullet-}$ is both electronically and thermodynamically stable and could be detected experimentally.

Armed with the basic terms bound with an electron attachment, we are prepared to discuss this process in the native and modified DNA fragments.

Electron Impact on Native DNA

In 2000, Sanche and coworkers showed that low-energy electrons can cause **single strand breaks (SSB)** and **double strand breaks (DSB)** in plasmid DNA (Boudaiffa et al. 2000). They irradiated DNA with an electron beam with an energy in the range of 3–20 eV and determined the yield of SSB and DSB production. The efficiency of those processes increased with the electron energy. However, the most surprising finding was the observation that electrons at energies below the ionization threshold could contribute to DNA damage, which is caused by rapid decay of transient molecular resonances located on the basic DNA units. These results suggested that secondary electrons produced by ionizing radiation can be lethal to the genome. In this manner, numerous theoretical and experimental studies on secondary electrons, carried out in many laboratories around the world, have been initiated.

One should keep in mind that the studies mentioned above were performed only for “dry” DNA under vacuum conditions, i.e., far from those encountered in cells. In fact, water plays a key role in radiation-induced DNA damage, particularly because it makes up 70–90 % of the mass of living cells. Thus, when a cell is subjected to irradiation, most of the radiation energy is absorbed by water. In aqueous solution, an excess electron swiftly becomes a e^-_{pre} and finally a e^-_{sol} .

In 2009, Lu’s group presented the mechanism of presolvated electron-induced damage to DNA nucleotides with the use of femtosecond, time-resolved, laser spectroscopy (Wang et al. 2009). The use of this technique enabled a real-time observation of the interaction of e^-_{pre} with DNA building blocks. The authors reported that bond breaking in nucleotides can be caused by transient molecular anion resonance formation induced by dissociative ultrafast-electron transfer (UET) of e^-_{pre} . Indeed, 2 years later the same group showed that the formation of DNA strand breaks induced by e^-_{pre} is even nearly twice as efficient as that induced by OH^{\bullet} radicals (Nguyen et al. 2011). These findings suggest that the UET reaction of e^-_{pre} may play a role in generating damage to DNA in aqueous environments. On the other hand, e^-_{pre} are just initial precursor states of an electron located in water with very short lifetimes and the end products are hydrated electrons (e^-_{sol}).

In 2005, Falcone et al. reported experimental studies that showed the influence of the hydration level on electron-induced DNA damage (Falcone et al. 2005). The authors investigated the yields of the products of the interaction between DNA in various hydration states and for dry or hydrated electrons: hydrogen (related

to the reaction of intra-spur electrons) and **5,6-dihydropyrimidines** (the result of electron attachment to cytosine and thymidine). In the lower tested hydration conditions (less than 22 mol of H₂O molecules per mole of nucleotide), the highest yields of 5,6-dihydropyrimidines and the lowest yields of hydrogen were found. At higher tested hydration levels (more than 22 mol of H₂O molecules per mole of nucleotide) an increase in the amount of hydrogen and decrease in yields of 5,6-dihydropyrimidines were observed.

The role of the aqueous environment was also examined by Kohanoff and coworkers (McAllister et al. 2015; Smyth and Kohanoff 2011). Using **ab initio molecular dynamics (AIMD)**, they demonstrated that in the natural DNA surroundings containing many water molecules, the excess electrons localize around the nucleobases in very short times (15 fs). Thus, they suggested that the excess electron created in close proximity to DNA (at least within 2 nm) can avoid hydration. In recent studies, McAllister et al. considered the impact of hydrogen bonding (between the bases and water molecules) and protonation by the water on the possibility of strand break production in DNA (McAllister et al. 2015). It has been shown that protonation causes a significant increase in the kinetic barrier of breaking both the C3'-O3' bond (in case of dCMPH⁻, dGMPH⁻, and dTMPH⁻) and glycosidic bond (in case of dCMPH⁻, dAMPH⁻, and dTMPH⁻) and hence it may prevent strand break production in DNA chains. It is worth noting that both hydrogen bonding to the nucleobase and the water molecule can be a donor of the proton and that the protonation preceded by electron attachment is favorable for many of the bases.

There are several experimental studies in which aqueous solutions of various DNA models, containing hydroxyl radical scavengers such as Tris-HCl or EDTA, were exposed to ionizing radiation. With the use of gel electrophoresis, Cecchini et al. demonstrated that e⁻_{sol} cannot cause strand breaks in a solution of unmodified DNA duplex in contrast to a double stranded DNA fragment labeled with 5-bromouracil (Cecchini et al. 2005). In 2013, Sanche's group reported that e⁻_{sol} generated by water radiolysis are unable to induce strand breaks in native plasmid DNA in contrast to plasmid DNA modified with cisplatin (Rezaee et al. 2013). Moreover, our data show that the electron-induced degradation of trimeric oligonucleotides (TXT) in aqueous solution occurs only if "X" is a brominated nucleobase (Westphal et al. 2015). Indeed, when the trimers contained brominated nucleosides (i.e., X = BrdU, BrdA, BrdG, BrdC), a series of products of electron-induced damage, such as bond breaks or **abasic sites**, were identified. However, when these short oligonucleotides contained native nucleosides (i.e., X = T, dA, dG, dC, dU), only the products of protonation (trimers containing 5,6-dihydrothymidine) were detected. It is worth noting that, in the past, the same trimeric model was exposed to low energy electrons under high vacuum (Park et al. 2012). Significant differences in the identified products for native trimers treated with excess electrons in the gas phase and in the aqueous solution best show the effect of the environment on the damaging potential of electrons toward DNA.

In 1982, it was proven that hydrated electrons cause a very slight inactivation of biologically active, single-stranded DNA of the bacteriophage ϕ X174 compared to

the inactivation induced by OH^\bullet and H^\bullet radicals (Nabben et al. 1982). The biological activity of ϕX174 DNA was presented as a function of gamma radiation dose (survival curves) for exposures carried out with or without an electron scavenger. The results of this experiment demonstrate a low efficiency of damage caused by the reaction of e^-_{sol} with bacteriophage DNA.

All the above mentioned findings demonstrate that e^-_{sol} exhibits a very limited reactivity toward native DNA, causing mostly just its protonation.

Why Use Modified Nucleosides of High Electron Affinity?

Knowing that native DNA is rather resistant to damage caused by irradiation-produced hydrated electrons, it is more understandable why the search for methods of sensitizing DNA to irradiation is necessary to enhance radiotherapy.

Modified nucleobases may play a key role in anticancer radiotherapy. They increase the sensitivity of DNA to ionizing radiation and are referred to as a “Trojan horse” because of their unique properties. These compounds can be readily incorporated into DNA without any modification of gene expression. This is because of their nontoxic properties, structural similarity to the native nucleobases, and the fact that they are recognized by polymerases (Goz 1977) and kinases (Lee and Cheng 1976). Before incorporation into the targeted cell or tissue, the modified nucleoside must be phosphorylated to maintain its radiosensitizing activity. Because of the fact that the nucleoside triphosphate is not chemically stable and has high polarity, the incorporation is also possible when the modified nucleoside is converted to a stable monophosphate nucleoside containing protective groups, such as carbonyloxymethyl (POM, POC), S-acyl-2-thioethyl (SATE), alkoxyalkyl (HDP, ODE), or cyclosaligenyl (cycloSal). These groups are removed inside the cell (enzymatically or chemically), which results in the release of the free nucleoside monophosphate. The resulting molecule can then be easily converted into a nucleoside triphosphate (Pradere et al. 2014).

To ensure the highest radiosensitivity, the derivative of nucleobase that increases the electron affinity, as well as facilitates the abstraction of its substituent, is required to be incorporated into the DNA strand. The substitution of a hydrogen atom with a halogen element changes the electrophilic properties of nucleobases. This results in an increase in their adiabatic electron affinity (AEA) compared to native nucleobases (Rak et al. 2015). Moreover, the highest value of AEA calculated for native nucleobases possesses uracil (Li et al. 2002a), while the 5-bromoderivative is the most radiosensitive halosubstituted 2'-deoxyuridine (Wetmore et al. 2001). The most studied modified nucleosides are 5-BrdU and 5-IdU. Undoubtedly, they are effective substitutes for thymine during replication and have been tested in the medical treatment of tumors such as head-and-neck cancer (Epstein et al. 1994), soft tissue sarcomas (Goffman et al. 1991; Robertson et al. 1995), and malignant gliomas (Greenberg et al. 1988). 5-Bromo- or 5-iodosubstituted uracil enhances the sensitivity of DNA to ionizing radiation, working as an “**electron-trap**” in the biopolymer strand (Li et al. 2002b). The probable mechanism of radiosensitization

for DNA containing 5-BrU is as follows: (i) the electron produced by ionizing radiation attaches to the modified nucleobase (which has the highest electron affinity in the strand), (ii) this is followed by formation of a bromouracil anion radical, (iii) and the **dehalogenation** of bromouracil anion radical results in the formation of a uracil-5-yl radical, whose reactivity will be discussed later in this chapter. Another derivative of nucleobases prone to increase the radiosensitivity of the nucleic acid strand is 5-fluoro-2'-deoxyuridine (5-FdU) (Gaba 2011). This analog has different cytotoxic and radiosensitizing properties than BrdU or IdU. This might be because of the size of the halogen. The van der Waals radius of Br and I are comparable to a methyl group while F is smaller and recalls a hydrogen atom, which makes the fluoroderivative similar to uracil rather than thymine and results in the limited incorporation of 5-FdU into DNA. However, this derivative is able to act as a suppressant of thymidylate synthase. 5-FdU interrupts the action of this enzyme and blocks the synthesis of thymine monophosphate, which results in the interference in DNA replication and inhibition of proliferation of tumor cells (Shewach and Lawrence 2006).

Dehalogenation of Modified Nucleobases

Independently of the final degradation reactions, most of the halogenated nucleoside induced damages begin with dehalogenation processes.

Electron attachment induced dehalogenation of **halonucleobases**, such as 5-bromouracil (5-BrU) and 5-iodouracil (5-IU), has long been a subject for theoretical studies. This is due to the discovery of their radiosensitizing properties, briefly discussed in the previous sections. It was shown experimentally that the attachment of a solvated electron to 5-bromo-2'-deoxyuridine initially leads to the formation of the respective anion radical (5-BrU^{•-}). 5-BrU^{•-} is thermodynamically unstable, therefore its creation is followed by dehalogenation as a second step, yielding bromide anion and uracil-5-yl radical (Fig. 3). This last species, being very reactive, possesses a significant potential to abstract hydrogen from adjacent sugar moieties, leading to single or double strand breaks as a result (Rak et al. 2015).

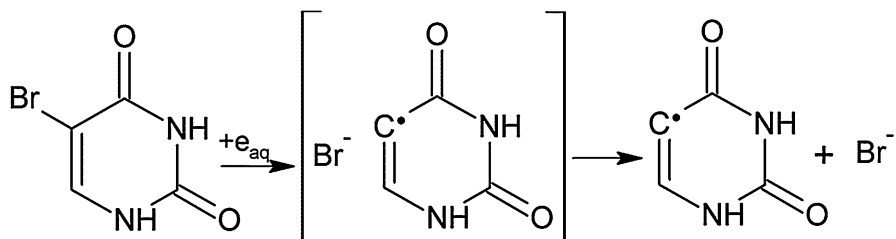


Fig. 3 Adding an excess electron to 5-BrU leads to the formation of an unstable anion radical that finally decomposes to uracil-5-yl radical and bromide anion

Table 1 Vertical and adiabatic EAs for the series of halouracils, uracil, and thymine (eV) (Li et al. 2002b; Wetmore et al. 2001)

	Gas phase				Solution		
	VAE ^a	VAE ^b	AEA ^a	AEA ^{b,c}	VAE ^a	AEA ^a	AEA ^b
U	-0.26	-0.35 ^d	0.18	0.18	1.62	2.30	2.02
T	-0.30	-	0.14	-	1.56	2.16	-
5-FU	-0.03	-0.11 ^d	0.45	0.48	1.82	2.42	2.21
5-CIU	0.14	0.06	0.58	0.60	1.94	2.53	2.26
5-BrU	0.17	0.11	0.64	0.63	1.98	2.60	2.44

^aWetmore et al. (2001)^bLi et al. (2002b)^cZPE corrected^dRecalculated for this chapter at the same B3LYP/6-31 + G(d) level

Both electron attachment and **dehalogenation** can be easily described by the use of computational chemistry methods. Molecular activity towards electrons can be characterized by the AEA, VEA, VDE values, defined in the second section of this chapter.

Starting from Wetmore's work (Wetmore et al. 2001), the adiabatic and vertical electron affinities of a series of **5-halouracils**, uracil, and thymine have been calculated both in the gas phase and solution (**polarizable continuum model, PCM**) using **DFT methods** (B3LYP/6-311G(2df,p) on B3LYP/6-31 + G(d,p) or higher level gas phase geometries). Their results, later verified positively with the use of a slightly different methodology (B3LYP functional and 6-31 + G(d) basis set with isodensity PCM for water solution) (Li et al. 2002b), are summarized therein (see Table 1). As some discrepancies were found between VEAs shown in the Li et al. and Wetmore et al. papers, we recalculated them, where necessary (see Table 1). Positive EA values imply that the anionic molecule is more stable than the corresponding neutral molecule, especially for 5-CIU and 5-BrU, and the stability grows in the following order: T < U < 5-FU < 5-CIU < 5-BrU (Wetmore et al. 2001).

The second step – dehalogenation – was investigated computationally for these compounds. The study suggested that, for the series of halouracils, the barrier induced by electron attachment dissociation to X⁻ and U* goes in the opposite order: 5-BrU < 5-CIU < 5-FU, with the activation energies for dehalogenation calculated as 1.9 kcal/mol, 4.0 kcal/mol, and 20.8 kcal/mol, respectively (Wetmore et al. 2001). We can therefore conclude that 5-BrU has the largest (among other halogenated nucleobases) potential to dehalogenate itself after electron attachment and therefore 5-BrU is the most effective electron-trap for electrons produced via ionizing radiation.

It is worth mentioning that this reaction has also been widely studied experimentally – by nanosecond laser pulse radiolysis and time resolved laser spectroscopy – and all these experimental results are in good accordance with the computations

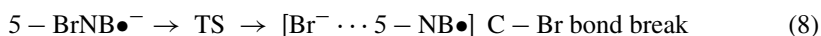
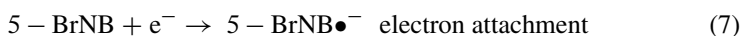
Table 2 AEA (in free enthalpy scale, eV), VDE [eV], kinetic barrier (ΔG^* , kcal/mol), and global driving force (ΔG_{tot} , kcal/mol) of methylated BrNBs, 5-SCNU, and 5-OCNU (Chomicz et al. 2013a, b)

	Gas				Solution			
	AEA	VDE	ΔG^*	ΔG_{tot}	AEA	VDE	ΔG^*	ΔG_{tot}
Me-8BrA	0.01 ^a	0.99	–	0.24	1.87 ^a	2.91	–	–63.69
Me-8BrG	0.02	0.06	0.44	–0.62	1.36	2.04	0.41	–63.94
Me-5BrC	0.27	0.76	0.55	–3.51	2.07	2.51	2.15	–66.40
Me-5BrU	0.49	1.02	2.38	–1.52	2.27	2.74	2.54	–64.96
5-SCNU	–	–	–	–	2.7	3.07	3.4	–67.5
5-OCNU	–	–	–	–	2.62	3.01	2.5	–83.7

^aConstrained geometry optimization with the C-Br distance frozen at 1.898 Å

showing that EA is much higher for halogen substituted uracils than for unmodified ones or thymine (Rak et al. 2015).

5-BrU seems to have a great potential to enhance DNA sensitivity, being the major electron capturing site during irradiation, but also other derivatives have been tested, including other brominated nucleobases (BrNB) and other uracil derivatives (XU). First of all, it has been shown experimentally that not only 5-BrU but also 5-BrC is a three times better electron scavenger compared to native cytosine (Razskazovskii et al. 1997). Also 8-bromopurine analogues were proven to possess electron affinities far exceeding those for nonhalogenated ones (Flyunt et al. 2000). Knowing that, in the experiment, 8-bromo derivatives of adenine or guanine capture electrons to form the C8 radical and bromide ion (Flyunt et al. 2000; Ioele et al. 2000), Chomicz et al. compared the electron induced dissociation of four nucleobase derivatives: 8-bromo-9-methyladenine (Me-8BrA), 8-bromo-9-methylguanine (Me-8BrG), 5-bromo-1-methylcytosine (Me-5BrC), and 5-bromo-1-methyluracil (Me-5BrU) both in the gas phase and PCM model, at the B3LYP/6-31++G(d,p) level of theory (Chomicz et al. 2012). Those calculations allowed the estimation of the thermodynamic and kinetic barriers (see ΔG_{tot} and ΔG^* in Table 2) of the three following reaction steps (Eqs. 7, 8, and 9):



The survey showed that all the considered BrNBs are far better electron acceptors compared to unsubstituted ones, prone to form isolated nucleobase radicals and

bromide anion. The calculated EAs were in all cases positive (even above 2 eV in the free enthalpy scale) but for purine derivatives smaller than for native nucleobases (less than 2 eV in aqueous solution). The thermodynamic and kinetic characteristics of molecular decomposition into a nucleobase radical and bromide anion were very promising (Chomicz et al. 2012). Here, we can also observe that pyrimidines are more likely to accept an excess electron than purines, and that anionic molecules are stabilized by a polar environment, which is the reason for the much higher electron affinity in solution. These findings were proven with the use of AIMD methodology, additionally showing that brominated purine decomposition should also be barrier-free (Wieczór et al. 2014).

What Next with the Nucleobase Radicals Produced?

Radicals generated in the DNA molecule can cause a wide range of damage such as single and double strand breaks, abasic sites, cyclopurines, and interstrand cross-links.

The first process occurring after the generation of a radical is usually a **hydrogen atom transfer**. The hydrogen target for abstraction depends on the position that the radical is localized on within the nucleobase. Radicals based on the carbon atom neighboring the *N*-glycosidic bond of the nucleotide (C6 in pyrimidines, C8 in purines) are likely to abstract a hydrogen atom from their own sugar residue. Due to the structure of DNA, 2', 3', and 5' hydrogen atoms can be the target as 1' and 4' hydrogen atoms are located on the opposite side of the 2'-deoxyribose ring (Fig. 4). In the case of radicals located on the C5 atom of pyrimidines, hydrogen atom transfer from the same nucleotide is hindered by steric repulsion between the sugar ring and the C5 atom of the pyrimidine. The conformational changes that are essential for such a reaction to occur make the kinetic barrier over 40 kcal/mol – impossible to overcome under cellular conditions (Chomicz et al. 2014b). Despite that, 5-substituted pyrimidines still exhibit radiosensitizing properties.

One possible pathway for pyrimidine radicals centered on the 5-position is abstraction of a hydrogen atom from the adjacent nucleotide, where 1' and 2' positions can be targeted (Schyman et al. 2007). These radicals can also abstract a hydrogen atom from a water molecule, producing a hydroxyl radical – which is known to be genotoxic (Fig. 4). Hydroxyl radicals generated in this reaction can immediately attach to the C5 carbon atom, making the spin density move to the C6-position, or they can go on to abstract hydrogen atoms from any sugar residue in the DNA chain (Chomicz et al. 2014c). The 5' and 3' sites have the lowest kinetic barriers for such an abstraction, although they are influenced to a large extent by the ever-changing local structure of water, and the barriers for other sites within the sugar moiety are too small to prevent abstraction there (Shukla et al. 2011).

Sugar centered radicals can lead to various kinds of damage depending on the site they are localized on. 3' and 5' radicals are centered on atoms bonded with phosphate groups and, as such, they can undergo homolytic O–P bond cleavage, which is tantamount to a strand break. C–O bond breakage is impossible in this

Fig. 4 Possible targets for radical attack by 5- (marked green) and 6-uridyl (blue), as well as hydroxyl (red) radical

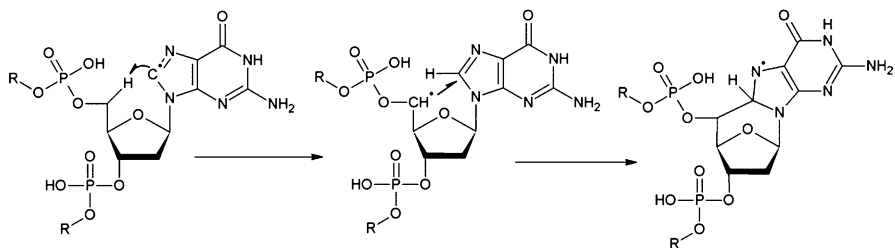
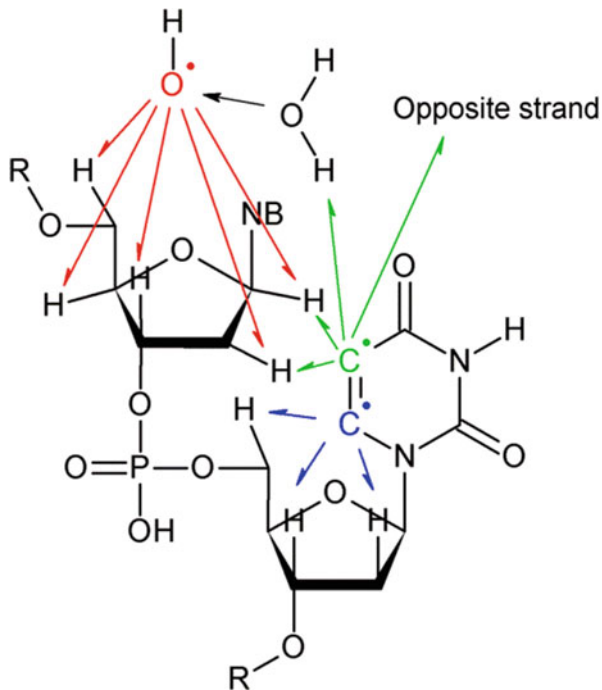


Fig. 5 The mechanism of 5',8-cyclo-2'-deoxyadenosine formation from 8-adenosinyl radical

case because elongating this bond leads to a phosphate radical along with a sugar centered open shell singlet, with an energy 60 kcal/mol higher than the substrate (Rak et al. 2011). When 5' radicals are located within a purine nucleotide they can also attack the C8 atom of the nucleobase to produce a **cyclopurine lesion** (Fig. 5). These reactions are more favorable kinetically and thermodynamically than 5' O-P bond cleavage (Chomicz et al. 2014a) and the lesion formed, although less cytotoxic than strand breaks, can still distort the DNA helix and thus impair cellular processes (Kropachev et al. 2014).

4' radicals can lead to strand breaks in single stranded DNA, as the 3' side C-O bond can be cleaved heterolytically to produce a cation radical, followed by hydrolysis to a new 4' radical, which can then undergo a similar reaction to break

the 5' side C–O bond (Giese et al. 1995). 2' radicals might in theory cause damage with a similar mechanism. However, there is no evidence of this pathway being feasible in DNA, as opposed to RNA, where the 2' hydroxyl group makes the 2' C–H bond weaker (Paul and Greenberg 2014). 1' radicals are known to transform into abasic sites, usually of **2-deoxyribonolactone** type, after oxidation. These alkali-labile sites can be incised by apurinic endonuclease 1 followed by forming a **Schiff base cross-link product** with DNA polymerase, which impedes further DNA repair (Greenberg 2007).

When the 5-pyrimidinyl radical is placed in a mismatched region it can also attack the opposite strand to form cross-links (Fig. 4). The mechanism proposed for the formation of this kind of lesion was based on a special zipper-like structure that the **mismatched DNA** can form (Dextraze et al. 2010). In general the conformation of DNA has a strong influence on the type of damage formed by 5-substituted pyrimidines.

Quantum Chemistry as a Tool to Search for New Therapeutics

Knowledge and experience in treating cancer and many decades of studies on the use of radiotherapy have allowed for a class of compounds referred to as radiosensitizers to be proposed for the purpose of sensitizing the DNA of tumorous cells. Radiosensitizers may become the future of modern radiotherapy, as selectively damaging tumorous cells using a lower dose of high-energy radiation ensures the highest effectiveness of treatment and limits side effects. The class includes the above mentioned 5-BrU, also in a nucleoside form 5-BrdU. The radiosensitizing mechanism of this compound on the molecular level is still not fully understood despite 50 years of studies and opportunities for clinical application.

It seems justified to search for compounds with sensitizing properties among other derivatives of nucleobases, rather than focusing on thymine derivatives or bromine modified nucleobases. This is especially true since anticancer therapy has already seen the use of analogues of pyrimidines such as 5-iodo-2'-deoxyuridine, 2',2'-difluoro-2'-deoxycitidine (gemcitabine, sold commercially as Gemzar), which increases the cytotoxicity in tumorous cells under radiation or another group: 5-fluorouracil (5-FU) or 5-fluoro-2'-deoxyuridine (5-FdU) metabolized into a monophosphate (5-FdUMP). 5-FdUMP acts as an inhibitor to thymidylate synthase and can be incorporated into DNA in the form of a triphosphate. Its radiosensitizing properties were studied clinically, too (Longley et al. 2003). Both in vitro and in vivo studies show that 4-thiothymidine has considerable potential as a photosensitizer in photodynamic therapy. Thanks to the substitution of the oxygen atom at the fourth position in thymine for a sulfur atom, the compound becomes sensitive to UVA radiation. Furthermore, experimental research shows that, despite certain reservations against clinical application due to the limited power of UVA radiation, this type of radiation is sufficient to induce the apoptosis of malignant skin cells when used with a particular concentration of the drug (Gemenetzidis et al. 2013).

Computer chemistry is a helpful tool in designing new sensitizers and understanding the mechanism of action of the known ones. Modifying a nucleobase by substituting a hydrogen atom for halogen substituents (Y: fluorine, chlorine, bromine, and iodine) at C5 (in pyrimidines) or C8 (purines) causes a change in the properties of the nucleobase as an electron acceptor and the durability of the C–Y bond (as described using DFT methods). Research has proven that the production of the 5-uridine radical due to the detachment of the electron acceptor substituent increases if the uracil is substituted with bromine. Research has also proven that modifying other bases with bromine increases their electron affinity compared to their native counterparts (Li et al. 2002b).

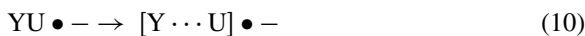
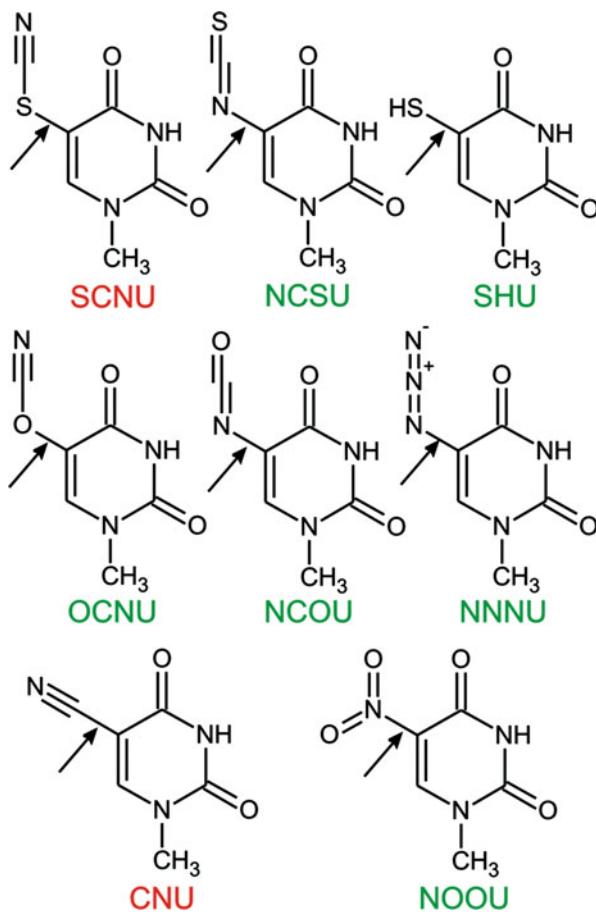
Chomicz et al. conducted theoretical research on the addition of an excess electron to model bromine-derived, methylated nucleobases (Me-8BrA, Me-5BrC, Me-8BrG, and Me-5BrU), which showed that all the tested nucleobases, once incorporated into a DNA strand, might sensitize the biopolymer, as indicated by the possibility of a stable valence radical anion being created both in the gas phase and the liquid phase and by the values of AEA and VDE (Table 2). The excess electron is usually located on the π^* orbital. Research shows that BrdA should undergo barrier-free degradation due to the lack of the production of a π^* radical anion. According to the authors, a C–Br bond in BrNB would break once a kinetic barrier of approximately 2.5 kcal/mol is exceeded and a stable bromine anion would be created then (Chomicz et al. 2012).

Simulations of the introduction of bromine-modified nucleotides, i.e., 3',5'-diphosphates of 8-bromo-2'-deoxyguanosine (8-BrdGDP) or 8-bromo-2'-deoxyadenosine (8-BrdADP), into a DNA strand indicate that this may sensitize the strand to high-energy radiation. The damage may occur as a result of the breaking of the O–P bond, leading to the creation of a closed-shelled carbonyl group and a phosphoryl radical or, more likely, cyclic mutations as a result of the induction of 5',8-cycloadenosine (cycloA) and 5',8-cycloguanosine (cycloG) (Chomicz et al. 2014a; Chomicz et al. 2013a).

Studies show that 5-bromine-substituted derivatives of pyrimidines are not the only compounds that may be useful in designing new anticancer drugs. Theoretical research on 8-bromo-2' deoxycytosine 3',5'-diphosphate (6-BrdCDP) and 6-bromo-2'-deoxyuridine 3',5'-diphosphate (6-BrdUDP) indicates that such modifications lead to barrier-free breakage of the C6-Br bond and the creation of a radical at the 6th carbon atom. Thanks to steric hindrance (ΔG^* activation barrier from 8.2 to 13.7 kcal/mol), this radical becomes stable by means of a hydrogen atom moving from the C3' position to C5' and similar structures forming through breaking, as reported in a study on the degradation of purine bromine nucleotides (Chomicz et al. 2014b; Golon et al. 2014).

Working under the assumption that a potential sensitizer should be a better electron acceptor than thymine, and the produced radical anion should easily degrade to a uridine radical, Rak et al. proposed two new, very promising sensitizers. The values of electron affinity (VDE and AEA) as well as thermodynamic and kinetic characteristics of reactions (Eqs. 10 and 11) were calculated for eight rationally chosen uracil derivatives (Fig. 6).

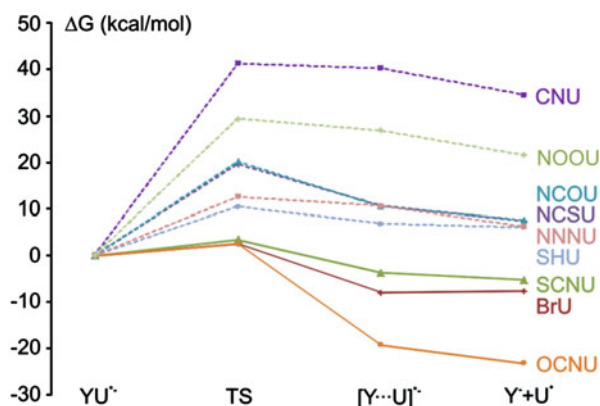
Fig. 6 5-Substituted derivatives of uracil (YU) with abbreviated names. Arrows indicate bonds that are potentially vulnerable to breakage after an excess electron is added



The results show that two of these derivatives, 5-thiocyanouracil (5-SCNU) and 5-cyanatouracil (5-OCNU), display very promising sensitizing properties, as indicated by negative values of ΔG and the activation barriers for reaction (Eq. 10) amounting to 2–3 kcal/mol only (Table 2). These calculations were confirmed experimentally using photoelectron spectroscopy (PES) for two chosen derivatives (5-SCNU and 5-CNU), for which computational studies predicted a different behavior. While it was possible to obtain the PES spectrum of the CNU^{•-} radical anion, the SCNU^{•-} anion radical spectrum was not registered, due to the SCNU^{•-} thermodynamical instability and its degradation into U[•] and SCN⁻ (Chomicz et al. 2013b) (Fig. 7).

Recent studies (Zdrowowicz et al. 2015) confirmed the sensitizing character of the uridine thiocyno derivative (SCNdU). The addition of an electron in an

Fig. 7 Energy profiles of degradation reactions for 5-substituted uracil derivatives



ESR experiment was found to lead to the production of uridine ($\text{dU}\bullet$) and $\text{dU-5-S}\bullet$ radicals. On the other hand, DFT calculations explained that the production of $\text{dU-5-S}\bullet$ and $\text{CN}\bullet$ is thermodynamically more beneficial than the creation of $\text{dU}\bullet$ and $\text{SCN}\bullet$. The activation barrier for breaking the C5-S and S-CN bonds in $\text{SCNdU}\bullet$ amounts to 8.7 kcal/mol and 4.0 kcal/mol, respectively. The computational data were confirmed by experimental research using ESR and LC-MS techniques, which observed the production of a dU-S-S-dU dimer. This suggests that SCNdU could be used as a potential radiosensitizer and would cause inter- and intrastrand cross-links within DNA.

Summary

Over 15 years have passed since the discovery that low-energy electrons are able to trigger single (SSB) and double (DSB) strand breaks in dry DNA. A number of experimental and theoretical studies aimed at explaining the mechanism of electron-induced strand break formation have been published over those years (for excellent reviews see e.g., Gu et al. 2012; Rak et al. 2008; Sanche 2005; Sanche 2009).

At first glance this phenomenon, especially the formation of DSB damage that correlates with cell death induced by ionizing radiation, has obvious practical connotations. In the current review we demonstrate, however, that as far as an aqueous environment is concerned, hydrated electrons are unable to introduce strand breaks in DNA. Although they react with nucleobases at the rate controlled by diffusion, swift and reversible proton transfer between water and the anion, resulting from electron attachment to the nucleobase, prevents the formation of SBs observed in dry DNA. On the other hand, if DNA is labeled with electron affinic, modified nucleosides, which undergo an efficient dissociative electron attachment (DEA), the nucleobase radicals formed within the DEA process are able to trigger SBs, as well as other serious DNA damage (intra- and interstrand cross-links) leading to lethal effects.

Thus, the introduction of modified nucleosides to tumor cells may constitute a foundation for a Trojan horse anticancer treatment dedicated to hypoxic conditions. Indeed, with the help of DFT modeling, we showed that bromonucleobases that swiftly release the bromide anion due to electron attachment could work as radiosensitizers. The nucleobase radicals that are left in DNA due to the DEA process to brominated nucleobases may be, in turn, involved in hydrogen transfer reactions that ultimately lead to strand breaks, cyclic purines, or cross-links.

Last, but not least, a computational approach that should result in new, efficient radiosensitizers is briefly characterized. Although radiotherapy is the most common technique employed against cancer, the number of radiosensitizers employed currently in clinics is surprisingly small. Therefore, access to new compounds sensitizing cancer cells to ionizing radiation seem to be a real obstacle for the further progress of therapy. Hence, computational chemistry methods, as demonstrated in this review, seem to be the methods of hope for solving this pressing issue.

Acknowledgments This work was supported by the Polish National Science Centre under Grant No. 2014/14/A/ST4/00405 (J.R.).

Bibliography

- Abel, B., Buck, U., Sobolewski, A., & Domcke, W. (2012). On the nature and signatures of the solvated electron in water. *Physical Chemistry Chemical Physics*, *14*, 22–34.
- Bielski, B. H., & Richter, H. W. (1977). A study of the superoxide radical chemistry by stopped-flow radiolysis and radiation induced oxygen consumption. *Journal of the American Chemical Society*, *99*(9), 3019–3023.
- Boudaiffa, B., Cloutier, P., Hunting, D., Huels, M. A., & Sanche, L. (2000). Resonant formation of DNA strand breaks by low-energy (3 to 20 eV) electrons. *Science*, *287*(5458), 1658–1660.
- Cecchini, S., Girouard, S., Huels, M. A., Sanche, L., & Hunting, D. J. (2005). Interstrand cross-links: A new type of gamma-ray damage in bromodeoxyuridine-substituted DNA. *Biochemistry*, *44*(6), 1932–1940.
- Chomicz, L., Rak, J., & Storonik, P. (2012). Electron-induced elimination of the bromide anion from brominated nucleobases. A computational study. *Journal of Physical Chemistry B*, *116*(19), 5612–5619.
- Chomicz, L., Leszczynski, J., & Rak, J. (2013a). Electron-induced degradation of 8-Bromo-2'-deoxyadenosine 3',5'-diphosphate, a DNA radiosensitizing nucleotide. *Journal of Physical Chemistry B*, *117*(29), 8681–8688.
- Chomicz, L., Zdrochowicz, M., Kasprzykowski, F., Rak, J., Buonaugurio, A., Wang, Y., & Bowen, K. H. (2013b). How to find out whether a 5-substituted uracil could be a potential DNA radiosensitizer. *Journal of Physical Chemistry Letters*, *4*(17), 2853–2857.
- Chomicz, L., Furmanchuk, A., Leszczynski, J., & Rak, J. (2014a). Electron induced single strand break and cyclization: A DFT study on the radiosensitization mechanism of the nucleotide of 8-bromoguanine. *Physical Chemistry Chemical Physics*, *16*, 6568.
- Chomicz, L., Golon, Ł., & Rak, J. (2014b). The radiosensitivity of 5- and 6-bromocytidine derivatives – Electron induced DNA degradation. *Physical Chemistry Chemical Physics*, *16*, 19424.
- Chomicz, L., Petrovici, A., Archbold, I., Adhikary, A., Kumar, A., Sevilla, M. D., & Rak, J. (2014c). An ESR and DFT study of hydration of the 2'-deoxyuridine-5-yl radical: A possible hydroxyl radical intermediate. *Chemical communications (Cambridge)*, *50*(93), 14605–14608.

- Daşu, A., & Denekamp, J. (1998). New insights into factors influencing the clinically relevant oxygen enhancement ratio. *Radiotherapy and Oncology*, 46(3), 269–277.
- Dextraze, M. E., Gantchev, T., Girouard, S., & Hunting, D. (2010). DNA interstrand cross-links induced by ionizing radiation: An unsung lesion. *Mutation Research*, 704(1–3), 101–107.
- Epstein, A. H., Lebovics, R. S., Goffman, T., Teague, D., Fuetsch, E. S., Glatstein, E., Okunieff, P., & Cook, J. A. (1994). Treatment of locally advanced cancer of the head and neck with 5'-iododeoxyuridine and hyperfractionated radiation therapy: Measurement of cell labeling and thymidine replacement. *Journal of the National Cancer Institute*, 86(23), 1775–1780.
- Falcone, J. M., Becker, D., Sevilla, M. D., & Swarts, S. G. (2005). Products of the reactions of the dry and aqueous electron with hydrated DNA: Hydrogen and 5,6-dihydropyrimidines. *Radiation Physics and Chemistry*, 72(2–3), 257–264.
- Flyunt, R., Bazzanini, R., Chatgililoglu, C., & Mulazzani, Q. G. (2000). Fate of the 2'-deoxyadenosin-5'-yl radical under anaerobic conditions. *Journal of the American Chemical Society*, 122(17), 4225–4226.
- Gaba, N. (2011). Halogenated pyrimidines. In *Radioprotectors and radiosensitizers*. Saarbrücken: LAP LAMBERT Academic Publishing GmbH & Co.
- Gemenetzidis, E., Shavorskaya, O., Xu, Y. Z., & Trigiant, G. (2013). Topical 4-thiothymidine is a viable photosensitizer for the photodynamic therapy of skin malignancies. *Journal of Dermatological Treatment*, 24(3), 209–214.
- Giese, B., Beyrich-Graf, X., Erdmann, P., Petretta, M., & Schwitter, U. (1995). The chemistry of single-stranded 4'-DNA radicals: Influence of the radical precursor on anaerobic and aerobic strand cleavage. *Chemistry and Biology*, 2(6), 367–375.
- Goffman, T., Tochner, Z., & Glatstein, E. (1991). Primary treatment of large and massive adult sarcomas with iododeoxyuridine and aggressive hyperfractionated irradiation. *Cancer*, 67(3), 572–576.
- Golon, Ł., Chomicz, L., & Rak, J. (2014). Electron-induced single strand break in the nucleotide of 5- and 6-bromouridine. A DFT study. *Chemical Physics Letters*, 612, 289–294.
- Goz, B. (1977). The effects of incorporation of 5-halogenated deoxyuridines into the DNA of eukaryotic cells. *Pharmacological Reviews*, 29(4), 249–272.
- Greenberg, M. M. (2007). Elucidating DNA damage and repair processes by independently generating reactive and metastable intermediates. *Organic and Biomolecular Chemistry*, 5(1), 18–30.
- Greenberg, H. S., Chandler, W. F., Diaz, R. F., Ensminger, W. D., Junck, L., Page, M. A., Gebarski, S. S., McKeever, P., Hood, T. W., & Stetson, P. L. (1988). Intra-arterial bromodeoxyuridine radiosensitization and radiation in treatment of malignant astrocytomas. *Journal of Neurosurgery*, 69(4), 500–505.
- Gu, J., Leszczynski, J., & Schaefer, H. F., III. (2012). Interactions of electrons with bare and hydrated biomolecules: From nucleic acid bases to DNA segments. *Chemical Reviews*, 112(11), 5603–5640.
- Ioele, M., Bazzanini, R., Chatgililoglu, C., & Mulazzani, Q. G. (2000). Chemical radiation studies of 8-bromoguanosine in aqueous solutions. *Journal of the American Chemical Society*, 122(9), 1900–1907.
- Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., & Forman, D. (2011). Global cancer statistics. *CA: A Cancer Journal for Clinicians*, 61(2), 69–90.
- Kropachev, K., Ding, S., Terzidis, M. A., Masi, A., Liu, Z., Cai, Y., Kolbanovskiy, M., Chatgililoglu, C., Broyde, S., Geacintov, N. E., & Shafirovich, V. (2014). Structural basis for the recognition of diastereomeric 5',8-cyclo-2'-deoxypurine lesions by the human nucleotide excision repair system. *Nucleic Acids Research*, 42(8), 5020–5032.
- Kumar, A., Walker, J. A., Bartels, D. M., & Sevilla, M. D. (2015). A simple ab initio model for the hydrated electron that matches experiment. *The Journal of Physical Chemistry A*, 119, 9148.
- Le Caer, S. (2011). Water radiolysis: Influence of oxide surfaces on H₂ production under ionizing radiation. *Water*, 3(1), 235–253.

- Lee, L. S., & Cheng, Y. C. (1976). Human deoxythymidine kinase. I. Purification and general properties of the cytoplasmic and mitochondrial isozymes derived from blast cells of acute myelocytic leukemia. *Journal of Biological Chemistry*, *251*(9), 2600–2604.
- Li, X., Cai, Z., & Sevilla, M. (2002a). DFT calculations of the electron affinities of nucleic acid bases: Dealing with negative electron affinities. *Journal of Physical Chemistry A*, *106*(8), 1596–1603.
- Li, X. F., Sanche, L., & Sevilla, M. D. (2002b). Dehalogenation of 5-halouracils after low energy electron attachment: A density functional theory investigation. *The Journal of Physical Chemistry. A*, *106*(46), 11248–11253.
- Longley, D. B., Harkin, D. P., & Johnston, P. G. (2003). 5-fluorouracil: Mechanisms of action and clinical strategies. *Nature Reviews Cancer*, *3*(5), 330–338.
- McAllister, M., Smyth, M., Gu, B., Tribello, G. A., & Kohanoff, J. (2015). Understanding the interaction between low-energy electrons and DNA nucleotides in aqueous solution. *Journal of Physical Chemistry Letters*, *6*(15), 3091–3097.
- Nabben, F. J., Karman, J. P., & Loman, H. (1982). Inactivation of biologically active DNA by hydrated electrons. *International Journal of Radiation Biology and Related Studies in Physics, Chemistry, and Medicine*, *42*(1), 23–30.
- Nguyen, J., Ma, Y., Luo, T., Bristow, R. G., Jaffray, D. A., & Lu, Q. B. (2011). Direct observation of ultrafast-electron-transfer reactions unravels high effectiveness of reductive DNA damage. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(29), 11778–11783.
- Oronsky, B. T., Knox, S. J., & Scicinski, J. (2011). Six degrees of separation: The oxygen effect in the development of radiosensitizers. *Translational Oncology*, *4*(4), 189–198.
- Park, Y., Polska, K., Rak, J., Wagner, J. R., & Sanche, L. (2012). Fundamental mechanisms of DNA radiosensitization: Damage induced by low-energy electrons in brominated oligonucleotide trimers. *Journal of Physical Chemistry B*, *116*(32), 9676–9682.
- Paul, R., & Greenberg, M. M. (2014). Independent generation and reactivity of uridin-2'-yl radical. *Journal of Organic Chemistry*, *79*(21), 10303–10310.
- Pradere, U., Garnier-Amblard, E. C., Coats, S. J., Amblard, F., & Schinazi, R. F. (2014). Synthesis of nucleoside phosphate and phosphonate prodrugs. *Chemical Reviews*, *114*(18), 9154–9218.
- Rak, J., Mazurkiewicz, K., Kobylecka, M., Storoniak, P., Haranczyk, M., Dąbkowska, I., Bachorz, R. A., Gutowski, M., Radisic, D., Stokes, S. T., Eustis, S. N., Wang, D., Xiang, L., Ko, Y. J., & Bowen, K. H. (2008). Stable valence anions of nucleic acid bases and DNA strand breaks induced by low energy electrons. In *Radiation induced molecular phenomena in nucleic acids*. London: Springer.
- Rak, J., Kobylecka, M., & Storoniak, P. (2011). Single strand break in DNA coupled to the O–P bond cleavage. *Journal of Physical Chemistry B*, *115*, 1911–1917.
- Rak, J., Chomicz, L., Wicz, J., Westphal, K., Zdrochowicz, M., Wityk, P., Żyndul, M., Makurat, S., & Golon, Ł. (2015). Mechanisms of damage to DNA labeled with electrophilic nucleobases induced by ionizing or UV radiation. *Journal of Physical Chemistry B*, *119*(26), 8227–8238.
- Razskazovskii, Y., Swarts, S. G., Falcone, J. M., Taylor, C., & Sevilla, M. D. (1997). Competitive electron scavenging by chemically modified pyrimidine bases in bromine-doped DNA: Relative efficiencies and relevance to intrastrand electron migration distances. *Journal of Physical Chemistry B*, *101*(8), 1460–1467.
- Reelfs, O., Karran, P., & Young, A. R. (2012). 4-Thiothymidine sensitization of DNA to UVA offers potential for a novel photochemotherapy. *Photochemical and Photobiological Sciences*, *11*(1), 148–154.
- Rezaee, M., Sanche, L., & Hunting, D. J. (2013). Cisplatin enhances the formation of DNA single- and double-strand breaks by hydrated electrons and hydroxyl radicals. *Radiation Research*, *179*(3), 323–331.
- Robertson, J. M., Sondak, V. K., Weiss, S. A., Sussman, J. J., Chang, A. E., & Lawrence, T. S. (1995). Preoperative radiation therapy and iododeoxyuridine for large retroperitoneal sarcomas. *International Journal of Radiation Oncology Biology Physics*, *31*(1), 87–92.

- Sanche, L. (2005). Low energy electron-driven damage in biomolecules. *European Physical Journal D*, 35(2), 367–390.
- Sanche, L. (2009). Low-energy electron interaction with DNA: Bond dissociation and formation of transient anions, radicals, and radical anions. In M. M. Greenberg (Ed.), *Radical and radical ion reactivity in nucleic acid chemistry*. Hoboken: Wiley.
- Schyman, P., Zhang, R., Eriksson, L. A., & Laaksonen, A. (2007). Hydrogen abstraction from deoxyribose by a neighbouring uracil-5-yl radical. *Physical Chemistry Chemical Physics*, 9, 5975–5979.
- Shewach, D. S., & Lawrence, T. S. (2006). Nucleoside radiosensitizers. In P. J. Godefridus (Ed.), *Deoxynucleoside analogs in cancer therapy*. Totowa: Humana Press.
- Shukla, P. K., Kumar, N., & Mishra, P. C. (2011). Hydrogen atom abstraction reactions of the sugar moiety of 2'-deoxyguanosine with an OH radical: A quantum chemical study. *International Journal of Quantum Chemistry*, 111(9), 2160–2169.
- Smyth, M., & Kohanoff, J. (2011). Excess electron localization in solvated DNA bases. *Physical Review Letters*, 106(23), 238108.
- von Sonntag, C. (2006). *Free-radical-induced DNA damage and its repair*. Berlin: Springer.
- Wang, C. R., Nguyen, J., & Lu, Q. B. (2009). Bond breaks of nucleotides by dissociative electron transfer of nonequilibrium prehydrated electrons: A new molecular mechanism for reductive DNA damage. *Journal of the American Chemical Society*, 131(32), 11320–11322.
- Westphal, K., Wiczak, J., Miloch, J., Kciuk, G., Bobrowski, K., & Rak, J. (2015). Irreversible electron attachment – A key to DNA damage by solvated electrons in aqueous solution. *Organic and Biomolecular Chemistry*, 13(41), 10362–10369.
- Wetmore, S., Boyd, R., & Eriksson, L. (2001). A theoretical study of 5-halouracils: Electron affinities, ionization potentials and dissociation of the related anions. *Chemical Physics Letters*, 343(1–2), 151–158.
- Wieczór, M., Wityk, P., Czub, J., Chomicz, L., & Rak, J. (2014). A first-principles study of electron attachment to the fully hydrated bromonucleobases. *Chemical Physics Letters*, 595, 133–137.
- Williams, K. J., Cowen, R. L., & Stratford, I. J. (2001). Hypoxia and oxidative stress. Tumour hypoxia – Therapeutic considerations. *Breast Cancer Research*, 3(5), 328–331.
- Yamaguchi, H., Uchihori, Y., Yasuda, N., Takada, M., & Kitamura, H. (2005). Estimation of yields of OH radicals in water irradiated by ionizing radiation. *Journal of Radiation Research*, 46(3), 333–341.
- Zdrowowicz, M., Chomicz, L., Żyndul, M., Wityk, P., Rak, J., Wiegand, T. J., Hanson, C. G., Adhikary, A., & Sevilla, M. D. (2015). 5-Thiocyanato-2'-deoxyuridine as a possible radiosensitizer: Electron-induced formation of uracil-C5-thiyl radical and its dimerization. *Physical Chemistry Chemical Physics*, 17(26), 16907–16916.

Jakub Rydzewski and Wieslaw Nowak

Contents

Introduction	1918
Adiabatic Molecular Dynamics	1920
Born-Oppenheimer Approximation	1920
Car-Parrinello Molecular Dynamics	1922
Machine Learning Molecular Dynamics	1923
Nonadiabatic Molecular Dynamics	1924
Ehrenfest Mean-Field Dynamics	1926
Surface Hopping	1927
Hybrid Quantum Mechanics/Molecular Mechanics (QM/MM)	1929
Applications to Large Biological Systems	1930
Photoactive Yellow Protein	1930
Rhodopsin	1931
Channelrhodopsin	1932
Organic Conjugated Materials	1933
Artificial Light-Harvesting Systems	1934
Software	1935
Newton-X	1935
Octopus	1936
CPMD	1936
Conclusion	1936
Bibliography	1937

Abstract

Excited-state molecular dynamics is a branch of theoretical physics and chemistry dedicated to simulating phenomena induced in molecules upon UV-visible

J. Rydzewski • W. Nowak (✉)
Institute of Physics, Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus
University, Toruń, Poland
e-mail: jr@fizyka.umk.pl; wiesiek@fizyka.umk.pl; Wieslaw.Nowak@fizyka.umk.pl

light absorption. This includes the determination of nuclei positions evolution in time along with the calculation of electronic excited-states potential energy surfaces. This chapter presents both adiabatic molecular dynamics and nonadiabatic molecular dynamics simulation methods as used in photochemistry. *Ab initio* molecular dynamics methods are briefly reviewed. A particular emphasis is on the recent developments in semiclassical algorithms and their applicability to studies of large molecular systems evolving in electronic excited states.

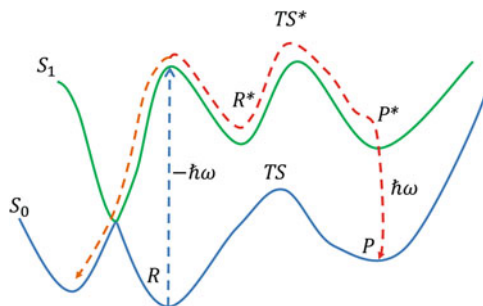
Introduction

Traditional molecular dynamics (MD), despite its great popularity, have few serious limitations. Among the most crucial ones is the treatment of nuclear evolution by classical equations of motion with the assumption that only a single electronic state, typically the ground state of a molecule, is occupied. In other words, MD is suitable for adiabatic processes (Marx and Hutter 2009). In classical MD, the interaction potential has to be known prior to the simulation, because the forces acting on nuclei have to be evaluated at each step of the numerical integration of equations of motions (Nowak 2012; Rydzewski et al. 2015). Despite these drawbacks (Remler and Madden 1990; Payne et al. 1992), traditional MD is well established as a tool to investigate many-body condensed matter systems, including biomolecules. The broadness and diversity of applications are documented in numerous monographs, articles, and reviews (Frenkel and Smit 2001; Sanbonmatsu and Tung 2007; Pérez et al. 2011; Orozco 2014; Perilla et al. 2015).

Gathering interaction potentials a priori is often difficult. One way to overcome this obstacle is by calculating the forces acting on the nuclei from electronic structure calculations performed “on-the-fly,” i.e., as the molecular trajectory is generated. This way, the electronic variables are considered as active degrees of freedom. For instance, in Born-Oppenheimer molecular dynamics nuclear degrees of freedom are propagated using ionic forces which are estimated at each iteration by approximately solving the electronic problem. In contrast – in the Car-Parrinello method the fictitious classical dynamics for electronic orbitals is explicitly introduced (Car and Parrinello 1985). These so-called *ab initio* MD calculations are much more computationally demanding, in contrast to traditional MD (Jasper et al. 2004a). Therefore, the study of full quantum dynamics of a polyatomic system is beyond computational possibilities, even in the next decades (Bornemann and Schütte 1998). A number of articles discussing *ab initio* MD have appeared over last years (Galli et al. 1991; Galli and Pasquarello 1993; Doltsinis and Marx 2002a; Grotendorst et al. 2002; Lorenz and Doltsinis 2012).

The interest in excited-state molecular dynamics simulations is growing, since many phenomena in modern science deal with electronic excited states: fluorescent microscopy, spectroscopy, labeling techniques, chemical lasers, photoactive enzymes, light-harvesting systems, optogenetics, photochemistry, just to name a few.

Fig. 1 Schematic comparison of nonadiabatic (orange) and adiabatic (red) reaction paths involving the ground state S_0 and first excited state S_1 . A nonadiabatic transition takes place through conical intersection (R reactants, TS transition state, P products, $*$ denoted excitation)



Among many applications of excited-state MD is computational photochemistry, which is described in a greater detail here, in order to introduce basic concepts and the formalism related to excited-state MD. Photochemical reactions can be divided into two categories: adiabatic and nonadiabatic (Fig. 1). Excited reactants R^* can overcome a potential barrier TS^* and turn into excited photoproducts P^* . After such a transition, P^* eventually return to their electronic ground state with an emission of a photon. In this process, a reaction path is on the same electronic excited state, which means that the reaction is adiabatic. On the other hand, a nonadiabatic reaction forms hot products in the electronic ground state by passing the conical intersection from the excited state to the ground state with no light emission. In other words, electronically nonadiabatic processes are defined as those in which the electronic state flips nonradiatively during dynamics (Fig. 1).

Therefore, excited-state MD is different from ground-state dynamics only when many electronic states must be considered simultaneously and can transfer a population among themselves. In such situations, the Born-Oppenheimer approximation breaks down, allowing nonadiabatic processes to take place. Nonadiabatic couplings (vibronic interactions) are well observed in molecules where two electronic states are degenerate. The nonadiabatic transition probabilities were introduced by Teller and Zener in 1930s (Zener 1932; Teller 1937). Teller also demonstrated the so-called conical crossings (also called molecular funnels or diabolic points) between potential energy surfaces in polyatomic molecules, which led to the formulation of the Jahn-Teller effect (Yarkony 1996).

Conical intersections are of high importance in a wide range of reactions, such as photoreactions, combustion reactions, and explosions (Bernardi et al. 1996; Baer and Billing 2003). Internal conversion at the conical intersections was showed by Zimmerman and Michl to be the fundamental feature to understand numerous photochemical mechanisms (Zimmerman 1966; Michl 1972a, b). In addition, the stability of DNA with respect to the UV irradiation is due to the presence of such conical intersection (Kang et al. 2002). These photochemical funnels provide pathways for ultrafast interstate crossing (in a femtosecond timescale).

This chapter presents selected main concepts and methods related to excited-state dynamics and is organized as follows. The theoretical formalism is presented in the sections “Adiabatic Molecular Dynamics” and “Nonadiabatic Molecular

Dynamics,” including the Born-Oppenheimer approximation and its limitations, the Ehrenfest dynamics, and surface hopping algorithms. In addition, the recently developed semiclassical Monte-Carlo approaches are also covered there. Selected practical calculations are presented in the **Applications** section, emphasizing advantages and drawbacks of the excited-state MD applicability to large molecular systems like proteins. The chapter ends with a **Conclusion** section, summarizing the current status of excited-state MD and pointing out the future prospects of this area of research.

Adiabatic Molecular Dynamics

Born-Oppenheimer Approximation

The Born-Oppenheimer (BO) approximation (Born and Oppenheimer 1927) is based on the assumption that the motions of nuclei and electrons in a molecule can be separated. It is usually introduced as an ansatz concerning the form of the molecular wavefunction. The BO approximation implies that the molecular wavefunction $\Psi(q, Q)$ is a product of the electronic and the nuclear factors:

$$\Psi(q, Q) = \psi(q; Q) \chi(Q) \quad (1)$$

where Q is a set of internal nuclear coordinates and q is a set of electronic coordinates. Here, $\chi(Q)$ is the nuclear wavefunction and $\psi(q; Q)$ is the electronic wavefunction, which also depends on fixed positions of the nuclei. The molecular Hamiltonian has the form (in the electrostatic approximation):

$$\mathcal{H} = -\sum_r \frac{\hbar^2}{2\mu_r} \frac{\partial^2}{\partial Q_r^2} - \frac{\hbar^2}{2m} \sum_i \frac{\partial^2}{\partial q_i^2} + V(q, Q) \quad (2)$$

where μ_r is a reduced mass for a r -th nucleus and m stands for the mass of an electron. One needs to remember that in Eq. 2 the Hamiltonian is simplified by neglecting the coupling between the rotation and the internal motions. By applying \mathcal{H} to the product $\psi(q; Q)\chi(Q)$, one gets:

$$\begin{aligned} \mathcal{H}\psi(q; Q) \chi(Q) &= \chi(Q) \left[V(q, Q) - \sum_i \frac{\hbar^2}{2m} \frac{\partial^2}{\partial q_i^2} \right] \psi(q; Q) \\ &\quad - \psi(q; Q) \left[\sum_r \frac{\hbar^2}{2\mu_r} \frac{\partial^2}{\partial Q_r^2} \right] \chi(Q) \\ &\quad - \sum_r \frac{\hbar^2}{2\mu_r} \left[2 \frac{\partial \psi}{\partial Q_r} \frac{\partial \chi}{\partial Q_r} + \chi \frac{\partial^2 \psi}{\partial Q_r^2} \right] \end{aligned} \quad (3)$$

The molecular Hamiltonian with the exclusion of the nuclear kinetic energy (electronic Hamiltonian) is the following:

$$\mathcal{H}_{el} = V(q, Q) - \frac{\hbar^2}{2m} \sum_i \frac{\partial^2}{\partial q_i^2} \quad (4)$$

$$\mathcal{H}_{el} \psi_k = U_k(Q) \psi_k \quad (5)$$

The eigenfunctions of the electronic Hamiltonian \mathcal{H}_{el} are so-called adiabatic wavefunctions. It is worth noting that, although \mathcal{H}_{el} does not contain differential operators with respect to the positions of nuclei Q , it depends on Q implicitly because its eigenvalues U_k and eigenfunctions ψ_k also depend on Q . The subscript k in the eigenfunctions and the corresponding eigenvalues is introduced to distinguish the different solutions of the Eq. 5. The ground states is for $k = 0$ and excited states are for $k > 0$.

Equation 3 now takes the form:

$$\begin{aligned} \mathcal{H} \psi_k(q; Q) \chi(Q) &= \psi_k(q; Q) \left[U_k(Q) - \sum_r \frac{\hbar^2}{2\mu_r} \frac{\partial^2}{\partial Q_r^2} \right] \chi(Q) \\ &\quad - \sum_r \frac{\hbar^2}{2\mu_r} \left[2 \frac{\partial \psi_k}{\partial Q_r} \frac{\partial \chi}{\partial Q_r} + \chi \frac{\partial^2 \psi_k}{\partial Q_r^2} \right] \end{aligned} \quad (6)$$

Finally, the definition of a nuclear Hamiltonian is following:

$$\mathcal{H}_{nuc} = U_k(Q) - \sum_r \frac{\hbar^2}{2\mu_r} \frac{\partial^2}{\partial Q_r^2} \quad (7)$$

$$\mathcal{H}_{nuc} \chi_{ku} = E_{ku} \chi_{ku} \quad (8)$$

where $U_k(Q)$ is an effective potential that constrains the nuclear motion. It is so-called the adiabatic potential energy surface (PES).

Now Eq. 3 can be written using the defined nuclear Hamiltonian:

$$\begin{aligned} \mathcal{H} \psi(q; Q) \chi_{ku}(Q) &= E_{ku} \psi_k(q; Q) \chi_{ku}(Q) \\ &\quad - \sum_r \frac{\hbar^2}{2\mu_r} \left[2 \frac{\partial \psi_k}{\partial Q_r} \frac{\partial \chi_{ku}}{\partial Q_r} + \chi_{ku} \frac{\partial^2 \psi_k}{\partial Q_r^2} \right] \end{aligned} \quad (9)$$

The BO approximation stands for neglecting the second term in Eq. 9:

$$\mathcal{H} \psi_k(q; Q) \chi_{ku}(Q) \approx E_{ku} \psi_k(q; Q) \chi_{ku}(Q) \quad (10)$$

Thus, Eq. 10 stands for the assumption that the $\psi_k \chi_{ku}$ vibronic states are approximate eigenstates of the molecular Hamiltonian.

The nonadiabatic couplings are introduced by the $H_{ku,lv}$ matrix elements:

$$H_{ku,lv} = \sum_r \frac{\hbar^2}{2\mu_r} \int_{-\infty}^{\infty} \chi_{ku}^* \left[t_{kl}^{(r)}(Q) + 2g_{kl}^{(r)}(Q) \frac{\partial}{\partial Q_r} \right] \chi_{lv} \dots dQ_s \dots \quad (11)$$

where

$$g_{kl}^{(r)} = \left\langle \psi_k \left| \frac{\partial}{\partial Q_r} \right| \psi_l \right\rangle \text{ and } t_{kl}^{(r)} = \left\langle \psi_k \left| \frac{\partial^2}{\partial Q_r^2} \right| \psi_l \right\rangle \quad (12)$$

The $g_{kl}^{(r)}$ terms do not couple different vibrational states that belong to the same electronic state. For $k = l$ in $t_{kl}^{(r)}$ a small correction is introduced to the U_k potential, which means that the approximate separation of the electronic and nuclear motions is not disrupted.

The BO approximation is based on neglecting all off-diagonal terms in nonadiabatic operators $g_{kl}^{(r)}$ and $t_{kl}^{(r)}$, i.e., neglecting interstate couplings, but keeping intrastate couplings between the electronic and nuclear terms (Eq. 12). The full adiabatic Born-Oppenheimer approximation is set by neglecting all diagonal terms too. In this approximation, the nuclei evolve on a single potential energy surface (either on the ground or the excited state), while the electrons adiabatically adjust to the nuclei.

Car-Parrinello Molecular Dynamics

The Car-Parrinello method for ab initio MD (Car and Parrinello 1985) has gained great popularity in recent years (Marx and Hutter 2009; Brunk and Rothlisberger 2015). It avoids expensive and explicit minimization of energy functionals required by the density functional theory (DFT) in the context of the Born-Oppenheimer approximation. As an alternative, it introduces a fictitious classical dynamics for the electronic orbitals (Car and Parrinello 1985). Car and Parrinello replaced the adiabatic motion of the electrons by a fictitious classical Lagrangian dynamics that oscillates around the energy minimum of the potential energy surface. In that sense, by introducing these oscillations again, the Car-Parrinello technique “undoes” the adiabatic Born-Oppenheimer approach that substituted the fast quantum-dynamically oscillatory behavior of the electronic wavefunction by the average of their motion around the minimum of the potential energy surface (Fig. 2). The fictitious classical dynamics is given by the following Lagrangian:

$$\mathcal{L}_{CP} = \frac{\mu}{2} \sum_{k=1}^N \dot{\psi}_k \cdot \dot{\psi}_k + \frac{1}{2} \sum_i M_i \dot{q}_i^2 - E_{KS}(\psi, q) + \sum_{k=1}^N \sum_{l=1}^N \Lambda_{kl} (\langle \psi_k, \psi_l \rangle - \delta_{kl}) \quad (13)$$

Here, the bracket $\langle \cdot, \cdot \rangle$ stands for the scalar product, ψ_k are the wavefunctions regarded as classical fields, M_i is the i -th ionic mass and, finally μ is the fictitious

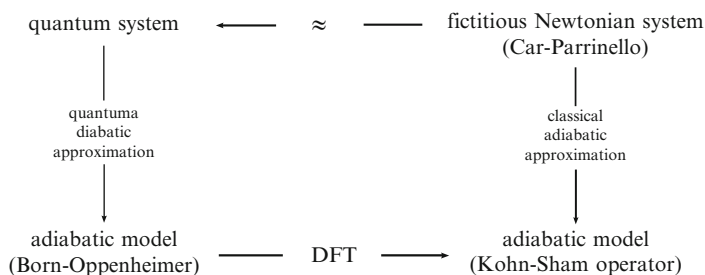


Fig. 2 Logic of the Car-Parrinello method (Figure based on Bornemann and Schütte (1998))

electronic mass parameter. The dot denotes differentiation with respect to time. The orthonormality of the given wavefunctions ψ_k and ψ_l is ensured by introducing the Lagrange parameters Λ_{kl} .

The study of Pastore et al. (1991) proves that the fictitious mass parameter μ (usually in the range of 300–800 a.u.) constitutes a kind of a control parameter. The bigger μ is chosen, the bigger the deviation of the Car-Parrinello method from the adiabatic model is. More importantly, μ introduces a time scale of order $\sqrt{\mu}$. This requires time steps in a simulation to be proportional to $\sqrt{\mu}$. For that reason, the Car-Parrinello quantum dynamics is restricted to time steps even order of magnitude smaller than the BO adiabatic dynamics, thus it is more computationally demanding (Kühne et al. 2007). One has to establish a compromise between the computational costs and accuracy. Moreover, it is worth remembering that maintaining adiabaticity requires that the fictitious mass of electrons μ is chosen small enough to suppress a significant energy transfer from the ionic to the electronic degrees of freedom.

It is worth mentioning that the nonadiabatic Car-Parrinello molecular dynamics is also possible by coupling the restricted open-shell Kohn-Sham (ROKS) excited state to the ground state using a surface hopping scheme (Doltsinis and Marx 2002b).

Ab initio molecular dynamics is not bound to any particular electronic structure method. It is obvious that advantages and drawbacks of a particular ab initio MD are strongly connected to the underlying approximate electronic structure algorithm chosen. Over the past years, a vast spectrum of different electronic structure methods became well established, i.e., DFT, Hartree-Fock, generalized valence bond, complete active space self-consistent field (CASSCF), second-order many-body perturbation theory (MP2), or full configuration interaction (FCI), all of them may be exploited in ab initio MD simulations, but the computational cost will vary enormously, especially for larger systems.

Machine Learning Molecular Dynamics

More recent approaches to calculate force fields “on-the-fly” are based on machine learning. They use neural networks (Behler and Parrinello 2007), Gaussian

processes (Bartók et al. 2010), or the Bayesian interference (Li et al. 2015) to fit the potential energy surface, after which atomic forces are obtained by analytical differentiation (Li et al. 2015). Such implementations are not formally limited to a small number of atoms; thus, open a way for ab initio molecular dynamics of proteins. In machine learning MD, simulating dynamics is not dependent on static force fields, i.e., predefined parametrized functional forms fitted on experimental data (Laio et al. 2000). One should notice that the mentioned first principles techniques are restricted nowadays only to one potential energy surface; therefore, this kind of an approach is currently feasible only to adiabatic processes. Moreover, some of these techniques require a training set to enable learning procedure during MD simulations. The training sets can be collected to optimize an efficiency of the learning procedure either by performing the classical MD (Karplus and Petsko 1990), Monte Carlo (Duane et al. 1987), or metadynamics (Laio and Parrinello 2002; Ensing et al. 2006; Laio and Gervasio 2008) simulations.

Nonadiabatic Molecular Dynamics

For a large number of physical/chemical problems, the Born-Oppenheimer approximation can be applied with a great success. On the other hand, to properly describe phenomena beyond the BO approximation one needs to introduce the nonadiabatic couplings in MD simulations. These include, for instance, charge transfer and photoisomerization reactions. According to Barbatti and Crespo-Otero (2014) semiclassical MD methods suitable for excited states should incorporate the following criteria:

1. Be computationally nondemanding
2. Be reliable and accurate in terms of excited states of different nature
3. Be independent from human intervention and preferably not system dependent
4. Provide a computational algorithm for calculating electronic structures near intersection seams with the ground state and between the excited states

The simplest approach to include the nonadiabatic terms in MD is to treat nuclear motions by classical mechanics. In such approximation, the atomic nuclei follow a given classical trajectory, while the electronic motions are described in terms of time-dependent total wavefunction that satisfies the time-dependent electronic Schrödinger equation. The semiclassical approach is often classified as trajectory ensemble methods, where the nuclear motion of the system is simulated by a swarm of classical trajectories (Tully and Miller 1976; Tully 1998; Hack and Truhlar 2000; Jasper et al. 2004b). Such semiclassical algorithms are popular, because their only quantum requirement is that electronic structure properties have to be computed “on-the-fly.” In other words, the semiclassical methods do not require the knowledge of the full configurational space – PES is probed only along the classical motion of the nuclei. The necessary electronic structure part can be approximated

using various DFT methods, such as ROKS, time-dependent DFT (TDDFT), or wavefunction-based approaches.

For the two-state system, the nonadiabatic couplings can be examined by the so-called Massey parameter (Desouter-Lecomte and Lorquet 1979; Gersonde et al. 1994; Nikitin and Zülicke 1978):

$$\xi = \frac{\Delta E l}{\hbar \dot{R}} \quad (14)$$

where l is the characteristic length over which wavefunctions corresponding to states change significantly. For value $\xi \gg 1$ (large energy gaps ΔE and small nucleic velocities \dot{R}), the nonadiabatic effects in the system can be neglected. In the Born-Oppenheimer approximation, the electron cloud has to adjust instantly to changes between the nuclear distances. The frequency of electronic motions ω can be calculated by dividing a corresponding energy gap ΔE by \hbar :

$$\Delta E = \hbar \omega \quad (15)$$

The frequency of electronic oscillations is high when the nuclei are close to each other, i.e., where a distance between the nuclei R is small. It is clear that such distortions in the electronic clouds will influence the nuclear trajectory. Thus, these are the reasons for the significance of the nonadiabatic effects in the discussed two-state system.

The general expression for measuring nonadiabatic effects for two states k and l of a particular system can be derived from the Hellman-Feynman theorem (Feynman 1939; Hellmann 1937):

$$\left\langle \psi_k \left| \frac{\partial}{\partial Q_r} \right| \psi_l \right\rangle = \frac{\left\langle \psi_k \left| \frac{\partial \mathcal{H}_{el}}{\partial Q_r} \right| \psi_l \right\rangle}{U_l - U_k} \quad (16)$$

The nonadiabatic couplings between two states k and l are larger where their PESs get closer to each other, thus $\Delta U = U_l - U_k \rightarrow 0$. It is worth noticing that ΔU cannot be 0 for diatomic molecule systems, because two distinct solutions cannot be satisfied for the same bond length R , which means that eigenvalues corresponding to different states cannot cross (cannot become an equal value). This is sometimes called an “avoided crossing” or “noncrossing rule.”

Switching from a nonadiabatic state to the other one by a single “hop” means staying in the same adiabatic state. The probability of hopping between distinct adiabatic states $P_{a_1 \rightarrow a_2}^d$ may be estimated from the so-called Landau-Zener formula, which was formulated separately in 1932 by L. Landau (1932), C. Zener (1932), E. Stueckelberg (1933), and E. Majorana (1932):

$$P_{a_1 \rightarrow a_2}^d \approx \exp\left(-\frac{\pi^2 \Delta E^2}{\hbar v \Delta S}\right) \quad (17)$$

Here, a_1 and a_2 are the chosen adiabatic states, ΔE is the energy gap between the two potential energy surfaces at the geometry of the closest approach, ΔS is the slope difference in the region of the avoided crossing and v is the nuclear velocity along the reaction coordinate. Therefore, the probability of a “hop” from one adiabatic BO surface to another increases with an increasing difference in the surface slopes and increasing velocity of the nuclear motion but also with decreasing energy gap ΔE . The less avoided the crossing, the larger the jump probability. Some early studies of excited-state dynamics in heme proteins (Li et al. 1993; Cottone et al. 2012) and photoactive enzymes (Kubiak and Nowak 2008) employed this formula.

Ehrenfest Mean-Field Dynamics

Incorporating vibronic properties is of high importance in all systems, where the nonadiabatic effect should not be neglected. In the plain simple semiclassical approximation, the total wavefunction $\Psi(q, t)$ consists of a linear combination of adiabatic eigenfunctions $\psi_k(q, Q)$ (Eq. 5):

$$\Psi(q, t) = \sum_k C_k(t) \psi_k(q, Q) \exp\left(-\frac{i}{\hbar} \int U_k(Q) dt\right) \quad (18)$$

Here, $|C_k(t)|^2$ is the probability of the system being in the adiabatic state k at time t . The adiabatic potential energy surface U_k can be replaced by the energy expectation value:

$$U_{eff} = \langle \Psi | \mathcal{H}_{el} | \Psi \rangle = \sum_k |C_k(t)|^2 U_k \quad (19)$$

So now, all the atoms experience the effective potential U_{eff} , which represents an average over the adiabatic potential energy surfaces multiplied by their expansion coefficients $|C_k(t)|^2$. Molecular dynamics with above substitution is called the Ehrenfest dynamics, and because the idea is to perform an average over adiabatic PESs, it is also referred to as mean-field dynamics (Ehrenfest 1927). An effective potential is most appropriate for strong vibronic interactions but leaves the system in an unphysical final state that corresponds to a quantum superposition of adiabatic states. Although, this method has been successfully used to study some chemical problems, i.e., excitation lifetimes of organic molecules (Liang et al. 2010), on the other hand it suffers from serious drawbacks. When the difference between the adiabatic states is large, then the dynamics must be primarily driven by only one surface and not an average potential. Moreover, a system, which is initially in a pure adiabatic state, after a transition through strong vibronic coupling, will be in a mixed state. The most significant issue is that the pure adiabatic character of a state cannot be recovered in the configuration space even asymptotically. Yet another problem is that the Ehrenfest methods violate microscopic reversibility (Grotendorst et al. 2002).

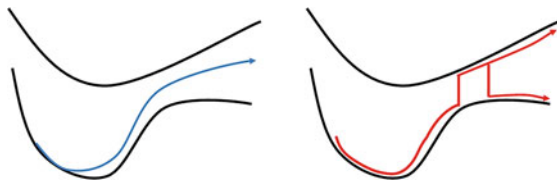


Fig. 3 Schematic comparison of two semiclassical molecular dynamics algorithms: the Ehrenfest mean-field method (*left, blue* – single trajectory) and Tully’s surface hopping method (*right, red* – swarm of trajectories) governed by two potential energy surfaces

Surface Hopping

The quantum wavepacket evolution in time can be approximated by a swarm of semiclassical trajectories propagating on adiabatic Born-Oppenheimer potential energy surfaces of multiple electronic states. The nonadiabatic events are represented by a stochastic algorithm which allows each trajectory to “hop” between other states during the time evolution. In this class of methods, the wavepacket probability distribution should be approximated by the statistics over an ensemble of trajectories in terms of a fraction in each electronic state at any time step (Fig. 3). This method, in contrast to the Ehrenfest dynamics, does not facilitate any averaged adiabatic potentials. At any moment in time, the system is propagated using one pure adiabatic potential energy surface that is selected according to its state population.

The time-dependent electronic state is following:

$$\Psi^{(r)}(q, Q, t) = \sum_k C_k^{(r)}(t) \psi_k^{(r)}(q; Q) \quad (20)$$

where the subscript (r) is used to emphasize the dependence on the trajectory. After taking Eq. 19 to the time-dependent Schrödinger equation, a set of coupled equations of motion for the complex nuclear state amplitudes $C_k^{(r)}(t)$ emerges:

$$i\hbar \dot{C}_k^{(r)}(t) = \sum_k C_k^{(r)}(t) \left\{ U_k(Q^{(r)}) \delta_{ij} - i\hbar \sum_\gamma g_{jk}^{(\gamma)}(Q^{(r)}) \dot{Q}_\gamma^{(r)} \right\} \quad (21)$$

where the first order nonadiabatic couplings $g_{jk}^{(\gamma)}$ are given by Eq. 12. These coupled equations can be solved along a nuclear classical trajectory r , evolving adiabatically in a given electronic state j . The given trajectory r can “hop” from the state j to k during the time interval $[t, t + dt]$ with the probability:

$$p_{jk}^{(r)}(t, t + dt) = -2 \int_t^{t+dt} \Re \left[\frac{C_k^{(r)}(\tau) C_j^{(r)*}(\tau) \kappa_{kj}^{(r)}(\tau)}{C_j^{(r)}(\tau) C_j^{(r)*}(\tau)} \right] d\tau \quad (22)$$

Here, $\kappa_{kj}^{(r)}(\tau) = \sum_{\gamma} g_{ij}^{(\gamma)}(Q^{(r)}) \dot{Q}_{\gamma}^{(r)}$. A stochastically driven jump between two distinct potential energy surfaces $j \rightarrow k$ takes place when a following condition is fulfilled:

$$\sum_{l \leq k-1} p_{jl}^{(r)} < \varrho < \sum_{l \leq k} p_{jl}^{(r)} \quad (23)$$

where $\varrho \in [0, 1]$ is a randomly generated number. This algorithm is a modification of the traditional surface hopping method (Tully and Preston 1971; Tully and Miller 1976), and it guarantees minimum number of jumps along each adiabatic trajectory. Thus, it is referred to as the “fewest switches surface hopping algorithm” (Tully 1990, 1998). This algorithm has emerged as a popular platform for performing excited-state MD over the recent years (Hudock et al. 2007; Hazra et al. 2010; Jiang and Sibert III 2012).

According to Tully, surface hopping algorithms suffer from the three major inconsistencies between quantum and classical mechanics (Tully 2012). The foremost issue in semiclassical molecular dynamics comes from discrepancies in feedback between the quantum and classical degrees of freedom. The reason for this is that the motions of particles treated classically imply changes of the Hamiltonian which govern the quantum particles and, therefore, drive nonadiabatic transitions. On the other hand, such changes of a quantum state have impact on the classical motion. It is so-called quantum back-reaction.

Yet another important divergence of surface hopping algorithms from the theory can be attributed to the quantum decoherence (Hammes-Schiffer and Tully 1994). The total wavefunction for a system and its bath have a phase to which contribute both fast and slow degrees of freedom. In the case of taking the classical limit for the slow degrees of freedom, the phase information corresponding to this part of system is gone. Unfortunately, this lost phase affects the branching ratios of the trajectories. The interference between the wavefunctions of the slow degrees of freedom for the different possible trajectories exists only when time is shorter than the decoherence times. Interesting idea for handling the quantum decoherence problem is the simultaneous-trajectory surface hopping (Shenvi et al. 2011).

The last concern with treating some degrees of freedom by the laws of classical mechanics and others with quantum mechanics is properly handling the detailed balance condition. Accordingly, the quantum system driven by finite temperature classical system will in limit approach an infinite temperature. This is a well-known issue, since the energy passes from the classical system to quantum system artificially. Besides the surface hopping algorithms, which handle detailed balance approximately, the Ehrenfest dynamics does not solve this problem in any form. This approximately corrects detailed balance results from frustrated surface hops (Fig. 4). A hop is frustrated if the system does not have enough energy to make the hop. The simplest solution suggested to deal with these frustrated hops is to ignore them (Tully 1990). Another solution is not to switch the adiabatic state but

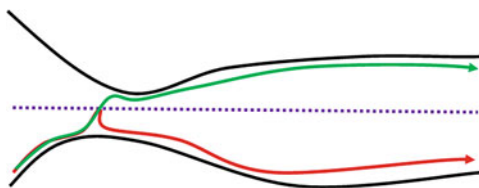


Fig. 4 Schematic comparison of the surface hopping trajectories in case of restricted total energy level (*violet line*): (*red*) omitting frustrated hop, if energy conservation condition is not fulfilled (Tully 1990); (*green*) allowing frustrated hop, even if such condition does not stand but Jasper condition stands (Jasper et al. 2002)

instead reverse the direction of the component of the velocity along the nonadiabatic coupling vector (Hammes-Schiffer and Tully 1994). One more suggestion proposed by Jasper et al. is to allow the transition to take place if an allowed hopping point is reachable within uncertainty time $\delta t = \hbar/2\Delta E$, where ΔE is the extra energy that the system needed to make the hop possible (Jasper et al. 2002).

Another recent modification of the fewest switches surface hopping was proposed in Gorshkov et al. (2013) and White et al. (2014). A part of this algorithm is similar to previous surface hopping methods. A swarm of classical trajectories is propagated, allowing the system to jump from a current pure adiabatic potential energy surface to the other state. The hopping rate is, however, arbitrary and the “hop” occurs only if the energy balance condition is fulfilled. Next, the classical trajectories are classified into groups according to their final electronic states and the numbers of jumps that occurred during their course. The gained data is used later in Monte-Carlo sampling to retain phase information and the hopping rates at the moments of hops.

Hybrid Quantum Mechanics/Molecular Mechanics (QM/MM)

The quantum mechanical approach to dynamics is much more computationally demanding than classical mechanics, thus it is not directly applicable for large biological molecules. The key idea to overcome this limitation was proposed by Warshel and Levitt (1976). The main assumption in that method is to subdivide a complex system in such a way that only a small, nevertheless chemically active region is treated quantum mechanically. The rest of the system is simulated with empirical potentials of the classical MD. In this way, even huge biological systems can be treated efficiently in terms of computational resources and for longer simulation times. This is the so-called QM/MM approximation and since its postulation the method has been applied to various complex biomolecules (Virshup et al. 2008; Beierlein et al. 2011; Shim et al. 2012; Brunk and Rothlisberger 2015).

The QM/MM methods were combined with both adiabatic and nonadiabatic quantum calculation methods (Brunk and Rothlisberger 2015).

The hybrid Hamiltonian takes the form:

$$\mathcal{H}^{QM/MM} = \mathcal{H}^{QM}(q) + \mathcal{H}^{MM}(Q) + \mathcal{H}^{QM-MM}(q, Q) \quad (24)$$

Here, \mathcal{H}^{QM} denotes the ab initio Hamiltonian, \mathcal{H}^{MM} is the classical Hamiltonian, which is described by a standard force field

$$\mathcal{H}^{MM} = \mathcal{H}_b^{MM}(Q) + \mathcal{H}_{nb}^{MM}(Q) \quad (25)$$

where \mathcal{H}_b^{MM} is the bonded term, which takes into account stretching, bending, and torsional terms, whereas \mathcal{H}_{nb}^{MM} is the nonbonded term, and it includes the van der Waals, hydrogen bonds, and electrostatic interactions. The major drawback of the QM/MM approximation is calculating the coupling term of the full Hamiltonian, \mathcal{H}^{QM-MM} . For instance, in QM/MM Car-Parrinello methods, the van der Waals interactions between QM and MM regions are described by the classical force field. The same assumption holds for the bond angles bending and dihedral distortions. In contrast, bonds between the QM and MM regions are treated with monovalent pseudopotentials.

Applications to Large Biological Systems

The most widely used approach to large biological systems, including proteins, membranes, and nucleic acids, is classical MD, dealing with a single adiabatic PES. Although, the MD method itself is very efficient, it is insufficient in describing numerous phenomena important for biological functions such as electronic excitations, bond formation and breaking, and charge transfer. Ab initio MD is, to some extent, applicable to the photophysical processes in structural biology and some representative examples of such studies are outlined here, together with an example from photoactive material science.

Photoactive Yellow Protein

Photoactive yellow protein (PYP) is a small protein that governs negative phototropism in bacteria *Halorhodospira halophila*. It contains p-coumaric acid covalently bound to the protein as a chromophore. It is a very well-characterized bacterial photoreceptor with a simple photocycle and some features in common with other photoactive proteins. The CASSCF (3-21G)-QM/MM method was applied to observe spontaneous isomerization of PYP (Groenhof et al. 2004).

This study has demonstrated that QM/MM MD with explicit surface hopping from the ground to the first excited-state can reproduce excited-state lifetimes in PYP acquired from experiments. Moreover, it helped to reveal the molecular mechanism responsible for the activity of the protein matrix on the PYP chromophore.

In another study by Isborn et al. (2012), QM/MM molecular dynamics simulations were used for computation of the absorption spectra of PYP in vacuo, aqueous solution, and protein environments. It was shown that the computed excitation energies of PYP require a QM region to be very large in order to achieve an agreement with the protein system treated entirely by quantum dynamics. The QM region consisting of the PYP chromophore, covalently linked to PYP, the layer of PYP, and water molecules, was treated with TDDFT. The dynamics of the QM region was simulated with BO dynamics. The rest of the PYP system was subject to the empirical force field. The interaction energy between the QM and MM was given by:

$$\begin{aligned}
 E^{QM-MM} = & \sum_{J=1}^{N_{QM}} \sum_{k=1}^{N_{MM}} \frac{Z_J q_k}{|\vec{r}_J - \vec{r}_k|} - \sum_{k=1}^{N_{MM}} \int q_k \frac{\rho(\vec{r})}{|\vec{r} - \vec{r}_k|} d\vec{r} \\
 & + \sum_{J=1}^{N_{QM}} \sum_{k=1}^{N_{MM}} \epsilon_{Jk} \left[\left(\frac{\sigma_{Jk}}{|\vec{r}_J - \vec{r}_k|} \right)^{12} - \left(\frac{\sigma_{Jk}}{|\vec{r}_J - \vec{r}_k|} \right)^6 \right] \quad (26)
 \end{aligned}$$

where the first term of Eq. 26 is the Coulomb interaction between the QM nuclei of charge Z_J with the MM charges q_k , the second term – the Coulomb interaction of the QM electron density $\rho(\vec{r})$ with the MM charges q_k , and the last term is the Lennard-Jones interaction of the QM atoms with the MM atoms to account for the van der Waals forces, where ϵ and σ are the interaction energy and interatomic distance for each pair of atoms, respectively (Isborn et al. 2012).

Isborn et al. observed mismatches between all the computed absorption spectra of PYP and experimental results. The authors listed possible reasons concerning this discrepancy: (i) treatment of errors by the TDDFT functional for various electronic environments, (ii) too few QM water molecules in QM/MM technique, and (iii) inaccurate molecular dynamics formulas regarding the interface between QM and MM regions (Eq. 26). Therefore, for both the ground and excited states of PYP it was suggested to include the majority of the simulated system in the QM region.

Rhodopsin

Rhodopsin is a highly specialized G protein-coupled receptor indispensable for transduction of light signals to the form of neuronal impulses. The protonated Schiff base of 11-*cis*-retinal, which is a chromophore of rhodopsin, upon absorption of light isomerizes to all-*trans* in an efficient process that takes 200 fs (Fig. 5).

Detailed mechanistic behavior of 11-*cis* to all-*trans* retinal isomerization was investigated by many hybrid molecular dynamics methods, including the QM/MM Car-Parrinello approach with ROKS of the first excited singlet state (Röhrig et al. 2004) and for a model π -electron system embedded in a real protein environment

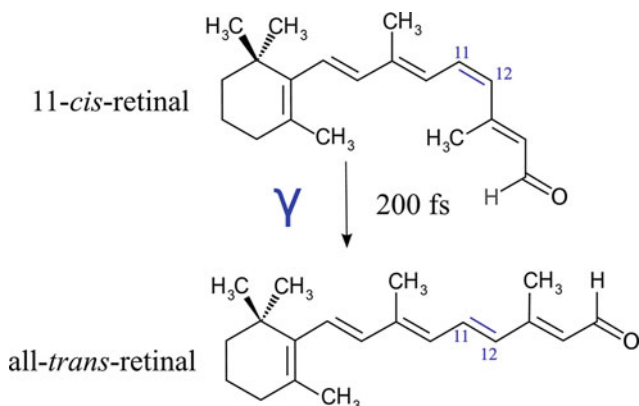


Fig. 5 Isomerization of 11-*cis*-retinal to all-*trans*-retinal upon light excitation (Scheme prepared by J. Rydzewski, NCU, Poland)

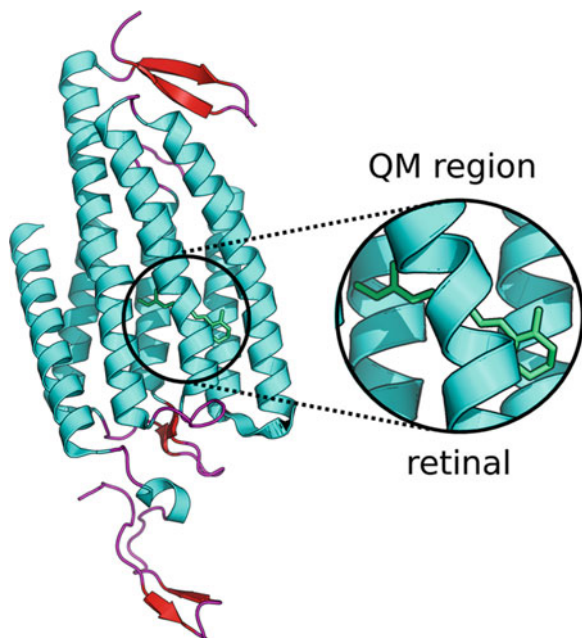
at the advanced CASSCF level (Hayashi et al. 2003). Moreover, the conical intersection present in bacteriorhodopsin has been studied computationally and verified experimentally by ultrafast femtosecond spectroscopy with the excellent agreement between both methods (Polli et al. 2010).

Channelrhodopsin

Channelrhodopsins (ChRs, Fig. 6) are light-gated cation channels derived from algae that have shown significant utility in optogenetics. Neurons expressing ChRs can be controlled via light event in freely moving mammals (Deisseroth 2011; Kato et al. 2012). In terms of structure, ChRs have seven transmembrane part and contain the light-isomerizable chromophore all-*trans*-retinal (Fig. 6). To optimize ChRs functions, neuroscientists engineered a variety of its mutants (Lin 2011). Their success is often based on trial-and-error, rather than conclusions drawn from mechanistic insights. The exact atomistic behavior of ChRs would be of high significance in the understanding of newly emerging research field – optogenetics. Excited state molecular dynamics simulations are the perfect tools to achieve that since the structural manipulations can be examined on the atomistic level.

In the recent paper, K. Welke et al. proposed a computational approach for engineering color-shifted (characterized by a shifted absorption maximum) optimized ChRs variants acquired from homology modeling of a bacteriorhodopsin template (Welke et al. 2011). The aim of their study was to enable the search for possible targets for mutations that shift the absorption maximum. The impact from amino acid mutations near the ChR retinal binding pocket was described by the use of the QM/MM approach with the chromophore of ChR treated with the CASSCF method. The rest of the system, including ChR, was simulated by the classical molecular

Fig. 6 Schematic representation of the QM/MM approach. The whole biological system (channelrhodopsin, membrane, solvate, retinal) is divided into two regions: QM – surrounding the retinal photoactive ligand and MM – concerning the rest of the system. Visualized using Pymol (Figure prepared by J. Rydzewski, NCU, Poland)



dynamics techniques. They concluded that the large hypsochromic shift can be traced back to two origins. First, while the single mutations in the ChR binding pocket effect on a color shift is rather small, a combination of mutations covers the majority of the hypsochromic shift. Second, the other significant contribution is a stronger interaction of the retinal with its counterions. Another study regarding ChRs showed that in the visible region the circular dichroism spectrum is dominated by the exciton coupling between the retinal chromophores in the two subunits of ChR (Pescitelli et al. 2014). In that study, the QM/MM model with QM part of the system treated with TDDFT reproduces the transition energies for ChR better than any other models examined therein. The authors are not aware of any published research concerning explicitly the excited state quantum dynamics of the ChRs proteins. Filling this gap perhaps makes a perfect opportunity for understanding in the near future the time evolution of optogenetic molecular systems.

Organic Conjugated Materials

Photoactive materials require understanding of their electronic properties in order to control their photo-induced pathways. Especially, irradiative intraband relaxation of organic conjugated materials (OCMs) can lead to various complex processes challenging to model, since such processes alleviate breakdown of the

BO approximation. In the paper by T. Nelson et al. it was showed that Tully's fewest switching surface hopping algorithm can be successfully applied to large polyatomic systems involving many excited states and multiple unavoided crossings between potential energy surfaces (Nelson et al. 2014). It was proved that the highly efficient intramolecular ultrafast energy transfer in poly-(phenylene ethynylene) dendrimers, modeled by two-, three-, and four-ring PPE chromophores, is consistent with experimental measurements. Moreover, a method that allows the analysis of the distribution of radiationless transitions due to cross and hop and proper dealing with so-called trivial unavoided crossings of the potential energy surfaces has been developed (Fernandez-Alberti et al. 2012). The term "trivial" unavoided crossing describes the situation when the nonadiabatic couplings take a form of sharp peaks.

Artificial Light-Harvesting Systems

In artificial photosynthetic devices, the efficient conversion of light into electricity is believed to occur on femto- and picosecond timescales. Up to now, it was assumed that this conversion involves an incoherent electron transfer process. However, in the recent studies, there have been indications showing that the coherent motion of electrons is one of the most important stages during the conversion process (Engel et al. 2007) in photosynthetic bacteria and higher plants (Lee et al. 2007). Rozzi et al. studied quantum dynamics of a photo-excited carotene-porphyrin-fullerene triad and provided evidence that the electron energy transfer from the porphyrin moiety to fullerene is strongly correlated with wavelike motions of electrons and nuclei on timescales not exceeding hundreds of femtoseconds. Quantum dynamics was performed by combining TDDFT and the Ehrenfest mean-field approach (Alonso et al. 2008; Andrade et al. 2009). The initial photoexcitation of the molecule was modeled by a rapid perturbation in the charge density localized on the porphyrin moiety. Next, the evolution of this perturbation and its effects on the charge density on the rest of the triad were monitored in time for two different simulation scenarios: quenched ionic positions in specific regions of the molecule and nuclei not restricted to move (Fig. 7). Therefore, determination of the effect of quantum dynamics of the system on the electron transfer was possible.

The results of quantum dynamics simulations of Rozzi et al. indicate a rapid charge density flow to the fullerene moiety within approximately 70 fs, but only in case of freely moving nuclei. The initial perturbation in form of the electronic wavepacket oscillates between the carotene-porphyrin group and the fullerene moiety. It was confirmed that the ionic movement drives the charge density transfer in the molecule. By quenching all the nuclei so they do not move during the quantum dynamics, the charge transfer was greatly suppressed.

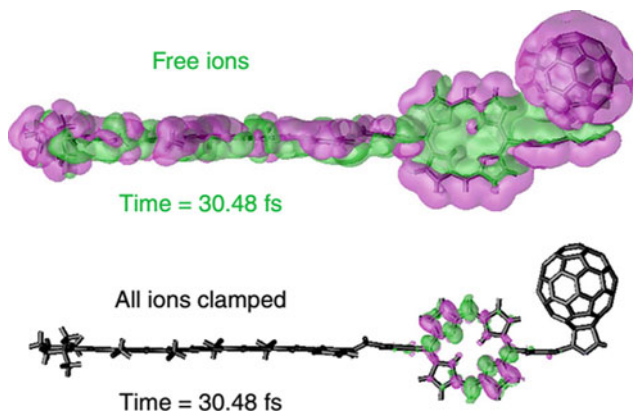


Fig. 7 Frames of the photo-induced perturbation used to model oscillations in charge density at a time of 30.48 fs the nuclei not restricted to change positions and all ions clamped. An increase in electronic charge density is colored in *magenta*, whereas a decrease is shown in *green* (Reprinted with permission of Nature Publishing Group from “Quantum coherence controls the charge separation in a prototypical artificial light-harvesting system” *Nat. Commun.* 2013, 4, 1602)

The given example shows a great importance of the quantum dynamics methods, i.e., the Ehrenfest method combined with TDDFT as a complementary tool to the ultrafast spectroscopy in probing the quantum behavior of the artificial photosynthetic devices.

Software

There are several programs suitable for performing quantum dynamics simulations, with a particular emphasis on nonadiabatic molecular dynamics or adiabatic molecular dynamics, capable of performing excited state dynamics on a single excited potential energy surface. An arbitrary list of newly emerging as well as established codes, which are appropriate for such dynamics studies for large biological systems, is presented below.

Newton-X

Newton-X (www.newtonx.org) implements a semiclassical surface hopping algorithm (see section “[Surface Hopping](#)”). It provides all steps of the nonadiabatic molecular dynamics simulation, i.e., generation of the initial conditions, computing the semiclassical trajectories and a statistical analysis of results. It must be interfaced with a program, which can calculate the potential energy surfaces by solving the Schrödinger equation for electrons, for instance: Gaussian, Turbomole,

or Gamess. Usually, the surface hopping method uses MCSCF, CC2, MRCI, or TDDFT techniques. Embedding the system in the molecular mechanics environment via the QM/MM approach is also possible.

Octopus

Octopus (www.tddft.org) employs norm-conserving pseudopotentials and real-space grid discretization (no basis sets) for performing Kohn-Sham TDDFT in real time. It can perform molecular dynamics simulations with use of Born-Oppenheimer dynamics, Ehrenfest dynamics, and Car-Parrinello dynamics. The only way to incorporate the nonadiabatic effects into the system of interest is by the Ehrenfest dynamics for nuclei, since the system evolves using a mean-field approximation. Octopus is capable of treating not only systems in three dimensions, but also in 2D and 1D for low-dimensional solids (Andrade et al. 2015; Castro et al. 2006; Marques et al. 2003).

CPMD

The CPMD program (www.cpmd.org) is an implementation of DFT based on plane waves and norm-conserving or ultra-soft pseudopotentials, particularly dedicated for both the adiabatic and nonadiabatic Car-Parrinello dynamics (see section “[Car-Parrinello Molecular Dynamics](#)”). It enables simulations of excited state molecular dynamics by the means of TDDFT (either using the density functional theory-based Ehrenfest approach or the linear response approximation to time-dependent Kohn-Sham density functional theory). It is suitable also for the hybrid QM/MM calculations with the use of the GROMOS force field. Tully’s fewest hopping algorithm is also implemented to simulate a complete swarm of classical trajectories on the potential energy surfaces. Moreover, the efficient Landau-Zener-Stückelberg theory of nonadiabatic transition has been implemented as well.

Conclusion

Excited state molecular dynamics as a tool for simulating molecular phenomena induced upon interaction with light is becoming more and more popular due to increasing computer power, implementations of new methods, and algorithm improvements. Some pilot applications in studies of biomolecules have appeared; nevertheless, there is already “plenty of space above the ground” for new applications. For instance, one can imagine the protein-ligand photo-dissociation process (Rydzewski and Nowak 2015) with a ligand capable of switching between the ground and excited potential energy surfaces during molecular dynamics simulations. Excited state molecular dynamics simulations, perhaps, will be soon mature enough for application to even larger molecular systems and more complex light-

induced phenomena in the near future. New application areas arise as biological phenomena reveal more and more their quantum nature (Arndt et al. 2009; Lambert et al. 2013).

Acknowledgements JR acknowledges a partial support from KwPV fellowship. The ICNT NCU infrastructure was used to prepare this chapter.

Bibliography

- Alonso, J. L., Andrade, X., Echenique, P., Falceto, F., Prada-Gracia, D., & Rubio, A. (2008). Efficient formalism for large-scale ab initio molecular dynamics based on time-dependent density functional theory. *Physical Review Letters*, *101*, 096403.
- Andrade, X., Castro, A., Zueco, D., Alonso, J. L., Echenique, P., Falceto, F., & Rubio, A. (2009). Modified Ehrenfest formalism for efficient large-scale ab initio molecular dynamics. *Journal of Chemical Theory and Computation*, *5*, 728–742.
- Andrade, X., Strubbe, D., De Giovannini, U., Larsen, A. H., Oliveira, M. J., Alberdi-Rodriguez, J., Varas, A., Theophilou, I., Helbig, N., & Verstraete, M. J. (2015). Real-space grids and the Octopus code as tools for the development of new simulation approaches for electronic systems. *Physical Chemistry Chemical Physics*, *17*, 31371.
- Arndt, M., Juffmann, T., & Vedral, V. (2009). Quantum physics meets biology. *Human Frontier Science Program Journal*, *3*, 386–400.
- Baer, M., & Billing, G. D. (2003). *Advances in chemical physics, the role of degenerate states in chemistry*. New York: Wiley.
- Barbatti, M., & Crespo-Otero, R. (2014). *Surface hopping dynamics with DFT excited states*. Berlin/Heidelberg: Springer.
- Bartók, A. P., Payne, M. C., Kondor, R., & Csányi, G. (2010). Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Physical Review Letters*, *104*, 136403.
- Behler, J., & Parrinello, M. (2007). Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical Review Letters*, *98*, 146401.
- Beierlein, F. R., Michel, J., & Essex, J. W. (2011). A simple QM/MM approach for capturing polarization effects in protein–Ligand binding free energy calculations. *Journal of Physical Chemistry B*, *115*, 4911–4926.
- Bernardi, F., Olivucci, M., & Robb, M. A. (1996). Potential energy surface crossings in organic photochemistry. *Chemical Society Reviews*, *25*, 321–328.
- Born, M., & Oppenheimer, R. (1927). Zur quantentheorie der molekeln. *Annals of Physics*, *389*, 457–484.
- Bornemann, F. A., & Schütte, C. (1998). A mathematical investigation of the Car-Parrinello method. *Numerische Mathematik*, *78*, 359–376.
- Brunk, E., & Rothlisberger, U. (2015). Mixed quantum mechanical/molecular mechanical molecular dynamics simulations of biological systems in ground and electronically excited states. *Chemical Reviews*, *115*, 6217.
- Car, R., & Parrinello, M. (1985). Unified approach for molecular dynamics and density-functional theory. *Physical Review Letters*, *55*, 2471.
- Castro, A., Appel, H., Oliveira, M., Rozzi, C. A., Andrade, X., Lorenzen, F., Marques, M. A., Gross, E. K. U., & Rubio, A. (2006). Octopus: A tool for the application of time-dependent density functional theory. *Physica Status Solidi B*, *243*, 2465–2488.
- Cottone, G., Lattanzi, G., Ciccotti, G., & Elber, R. (2012). Multiphoton absorption of myoglobin–nitric oxide complex: Relaxation by D-NEMD of a stationary state. *Journal of Physical Chemistry B*, *116*, 3397–3410.
- Deisseroth, K. (2011). Optogenetics. *Nature Methods*, *8*, 26–29.

- Desouter-Lecomte, M., & Lorquet, J.-C. (1979). Nonadiabatic interactions in unimolecular decay. IV. Transition probability as a function of the Massey parameter. *The Journal of Chemical Physics*, *71*, 4391–4403.
- Doltsinis, N. L., & Marx, D. (2002a). First principles molecular dynamics involving excited states and nonadiabatic transitions. *Journal of Theoretical and Computational Chemistry*, *1*, 319–349.
- Doltsinis, N. L., & Marx, D. (2002b). Nonadiabatic Car-Parrinello molecular dynamics. *Physical Review Letters*, *88*, 166402.
- Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. (1987). Hybrid Monte Carlo. *Physics Letters B*, *195*, 216–222.
- Ehrenfest, P. (1927). Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik. *Zeitschrift für Physik*, *45*, 455–457.
- Engel, G. S., Calhoun, T. R., Read, E. L., Ahn, T.-K., Mančal, T., Cheng, Y.-C., Blankenship, R. E., & Fleming, G. R. (2007). Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature*, *446*, 782–786.
- Ensing, B., De Vivo, M., Liu, Z., Moore, P., & Klein, M. L. (2006). Metadynamics as a tool for exploring free energy landscapes of chemical reactions. *Accounts of Chemical Research*, *39*, 73–81.
- Fernandez-Alberti, S., Roitberg, A. E., Nelson, T., & Tretiak, S. (2012). Identification of unavoided crossings in nonadiabatic photoexcited dynamics involving multiple electronic states in polyatomic conjugated molecules. *Journal of Chemical Physics*, *137*, 014512.
- Feynman, R. P. (1939). Forces in molecules. *Physics Review*, *56*, 340.
- Frenkel, D., & Smit, B. (2001). *Understanding molecular simulation: From algorithms to applications*. San Diego: Academic Press.
- Galli, G., & Pasquarello, A. (1993). First-principles molecular dynamics. In M. P. Allen & D. J. Tildesley (Eds.), *Computer simulation in chemical physics*. Dordrecht: Kluwer.
- Galli, G., Parrinello, M., Meyer, M., & Pontikis, V. (1991). Ab-initio molecular dynamics: Principles and practical implementation. In M. Meyer & V. Pontikis (Eds.), *Computer simulation in materials science. Proceedings of NATO ASI*, Kluwer.
- Gersonde, I. H., Hennig, S., & Gabriel, H. (1994). Nonadiabatic transitions in the photodissociation of HCl. *Journal of Chemical Physics*, *101*, 9558–9564.
- Gorshkov, V. N., Tretiak, S., & Mozysky, D. (2013). Semiclassical Monte-Carlo approach for modelling non-adiabatic dynamics in extended molecules. *Nature Communications*, *4*. doi:10.1038/ncomms3144
- Groenhof, G., Bouxin-Cademartory, M., Hess, B., De Visser, S. P., Berendsen, H. J., Olivucci, M., Mark, A. E., & Robb, M. A. (2004). Photoactivation of the photoactive yellow protein: Why photon absorption triggers a trans-to-cis isomerization of the chromophore in the protein. *Journal of the American Chemical Society*, *126*, 4228–4233.
- Grotendorst, A. M. J., Marx, D., & Muramatsu, A. (2002). *Quantum simulations of complex many-body systems: From theory to algorithms* (NIC series, Vol. 10). Jülich: Research Centre Jülich. www.fz-juelich.de/nic. Printer: Graphische Betriebe, Forschungszentrum Jülich.
- Hack, M. D., & Truhlar, D. G. (2000). Nonadiabatic trajectories at an exhibition. *The Journal of Physical Chemistry A*, *104*, 7917–7926.
- Hammes-Schiffer, S., & Tully, J. C. (1994). Proton transfer in solution: Molecular dynamics with quantum transitions. *Journal of Chemical Physics*, *101*, 4657–4667.
- Hayashi, S., Tajkhorshid, E., & Schulten, K. (2003). Molecular dynamics simulation of bacteriorhodopsin's photoisomerization using ab initio forces for the excited chromophore. *Biophysical Journal*, *85*, 1440–1449.
- Hazra, A., Soudackov, A. V., & Hammes-Schiffer, S. (2010). Role of solvent dynamics in ultrafast photoinduced proton-coupled electron transfer reactions in solution. *Journal of Physical Chemistry B*, *114*, 12319–12332.
- Hellmann, H. (1937). *Einführung in die quantenchemie*. Leipzig: Deuticke.
- CPMD. (n.d.). <http://www.cpmc.org/>
- Hudock, H. R., Levine, B. G., Thompson, A. L., Satzger, H., Townsend, D., Gador, N., Ullrich, S., Stolow, A., & Martinez, T. J. (2007). Ab initio molecular dynamics and time-resolved

- photoelectron spectroscopy of electronically excited uracil and thymine. *The Journal of Physical Chemistry A*, *111*, 8500–8508.
- Isborn, C. M., Gotz, A. W., Clark, M. A., Walker, R. C., & Martínez, T. J. (2012). Electronic absorption spectra from MM and ab initio QM/MM molecular dynamics: Environmental effects on the absorption spectrum of photoactive yellow protein. *Journal of Chemical Theory and Computation*, *8*, 5092–5106.
- Jasper, A. W., Stechmann, S. N., & Truhlar, D. G. (2002). Fewest-switches with time uncertainty: A modified trajectory surface-hopping algorithm with better accuracy for classically forbidden electronic transitions. *Journal of Chemical Physics*, *116*, 5424–5431.
- Jasper, A., Kendrick, B. K., Mead, C. A., & Truhlar, D. G. (2004a). Non-Born-Oppenheimer chemistry: potential surfaces, couplings, and dynamics. In X. Yang & K. Liu (Eds.), *Modern trends in chemical reaction dynamics part I*. Singapore: World Scientific. Chapter 8.
- Jasper, A. W., Zhu, C., Nangia, S., & Truhlar, D. G. (2004b). Introductory lecture: Nonadiabatic effects in chemical dynamics. *Faraday Discussions*, *127*, 1. doi:[10.1039/b405601a](https://doi.org/10.1039/b405601a).
- Jiang, R., & Sibert, E. L., III. (2012). Surface hopping simulation of vibrational predissociation of methanol dimer. *Journal of Chemical Physics*, *136*, 224104.
- Kang, H., Lee, K. T., Jung, B., Ko, Y. J., & Kim, S. K. (2002). Intrinsic lifetimes of the excited state of DNA and RNA bases. *Journal of the American Chemical Society*, *124*, 12958–12959.
- Karplus, M., & Petsko, G. A. (1990). Molecular dynamics simulations in biology. *Nature*, *347*, 631–639.
- Kato, H. E., Zhang, F., Yizhar, O., Ramakrishnan, C., Nishizawa, T., Hirata, K., Ito, J., Aita, Y., Tsukazaki, T., Hayashi, S., Hegemann, P., Maturana, A. D., Ishitani, R., Deisseroth, K., & Nureki, O. (2012). Crystal structure of the channelrhodopsin light-gated cation channel. *Nature*, *482*, 369–374. doi:[10.1038/nature10870](https://doi.org/10.1038/nature10870).
- Kubiak, K., & Nowak, W. (2008). Molecular dynamics simulations of the photoactive protein nitrile hydratase. *Biophysical Journal*, *94*, 3824–3838. doi:[10.1529/biophysj.107.116665](https://doi.org/10.1529/biophysj.107.116665).
- Kühne, T. D., Krack, M., Mohamed, F. R., & Parrinello, M. (2007). Efficient and accurate Car-Parrinello-like approach to Born-Oppenheimer molecular dynamics. *Physical Review Letters*, *98*, 066401.
- Laio, A., & Gervasio, F. L. (2008). Metadynamics: A method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports on Progress in Physics*, *71*, 126601.
- Laio, A., & Parrinello, M. (2002). Escaping free-energy minima. *Proceedings of the National Academy of Science*, *99*, 12562–12566.
- Laio, A., Bernard, S., Chiarotti, G. L., Scandolo, S., & Tosatti, E. (2000). Physics of iron at Earth's core conditions. *Science*, *287*, 1027–1030.
- Lambert, N., Chen, Y.-N., Cheng, Y.-C., Li, C.-M., Chen, G.-Y., & Nori, F. (2013). Quantum biology. *Nature Physics*, *9*, 10–18.
- Landau, L. D. (1932). Zur theorie der energieübertragung. II. *Physics of the Soviet Union*, *2*, 28.
- Lee, H., Cheng, Y.-C., & Fleming, G. R. (2007). Coherence dynamics in photosynthesis: Protein protection of excitonic coherence. *Science*, *316*, 1462–1465.
- Li, H., Elber, R., & Straub, J. E. (1993). Molecular dynamics simulation of NO recombination to myoglobin mutants. *Journal of Biological Chemistry*, *268*, 17908–17916.
- Li, Z., Kermodé, J. R., & De Vita, A. (2015). Molecular dynamics with on-the-fly machine learning of quantum-mechanical forces. *Physical Review Letters*, *114*, 096405.
- Liang, W., Isborn, C. M., Lindsay, A., Li, X., Smith, S. M., & Levis, R. J. (2010). Time-dependent density functional theory calculations of Ehrenfest dynamics of laser controlled dissociation of NO⁺: Pulse length and sequential multiple single-photon processes. *The Journal of Physical Chemistry A*, *114*, 6201–6206.
- Lin, J. Y. (2011). A user's guide to channelrhodopsin variants: Features, limitations and future developments. *Experimental Physiology*, *96*, 19–25.
- Lorenz, C., & Doltsinis, N. (2012) Handbook of Computational Chemistry. *Molecular dynamics simulation: from "ab initio" to "coarse grained"*.

- Majorana, E. (1932). Atomi orientati in campo magnetico variabile. *Il Nuovo Cimento* (1924–1942), 9, 43–50.
- Marques, M. A., Castro, A., Bertsch, G. F., & Rubio, A. (2003). Octopus: A first-principles tool for excited electron–ion dynamics. *Computer Physics Communications*, 151, 60–78.
- Marx, D., & Hutter, J. (2009). *Ab initio molecular dynamics: Basic theory and advanced methods*. Cambridge: Cambridge University Press.
- Michl, J. (1972a). Photochemical reactions of large molecules. II. Application of the model to organic photochemistry. *Molecular Photochemistry*, 4, 257–286.
- Michl, J. (1972b). Photochemical reactions of large molecules. I. A simple physical model of photochemical reactivity. *Molecular Photochemistry*, 4, 243–257.
- Nelson, T., Fernandez-Alberti, S., Roitberg, A. E., & Tretiak, S. (2014). Nonadiabatic excited-state molecular dynamics: Modeling photophysics in organic conjugated materials. *Accounts of Chemical Research*, 47, 1155–1164. doi:10.1021/ar400263p.
- Nikitin, E. E., & Züllicke, L. (1978). *Theory of chemical elementary processes*. Institute of Chemical Physics, Academy of Sciences of USSR. New York: Springer-Verlag.
- Nowak, W. (2012). Applications of computational methods to simulations of proteins dynamics. In *Handbook of computational chemistry* (pp. 1127–1153). Dordrecht: Springer.
- Orozco, M. (2014). A theoretical view of protein dynamics. *Chemical Society Reviews*, 43, 5051–5066.
- Pastore, G., Smargiassi, E., & Buda, F. (1991). Theory of ab initio molecular-dynamics calculations. *Physical Review A*, 44, 6334.
- Payne, M. C., Teter, M. P., Allan, D. C., Arias, T. A., & Joannopoulos, J. D. (1992). Iterative minimization techniques for ab initio total-energy calculations: Molecular dynamics and conjugate gradients. *Reviews of Modern Physics*, 64, 1045.
- Pérez, A., Luque, F. J., & Orozco, M. (2011). Frontiers in molecular dynamics simulations of DNA. *Accounts of Chemical Research*, 45, 196–205.
- Perilla, J. R., Goh, B. C., Cassidy, C. K., Liu, B., Bernardi, R. C., Rudack, T., Yu, H., Wu, Z., & Schulten, K. (2015). Molecular dynamics simulations of large macromolecular complexes. *Current Opinion in Structural Biology*, 31, 64–74.
- Pescitelli, G., Kato, H. E., Oishi, S., Ito, J., Maturana, A. D., Nureki, O., & Woody, R. W. (2014). Exciton circular dichroism in channelrhodopsin. *Journal of Physical Chemistry B*, 118, 11873–11885. doi:10.1021/jp505917p.
- Polli, D., Altoè, P., Weingart, O., Spillane, K. M., Manzoni, C., Brida, D., Tomasello, G., Orlandi, G., Kukura, P., & Mathies, R. A. (2010). Conical intersection dynamics of the primary photoisomerization event in vision. *Nature*, 467, 440–443.
- Remler, D. K., & Madden, P. A. (1990). Molecular dynamics without effective potentials via the Car-Parrinello approach. *Molecular Physics*, 70, 921–966. doi:10.1080/00268979000101451.
- Röhrig, U. F., Guidoni, L., Laio, A., Frank, I., & Rothlisberger, U. (2004). A molecular spring for vision. *Journal of the American Chemical Society*, 126, 15328–15329.
- Ryzewski, J., & Nowak, W. (2015). Memetic algorithms for ligand expulsion from protein cavities. *Journal of Chemical Physics*, 143, 124101.
- Ryzewski, J., Jakubowski, R., & Nowak, W. (2015). Communication: Entropic measure to prevent energy over-minimization in molecular dynamics simulations. *Journal of Chemical Physics*, 143, 171103.
- Sanbonmatsu, K. Y., & Tung, C.-S. (2007). High performance computing in biology: Multimillion atom simulations of nanoscale systems. *Journal of Structural Biology*, 157, 470–480.
- Shenvi, N., Subotnik, J. E., & Yang, W. (2011). Simultaneous-trajectory surface hopping: A parameter-free algorithm for implementing decoherence in nonadiabatic dynamics. *Journal of Chemical Physics*, 134, 144102.
- Shim, S., Rebentrost, P., Valleau, S., & Aspuru-Guzik, A. (2012). Atomistic study of the long-lived quantum coherences in the Fenna-Matthews-Olson complex. *Biophysical Journal*, 102, 649–660.
- Stueckelberg, E. K. G. (1933). *Theorie der unelastischen Stöße zwischen Atomen*. Basel: Birkhäuser.

- Teller, E. (1937). The crossing of potential surfaces. *Journal of Physical Chemistry*, *41*, 109–116.
- Tully, J. C. (1990). Molecular dynamics with electronic transitions. *Journal of Chemical Physics*, *93*, 1061–1071.
- Tully, J. (1998). Mixed quantum–classical dynamics. *Faraday Discussions*, *110*, 407–419.
- Tully, J. C. (2012). Perspective: Nonadiabatic dynamics theory. *Journal of Chemical Physics*, *137*, 22A301. doi:[10.1063/1.4757762](https://doi.org/10.1063/1.4757762).
- Tully, J. C. (1976). Nonadiabatic processes in molecular collisions. In W. H. Miller (Ed.), *Dynamics of molecular collisions. Vol. 2: Modern Theoretical Chemistry*.
- Tully, J. C., & Preston, R. K. (1971). Trajectory surface hopping approach to nonadiabatic molecular collisions: The reaction of H⁺ with D₂. *Journal of Chemical Physics*, *55*, 562–572.
- Virshup, A. M., Punwong, C., Pogorelov, T. V., Lindquist, B. A., Ko, C., & Martinez, T. J. (2008). Photodynamics in complex environments: Ab initio multiple spawning quantum mechanical/molecular mechanical dynamics†. *Journal of Physical Chemistry B*, *113*, 3280–3291.
- Warshel, A., & Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology*, *103*, 227–249.
- Welke, K., Frähmcke, J. S., Watanabe, H. C., Hegemann, P., & Elstner, M. (2011). Color tuning in binding pocket models of the chlamydomonas-type channelrhodopsins. *Journal of Physical Chemistry B*, *115*, 15119–15128.
- White, A. J., Gorshkov, V. N., Wang, R., Tretiak, S., & Mozyrsky, D. (2014). Semiclassical Monte Carlo: A first principles approach to non-adiabatic molecular dynamics. *Journal of Chemical Physics*, *141*, 184101. doi:[10.1063/1.4900988](https://doi.org/10.1063/1.4900988).
- Yarkony, D. R. (1996). Diaboliocal conical intersections. *Reviews of Modern Physics*, *68*, 985.
- Zener, C. (1932). Non-adiabatic crossing of energy levels. In *Proceedings of the Royal Society of London A: Mathematical, physical and engineering sciences* (pp. 696–702). London: The Royal Society.
- Zimmerman, H. E. (1966). On molecular orbital correlation diagrams, the occurrence of Möbius systems in cyclization reactions, and factors controlling ground-and excited-state reactions. I. *Journal of the American Chemical Society*, *88*, 1564–1565.

Ab Initio Investigation of Photochemical Reaction Mechanisms: From Isolated Molecules to Complex Environments

46

Igor Schapiro, Patrick Zakhia El-Khoury, and Massimo Olivucci

Contents

Introduction	1944
Methods in Computational Photochemistry	1948
Multiconfigurational Quantum Chemistry	1948
QM/MM Methodology	1954
Case Studies of Isolated Chromophores	1956
Photoisomerization in a Rhodopsin Chromophore Model	1956
Photofragmentation Through a Conical Intersection: The Photodenitrogenation of a Bicyclic Azoalkane	1962
Charge Transfer and Intermolecular Hydrogen Transfer Mediated by a Conical Intersection: Quenching the Fluorescence of Bicyclic Azoalkanes	1965
Case Studies of Chromophores Embedded in a Molecular Environment	1971
Green Fluorescent Proteins (GFP)	1971
Understanding the Spectral Tuning in Retinal Proteins	1974
Deactivation Mechanism in Cytosine-Guanine DNA Base Pair	1980
Absorption Spectra of a Coumarin in Solution	1981
Excited State Molecular Dynamics	1983
Tracking the Photoisomerization of Retinal in Different Environments	1983
Conclusion	1986
Bibliography	1987

I. Schapiro (✉) • P.Z. El-Khoury
Chemistry Department, Bowling Green State University, Bowling Green, OH, USA

M. Olivucci
Dipartimento di Chimica, Università di Siena, Siena, Italy
e-mail: molivuc@bgsu.edu

Abstract

This chapter focuses on the computational investigations of light-induced chemical reactions in different systems ranging from organic molecules in vacuo to chromophores in complex protein environments. The aim is to show how the methods of computational photochemistry can be used to attain a molecular-level understanding of the mechanisms of photochemical and photophysical transformations. Following a brief introduction to the field, the most frequently used quantum chemical methods for mapping excited state potential energy surfaces and for studying the mechanism of photochemical reactions in isolated molecules are outlined. In the following sections, such methods and concepts are further developed to allow the investigation of photo-induced reactions in solution and in the protein environment.

Introduction

Since the early 1990s, quantum chemical methods based on *ab initio* multi-configurational wavefunctions have been used to investigate the mechanisms of photochemical organic reactions and have led to the development of the field of computational photochemistry (Bernardi et al. 1996b). As a consequence it is now possible, using available program packages (examples of which include Molcas (Andersson et al. 2009; Aquilante et al. 2009), Gaussian (Frisch et al. 2009), Gamess-US (Gordon et al. 2005; Schmidt et al. 1993), Columbus (Lischka et al. 2006), Molpro (Werner et al. 2008), and Q-Chem (Adams et al. 2009; Shao et al. 2006)), to locate excited state equilibrium structures, conical intersections, or singlet/triplet crossings, and map ground or excited state minimum energy paths at a level of accuracy that often allows for comparison with spectral data (Kutateladze 2005; Olivucci and Sinicropi 2005). This offers chemists a powerful tool to rationalize and predict such complex processes as the light-induced response of a molecule.

There are two events that may happen when light energy is absorbed by a molecule: wastage or exploitation. Some technological application may require molecules whose structures are unaffected by light absorption, those that efficiently dispose of the excess electronic energy (e.g., through light emission and internal conversion). Other applications may require molecules that exploit light to drive specific chemical, conformational and electronic changes (e.g., through photochemical reactions or light-induced energy or electron transfer). It can be easily recognized that nature provides many examples of effective light energy wastage and exploitation. Therefore, the understanding of these events can be considered a basic requirement for the rational design of efficient photochemical reactions, artificial photosynthetic systems, and molecular-level devices and machines. Below, we will present case studies from small molecules in the gas phase to molecules in complex environments such as biological photoreceptors.

The computational investigation of a light-induced reactive process requires the construction of the associated photochemical reaction path (Olivucci and Sinicropi 2005). In its simplest form, this is defined by a branched minimum energy path that connects the Franck–Condon (FC) point A^* on the potential energy surface of the light-absorbing state with one (or more) primary photoproduct (B) or original reactant (A). As illustrated in Fig. 1 left, if the relaxation of A^* leads to a stable excited state equilibrium species (structure B^*) this may form a luminescent state, for instance a fluorescent state (FS), thus leading to radiative deactivation. High fluorescence quantum yields require long excited state lifetimes that, among other factors, could be imposed by high energy barriers restraining further evolution of A^* . If this is not the case, B^* can evolve beyond a transition state (TS), to produce a different excited state species (C^*) that may also decay radiationally. If this is the case, one has an excited state conformational or chemical change that might lead to a different fluorescence emission. One of the most prominent examples of this so-called dual fluorescence is that produced by 4-(N, N-dimethylamino)benzonitrile in polar solvents (Grabowski et al. 2003). Alternatively (see Fig. 1 right), the excited state transition structure could connect B^* to a point where efficient radiationless decay (internal conversion) takes place. In the limiting case of a negligible or non-existent barrier, A^* may promptly reach such point on a picoseconds time-scale. For internal conversion, this process is usually associated with regions where there is a limited gap between ground and excited states, a large overlap between a low lying vibrational state of the photoexcited molecule and a suitable ground state vibrational mode (Simons 1983). These conditions are satisfied in the regions surrounding a

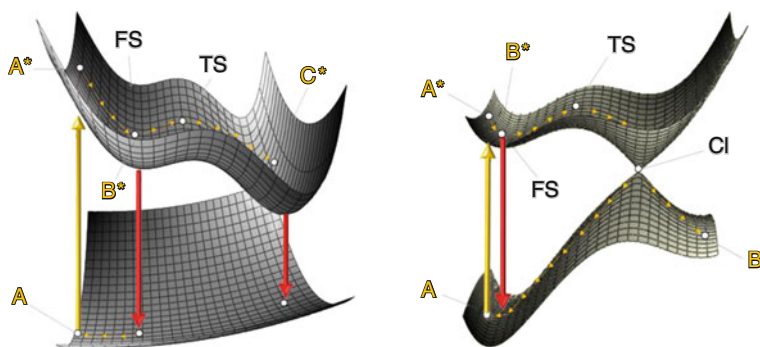


Fig. 1 Schematic structure of the photochemical reaction path. Left: general structure of the excited state energy surface in the vicinity of the Franck–Condon point. A^* is geometrically identical to the minimum on the ground state, but has the electronic wave function of the excited state. The species B^* is a relaxed form of A^* that may emit fluorescence and/or evolve towards a different region of the excited state potential energy surface through the transition state (TS), producing a different emitting or reactive species C^* . Right: intersecting ground and excited state potential energy surfaces. The arrows indicate the direction of the minimum energy path connecting A^* to the conical intersection (CI) and then to the photoproduct B or original reactant A

conical intersection (CI). Therefore, evolution towards a CI leads to an efficient decay to the ground state (decay is common in the whole region surrounding the CI). When the relaxation following decay results exclusively in reactant back-formation (CI \rightarrow A stream of arrows) one has internal conversion. In contrast, when the ground state relaxation produces a novel species B (a primary photoproduct) by following a different relaxation path, one has a photochemical event (left to right stream of arrows).

CIs provide a common key mechanistic entity for the description of a photochemical reaction and ultrafast internal conversion (Turro et al. 2009). The potential energy surface region surrounding each CI point can be seen as an efficient radiationless deactivation channel (Bernardi et al. 1996a; Migani and Olivucci 2004; Robb et al. 1995) that, in photochemistry, is called a funnel. A photochemical funnel (Braslavsky 2007) corresponds to a molecular structure in the vicinity of a CI that would decay within one half of a vibrational period. For this reason computer simulations appear to be the only “direct” source of information about the structure of a molecule in a funnel. Notice that one has a CI only when the ground and the excited state have the same spin. If this is not the case and the electronic states have different spins the process is called intersystem crossing (Simons 1983; Turro et al. 2009).

In Fig. 1 we show that the entire “photochemical reaction path” is defined and computed in terms of a set of connected minimum energy paths (MEP) corresponding to steepest descent lines. In particular, the path starting at A* on the spectroscopic state and ending at the photoproduct energy minimum B on the ground state energy surface is constructed by joining two MEPs. A first MEP connects FC point to the CI. A second MEP connects the CI to the photoproduct B. A third MEP can be computed that starts at CI and describes the reactant reconstitution process responsible for partial return of the photoexcited species to the original ground state minimum (i.e., internal conversion). The mapping of the excited state potential energy surface of chemically different organic chromophores, strongly supports the idea that the mechanistic scenario described above (i.e., involving decay at a CI), is common. This was first realized after the development of CASSCF gradients (Schlegel and Robb 1982) in the early 1980s, and through a systematic computational research effort carried out in the 1990s (Bernardi et al. 1990; 2006). Also a number of algorithms were developed in order to locate the conical intersections and characterize the associated geometries (Anglada and Bofill 1997; Bearpark et al. 1994; De Vico et al. 2005; Farazdel and Dupuis 1991; Koga and Morokuma 1985; Levine et al. 2006; Ragazos et al. 1992; Yarkony 1993). Since the conical intersections are high-dimensional entities and not single points, the algorithms minimize the energy along the intersection seam. The points with the lowest energy of the seam are referred to as minimum energy conical intersections (MECI). The decay processes that are mediated by conical intersection are ultrafast and therefore too fast to allow the molecule to travel along the seam to reach the MECI. Thus transitions can occur at intersections which are higher in energy than MECI. This idea prompted the development of advanced methods to optimize stationary points along the multidimensional seam of intersection and determine the

structure of more extended segments of the crossing (Laino and Passerone 2004; Sicilia et al. 2007, 2008). A better way of defining the conical intersection point during the decay is to perform molecular dynamics simulation, which considers the kinetic energy of the molecule and can explore different regions of the excited state surface provided a number of different initial conditions. The methods solve the classical equations of motions and use forces from electronic structure methods to move on a single potential energy surface. In order to account for non-adiabatic transitions, a variety of approaches have been produced. Trajectory surface hopping algorithms are widely used to accomplish this task. They have in common that a transition is induced, i.e., a hop between surfaces, when the non-adiabatic coupling becomes strong. There are several examples where this surface hopping algorithm has been successfully applied to study photochemical events (Barbatti and Lischka 2008; Barbatti et al. 2008; Groenhof et al. 2004, 2008; Schaefer et al. 2007; Schapiro et al. 2009; Virshup et al. 2009; Weingart 2007; Weingart et al. 2004). A more sophisticated approach that treats both electrons and nuclei quantum mechanically was implemented by Martinez and co-workers (Ben-Nun and Martinez 1998; Levine et al. 2006). Herein an additional basis function describing the nuclei is created or spawned (therefore the name “multiple-spawning”) when strong non-adiabatic coupling between different states arises.

Below we discuss selected case studies of a series of processes investigated by means of computational photochemistry. We begin with isolated molecules in vacuum and conclude with chromophores in complex environments. The first case study focuses on reduced models of the retinal chromophore of visual photoreceptors. Subsequently the photodenitrogenation, intermolecular charge transfer and hydrogen transfer in azoalkanes are described. The following section deals with chromophores in complex environments. An overview will be given starting with GFP and asFP595 which is a GFP-like protein (Fig. 2) whose fluorescence can be switched on by Z/E photoisomerization. This is followed by an extensive discussion of a simulation of the photoisomerization of rhodopsin, a membrane protein characterized by seven α -helices (Fig. 2). The cationic Rh chromophore is the 11- cis stereoisomer of the protonated Schiff base of retinal (PSB11), and is connected through one covalent bond to the Lys296 residue. Upon absorption of visible light, PSB11 undergoes a Z/E photoisomerization and produces (stereoselectively) the corresponding all- trans chromophore (PSBAT), which ultimately triggers the Rh activity (Kandori et al. 2001; Mathies and Lugtenburg 2000; Palczewski et al. 2000; Teller et al. 2001). Computational studies of bR, an archaeal rhodopsin with a red-shifted absorption maximum (λ_{\max}^a) relative to Rh, and with a PSBAT chromophore, rather than PSB11, are described. Another class of molecules that have been excessively studied by tools of computational photochemistry are the deoxyribonucleic acids (DNA). We will illustrate how quantum chemistry is applied to unveil a mechanism of photodeactivation in a cytosine-guanine base pair. In the next case study the calculation of absorption spectra, including solvent effects and a novel method for taking vibrational contributions into account, is applied to coumarin C153. Finally, we highlight molecular dynamics simulations used to understand the isomerization mechanism in retinal containing proteins (Rh and bR).

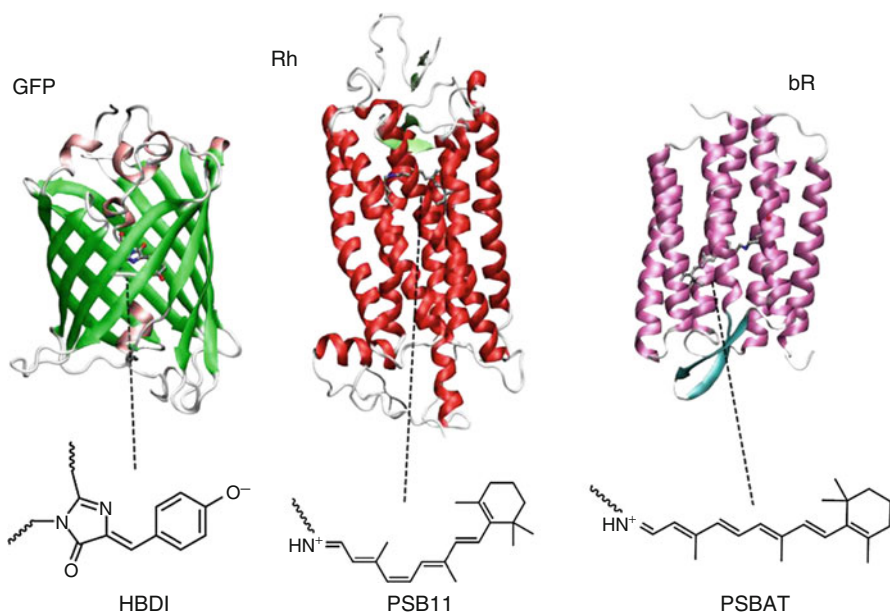


Fig. 2 Cartoon representation of the three-dimensional structures of three different photoresponsive proteins. Left: GFP with its HBDI anionic chromophore. Center the bovine visual pigment rhodopsin together with the cationic PSB11 chromophore. Right: bacteriorhodopsin and its PSBAT chromophore

Methods in Computational Photochemistry

Multiconfigurational Quantum Chemistry

Among the possible quantum chemical methodologies (e.g., semi-empirical or *ab initio*) used to model excited state species, below we mainly deal with *ab initio* (i.e., from first principles of quantum mechanics) approaches. In particular, to describe the evolution of intermediates and species whose geometries are far from equilibrium, one employs *ab initio* multiconfigurational methods. These methods are based on the multideterminant configuration interaction ansatz (i.e., where the electronic wavefunction is represented by a linear combination of determinants rather than by a single determinant) being the methods of choice when, in a photochemical process, a molecule rapidly changes its electronic structure and enters regions of crossing between different potential energy surfaces. The favorable properties of these methods are: (1) the accuracy can be systematically improved until they provide the exact solution of the non-relativistic Schrödinger equation (when a full-configuration interaction wavefunction and basis set close to the Hartree limit are used), (2) they are unbiased in the sense that they do not contain empirical parameters, and (3) they are suitable for dealing with potential energy

surface crossings between electronic states of the same or different spin-multiplicity (e.g., when the wavefunction is written in terms of Slater determinants) and there is a rapid change of electronic structure. The drawback is that these types of calculations are computationally expensive even though, nowadays, progress is being made to extend their applicability (Aquilante et al. 2009).

The ab initio complete-active-space self-consistent-field (CASSCF) method (Roos 1987) is a multiconfigurational method that offers maximum flexibility for an unbiased description of the electronic and geometrical features of molecules in their ground and excited electronic states. This is a particularly powerful method because the orbital coefficients are optimized together with the weights of a linear combination of all possible configurations constructed from a selected set of electrons and orbitals (Roos and Taylor 1980; Schmidt and Gordon 1998; Siegbahn et al. 1981). Furthermore, the CASSCF wavefunction can be used for subsequent second-order perturbation theory computations (CASPT2) of the dynamic correlation energy, (Andersson et al. 1990) ultimately allowing for a quantitative simulation and prediction of excitation energies.

In order to illustrate the accuracy that one can reach by a CASPT2//CASSCF protocol (the ‘//’ sign indicates that the molecular structure is determined (optimized) at the CASSCF level but that the energy is computed at the CASPT2 level), we report the results of the simulation of the observed gas-phase spectra of PSBAT in Fig. 3 (Andruniw and Olivucci). The spectrum was simulated according to the procedure described in ref. Andruniw and Pawlikowski (1998). The $S_0 \rightarrow S_1$ electronic excitation is found to be 524 nm at the CASPT2//CASSCF/6-31G* level of theory (the 6-31G*) Pople-type acronym refers to the specific atomic basis set used in the calculation). In 2006 the experimental spectrum of PSBAT was measured by Andersen and co-workers showing several overlapping bands that were initially attributed to the vibrational excitations in S_1 (Nielsen et al. 2006). The most intense peak was at 590–600 nm meaning the error of the CASPT2//CASSCF protocol is nearly 6 kcal · mol⁻¹. However, very recently the same group has repeated the gas phase measurement finding a remarkably broad absorption with a flat top ranging from 530 to 610 nm (Rajput et al. 2010). By probing different retinal PSB analogues it was shown that the broadness of the band is due to a highly flexible C5–C6–C7–C8 dihedral angle. The C6–C7 bond connects the polyene chain with the β -ionone ring and has either a 6s- cis or 6s- trans conformation. The shorter wavelength of 530 nm was assigned to the 6s- cis-conformer and the longer wavelength of 610 nm to the 6s- trans-conformer. Compared to the new absorption maximum the CASPT2//CASSCF value is in very good agreement (0.6 kcal · mol⁻¹).

The general assumption underlying the validity of the CASPT2//CASSCF protocol is that the CASSCF level of theory is accurate enough to provide qualitatively correct ground and excited state equilibrium structures, and in particular, MEPs. Currently this is of practical importance since only the CASSCF energy gradient can be evaluated analytically while the analytical CASPT2 gradients are not available for molecules of the size of interest in photobiology. This approximation is still awaiting a thorough evaluation. However, in certain cases, the equilibrium structures

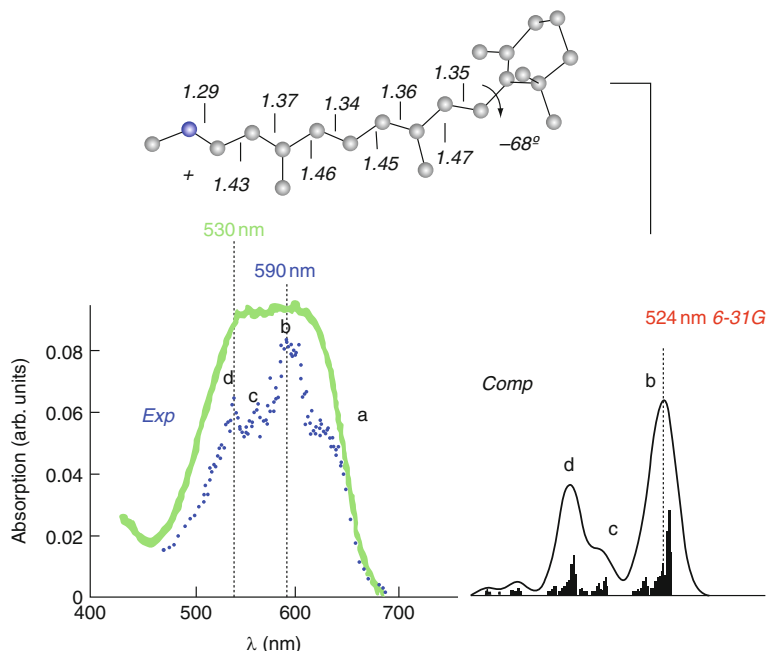


Fig. 3 The $S_0 \rightarrow S_1$ absorption band of gas-phase PSBAT computed on the basis of excited state CASSCF frequency calculations (the frequencies have been scaled using a 10% factor). Top: the equilibrium geometry of the PSBAT is computed at the CASSCF/6-31G* level of theory with a 12 electrons in 12 orbitals active space. Left: experimental absorption spectra recorded by Andersen and co-workers (Nielsen et al. 2006; Rajput et al. 2010). Right: simulation of the first order line spectra and of the spectral band by Lorentzian fitting. The maximum “b” is assigned to the vertical transition and it is < 600 nm (~ 590 nm). The vertical excitation energy resulting in the predicted λ_{\max}^a value is computed by single-point CASPT2/6-31G* calculation on the CASSCF equilibrium geometry. Normal modes involving mixed β -ionone ring and chain torsion and bending, as well as C-CH₃ motion in simulated spectrum are difficult to assign because of overlapping bands

on a CASPT2 potential energy surface may differ significantly from the same structure calculated at the CASSCF level. In order to illustrate this point, the energy profile along a CASPT2//CASSCF minimum energy path computed for a simplified HBDI chromophore in the excited state is compared to the CASSCF energy profile in Fig. 4 (Martin et al. 2004).

Recently (Valsson and Filippi2010) reported the photoisomerization of retinal models using high-level quantum chemical methods, among others CASPT2, which include dynamic electron correlation. The two and three double bond truncated retinal models optimized at the CASPT2 level and numerical gradients. It was found that all bonds were equally lengthened and therefore the rotation around several bonds became possible. However, not all of the rotations lead to radiationless decay to the ground state. Some of the pathways on the excited state involved single bond rotation up to 45° and ended up in energy wells on the excited state (see Fig. 1 left). This finding is in contrast with the widely accepted reaction

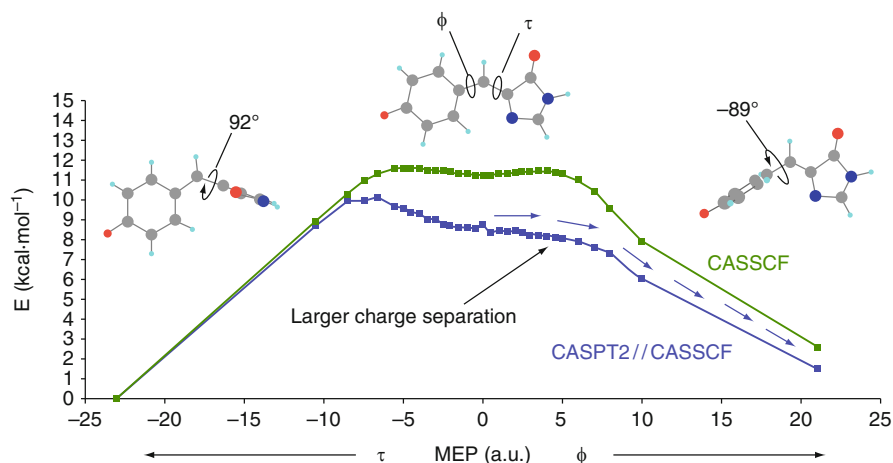


Fig. 4 CASSCF (green squares) and CASPT2 (blue squares) energy profiles for the electronic S_1 state for the torsional relaxation. The stream of arrows indicates the energetically favored relaxation path. A change in shape of the reaction energy profiles after dynamic electron correlation correction (i.e., passing from CASSCF to CASPT2) has also been documented in other systems. See a discussion in ref. Sinicropi et al. (2003) (Redrawn with permission from ref. Martin et al. 2004)

coordinate established at the CASSCF level which is driven subsequently by two uncoupled modes. The first mode is responsible for the inversion of the bond length alternation in the excited state whereas the second mode is mainly the rotation of the central double bond. It is shown that this path also exists in the CASPT2 potential energy surface but it is reached rather from relaxing directly out of the Franck–Condon point. These controversial results on retinal chromophore in vacuo outline the importance of the dynamic electron correlation and the choice of the QM method that need to be verified for more realistic models. The size of the chromophore models should be extended towards the full retinal. However, the lack of the analytical gradients for CASPT2 makes these calculations extremely computationally demanding.

The difference between the CASPT2 and CASSCF geometries has also been investigated for both ground and excited state models of the PSB11 chromophore. The main difference in the results between these treatments is found in the ground state equilibrium geometry, where one finds a bond length alternation (BLA) pattern along the polyene chain which is less pronounced at the CASPT2 level (indicating a more extended delocalization) than at the CASSCF level (Page and Olivucci 2003). However, these studies were carried out with limited basis sets. Recently, different quantum chemical studies of the λ_{\max}^a of biological chromophores have employed density functional theory (DFT) rather than CASSCF ground state equilibrium geometries. For PSB11 and PSBAT these levels yield significantly different BLA values. Blomgren et al. (Blomgren and Larsson 2005) have proposed that this difference originates from the dynamic electron correlation which is (partially)

included in DFT functionals. The changes in the BLA computed for PSB11 are also reflected in the magnitude of the twisting of the C6–C7 single bond with respect to the backbone (see Fig. 2) and consequently in the conformation of the β -ionone ring. The ring is nearly planar in the DFT geometry but highly twisted in the CASSCF geometry as reported by Bravaya et al. (2007). Using the PBE0/cc-pVDZ calculations these authors find that PSB11 has two conformers featuring a -39° and $+170^\circ$ twisting of the C5–C6–C7–C8 dihedral angle corresponding to *s-cis* and *s-trans* isomers of the β -ionone ring, respectively. These results can be compared with the CASSCF/6-31G* data presented by Cembran et al. (2005) for the same chromophore. The CASSCF/6-31G* level of theory yields a single highly twisted PSBAT conformer with a -69° C5–C6–C7–C8 dihedral angle.

The gas-phase λ_{\max}^a (i.e., the corresponding vertical excitation energy) of PSB11 has been computed with different protocols that are also a matter of debate. Using a modified multiconfiguration quasi-degenerate perturbation theory aug-MCQDPT2 for excitation energies and DFT optimized geometry at PBE0/cc-pVDZ level for PSB11 (Bravaya et al. 2007) yields a 599 nm λ_{\max}^a , which appeared to be a very good estimate of the experimentally observed value for the gas-phase PSB11 (~ 610 nm) (Nielsen et al. 2006). In the light of the refined experimental absorption maximum at 530 nm (Rajput et al. 2010) the deviation is $7 \text{ kcal} \cdot \text{mol}^{-1}$. The use of a computationally consistent (i.e., with geometry and energy computed using the same method) TDDFT//DFT protocol (TD-B3LYP/6-31 + G*//B3LYP/6-31G) yields a 537 nm λ_{\max}^a (Wanko et al. 2005) and therefore performs better than mixed wavefunction-based/density-functional approaches. Also the CASPT2/ANO//MP2/6-31G** calculated result is far off (606 nm) which is based on a protonated PSB11 model rather than an alkylated (N-dimethyl) like in the aug-MCQDPT2//PBE0 study, which is more consistent with the measured chromophore. In contrast the CASPT2//CASSCF/6-31G* method yields 524 nm (Andruniw and Olivucci) which appears to be close to the recent experimental excitation energy of 530 nm (Rajput et al. 2010). TDDFT is a method that has much promise in studying excited states, but currently has limitations. It has been shown to be effective when used to compute valence-excited states (Dreuw and Head-Gordon 2005). However, when the excited state being studied has charge transfer (CT) character (either intermolecular or intramolecular) or if the molecule has an extended π system, TDDFT cannot reliably predict excitation energies. It has failed to accurately describe the CT states of several biologically relevant molecules, some of which include porphyrin complexes (Dreuw and Head-Gordon 2005). The failure is extensive. Not only is the energy gap inaccurate but also the topography of the excited state potential energy surface. This can be seen from the results of a study conducted in our laboratory (Fantacci et al. 2004), the results of which are displayed in Fig. 5. A MEP was calculated for a truncated model of the retinal chromophore (only three conjugated double bonds were included) with CASSCF. Energies were then recalculated at geometries along the path using CASPT2//CASSCF and various density functionals using TDDFT//CASSCF. Near the FC point, where the wavefunction changes rapidly, a potential energy well and maximum were found using TDDFT. These are features of an excited state energy surface that would not

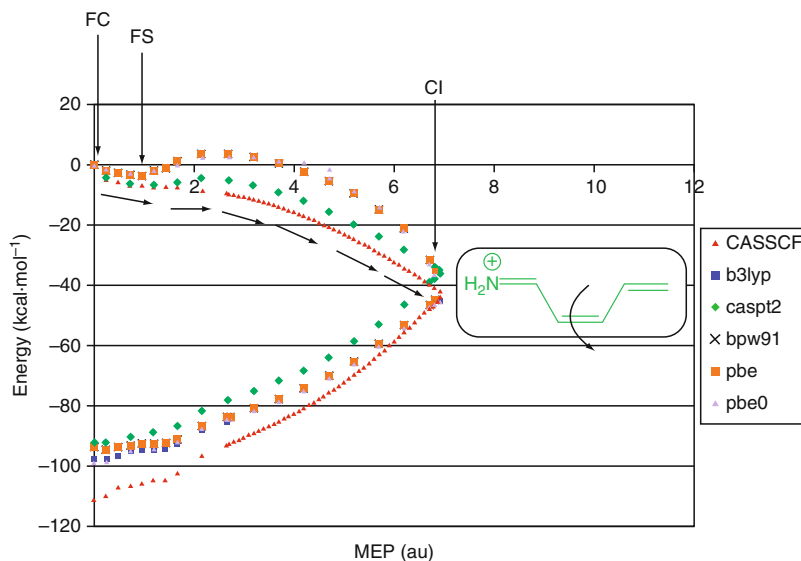


Fig. 5 Energy profiles along the S_1 photoisomerization coordinate of a simplified PSB11 model computed at the CASPT2//CASSCF/6-31 G* level. The energy at the Franck–Condon point is taken as the reference for the comparison. The fluorescent state (FS) corresponds to the local minimum along the TDDFT curve. The local maximum of the TDDFT curve corresponds to a model structure whose dihedral angle around the central isomerizing double bond measures 15° . There is no maximum or energy well along the CASSCF energy profile and these features are much less distinct along the CASPT2//CASSCF energy profile. In the PSB11 chromophore the value of the C5–C6–C7–C8 dihedral angle strongly affects the vertical excitation energy of the gas-phase retinal chromophore. By changing it one can change the excitation energy by up to 0.7 eV. The PBE0/cc-pVDZ bond alternation value is close to the value found for N-methyl-11-cis PSBR obtained by Sekharan et al. (2006) with the MP2 method. As suggested by Wanko et al. (2005) the hybrid exchange-correlation functionals like the PBE0 used in the aforementioned work, should provide BLA values very close to those computed at the CASPT2 level of theory. Also the similarity of MP2/6-31G* and CASPT2/6-31G* equilibrium geometry was directly demonstrated for short retinal chromophore models (i.e., the pentadieniminium cation) (Page and Olivucci 2003). Of course, no CASPT2 optimization of the full retinal chromophore with adequate (e.g., large ANO) basis sets has been reported up to now and in our view the issue regarding the correct ground state geometry of the retinal chromophore has still to be settled (Redrawn with permission from ref. Fantacci et al. 2004)

be expected to be exhibited by a molecule that undergoes an ultrafast femtosecond-scale photochemical reaction, so it is presumed that they are artifacts resulting from the use of TDDFT. The CASPT2//CASSCF surface is more accurate, exhibiting a plateau, with only a minuscule potential energy well. Another benchmark study came to similar conclusions, finding that the use of TDDFT led to a qualitatively incorrect description of the excited state potential at geometries distant from the FC point when the method was used on model retinal chromophores (Wanko et al. 2004). The source of the problem can be attributed to the CT nature of retinal's first excited state.

Problems with TDDFT can be found at all points of a photochemical reaction path, from the FC point to the CI. At the CI the quality of the immediately surrounding ground and excited state potential energy surface (PES) is inaccurate (Levine et al. 2006). A CI must have two branching directions (i.e., there must be two different motions that the molecule can undergo that will cause the energies of the states to separate). TDDFT, in contrast, describes a CI where only one branching direction lifts the degeneracy of states. An additional problem is that near the CI the slope of the PES along the single degeneracy-lifting coordinate is much steeper than that obtained with more reliable methods, as shown in the case of butadiene (Levine et al. 2006).

There are many reasons why TDDFT is not appropriate to study excited states, especially when double excitation or CT states are involved. For further reading we recommend the review by Dreuw et al. (Dreuw and Head-Gordon 2005). CASPT2 is much more reliable, in terms of accuracy, but cannot be used for optimizations, unless the molecule being studied is very small, in which case numerical gradients can be calculated with CASPT2. In the majority of cases, though, the best compromise will be CASPT2//CASSCF.

QM/MM Methodology

The quantum mechanical methods described above are limited in the application by the size of the molecule. In 1976 Warshel (Warshel and Levitt 1976) proposed a hybrid treatment, which allowed to treat biological macromolecules such as proteins that are too large for quantum chemical calculations. The idea is to divide the macromolecule into two or more subsystems which have different sizes and functions. The smaller subsystem, where a change of electronic structure or a chemical reaction occurs, is described by a quantum mechanical (QM) approach (i.e., by a suitable quantum chemical method), whereas the larger residual subsystem is treated using a molecular mechanics (MM) force field. Fortunately, because photochemical reactions are localized only to a specific area (i.e., the chromophore) electronic structure methods can be applied to this significantly smaller region, and the remaining part can be treated on a lower level such as a classical force field. This approach allows the study of large systems but also introduces new problems, the most important of which regards the treatment of the frontier between the two subsystems. In the past three decades much work has been done to solve this problem and for a comprehensive outline the reader is referred to the review by Senn and Thiel (2009).

Using the multiconfigurational approach combined with classical or molecular mechanics treatment of the environment a CASPT2//CASSCF/MM protocol was established (the//separator has the same meaning as above). Thus geometry optimization and molecular dynamics simulations are carried out at the CASSCF/MM level (this notation indicates that the QM part of the molecule is described at the CASSCF level while the rest is described with an MM force field such as AMBER (Case et al. 2002; Ponder and Case 2001), CHARMM, (Brooks et al. 1983)

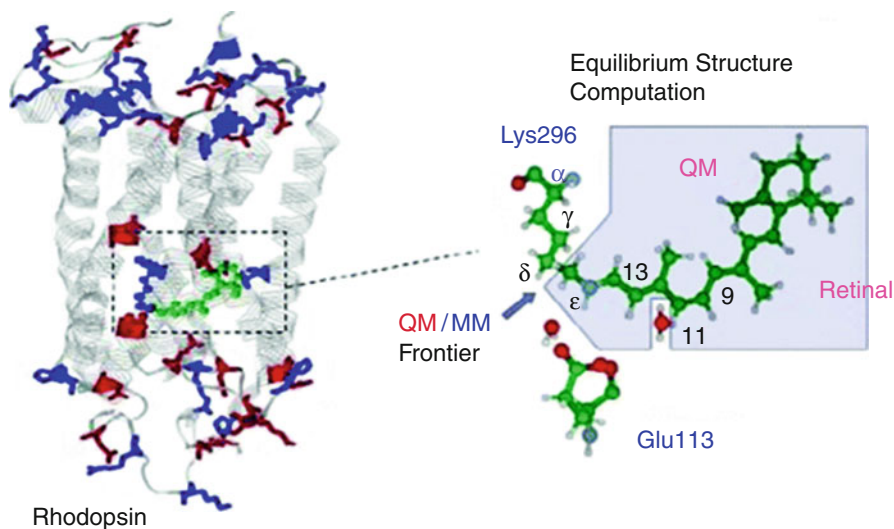


Fig. 6 Main parts of a QM/MM model for Rh. The protein structure is usually obtained from the Protein Data Bank (Berman et al. 2000). The QM part corresponds to the chromophore plus a part of the side-chain of the amino acid that is linked to it. The QM/MM frontier region usually corresponds to a weakly polarized C–C bond that, for Rh, is chosen to correspond to the $C_\delta - C_\epsilon$ bond of the Lys296 side chain in Rh. In this bond the QM atom (C_ϵ) is capped with an H link-atom and the MM atom (C_δ) is held in place by MM potentials that are specifically parameterized for this region. This MM parameterization of the QM/MM frontier region is specific to the chromophore and the side chain

or OPLS (Jorgensen and Tirado-Rives 1988)). Of course, the application of this methodology requires the construction of a suitable protein model. In order to study the Rh pigment, a CASPT2//CASSCF/AMBER protocol based on the QM/MM link-atom electrostatic-embedding scheme (Senn and Thiel 2009) (see Fig. 6) has been implemented and successfully applied for the first time in our laboratory (Ferré and Olivucci 2003a, b; Ferré et al. 2004).

A meaningful calculation of the spectroscopic parameters can be achieved by running a long time ground state QM/MM trajectory simulation and computing CASPT2 vertical excitation energies at regular time intervals. In this way the λ_{\max}^a , fluorescence maximum (λ_{\max}^f) or phosphorescence maximum could be obtained by averaging these results (for an application of this concept using other quantum chemical methods see refs Hoffmann et al. (2006) and Pistolesi et al. (2009)). However, because of the cost of CASPT2//CASSCF calculations this strategy appears to be unpractical. The calculation of quantities such as the λ_{\max}^a and λ_{\max}^f cannot be systematically carried out using this approach. Yet, as explicitly mentioned above, the CASPT2//CASSCF level of QM treatment is required if one wants an unbiased and rigorous quantitative description of the electronic excited state of an organic chromophore/fluorophore. One possible solution to this problem is to prepare a model of the system that provides a sufficiently accurate

representation of the time-averaged configuration of the chromophore/fluorophore and its environment. Using this model the chromophore would effectively feel the average steric and electrostatic field imposed by the protein.

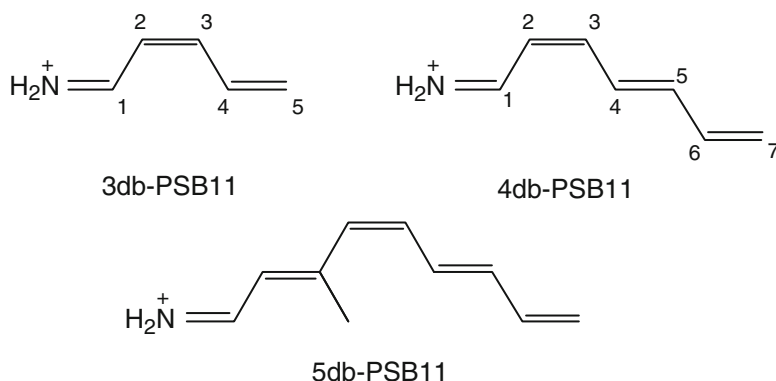
The case studies given in section “[Case Studies of Chromophores Embedded in a Molecular Environment](#)” show that within certain accuracy QM/MM models are viable for different proteins. Indeed, we have shown that models based on the experimental crystallographic structure (i.e., where the protein is not fully relaxed or equilibrated) allow the prediction of the spectroscopic parameters with an accuracy of a few kcal · mol⁻¹ for systems that differ in protein structure and in the chemical nature of the chromophore/fluorophore.

Case Studies of Isolated Chromophores

Photoisomerization in a Rhodopsin Chromophore Model

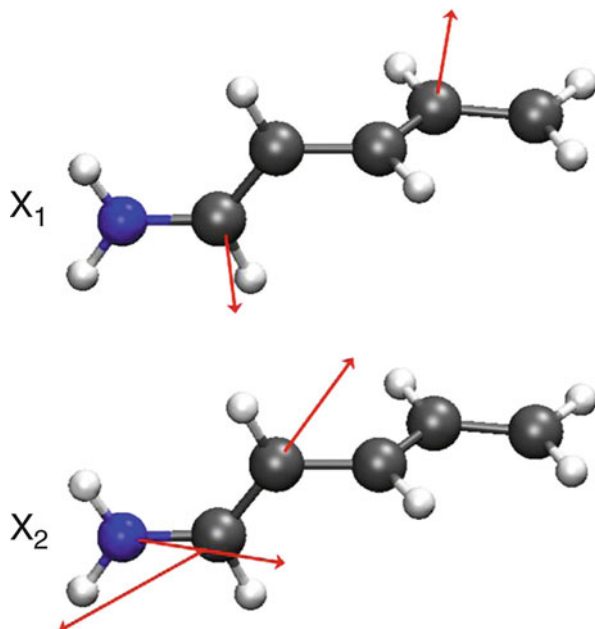
The protonated Schiff base of 11- cis-retinal (PSB11) is the chromophore in the visual receptor rhodopsin. A light-induced cis/trans isomerization of the chromophore triggers the primary event in vision, which is a series of conformational changes of the protein. This photoreaction is considered as the archetype of a chemical reaction optimized by nature to achieve a specific molecular response. Hence, it was in focus of numerous computational and experimental investigations.

The first detailed structure of the excited and ground state potential energy surfaces of the rhodopsin chromophore model was obtained for the three double-truncated 11- cis-retinal (3db-PSB11). In particular the structure of the excited and ground state reaction path branches has been fully elucidated. Furthermore the reduced dimension of the model (Scheme 1) has allowed for the computations of ab-initio CASSCF semi-classical trajectories and the evaluation of the excited state lifetime and time scale of the photochemical isomerization (Weingart et al. 2004).



Scheme 1 Reduced models of PSB11

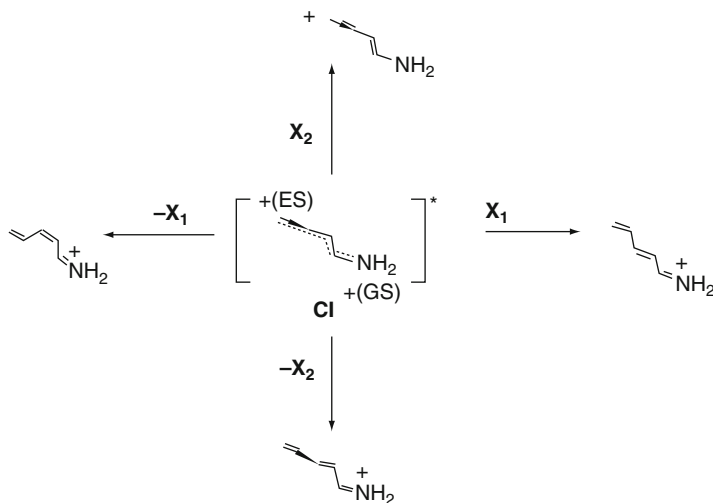
Fig. 7 Branching (or g,h) plane vectors for the CI structure of 3db-PSB11. The X_1 and X_2 vectors correspond to the derivative coupling (or non-adiabatic coupling) and gradient difference vectors between the S_1 and S_0 states (Redrawn with permission from ref. Migani and Olivucci 2004)



The results demonstrated that 3db-PSB11 provides a reasonable model for more realistic structures. In particular, the two-state, two-mode nature of the reaction coordinate computed and observed (both in solution and in the protein) is maintained in the minimal model and the computed ultrafast excited state dynamics is still characterized by two different timescales corresponding to a the initial stretching relaxation (i.e., an inversion of the single bond/double bond length alternation) and to the following torsional deformation (about the central $C_2 - C_3$ bond), respectively.

In Fig. 7 we plot the branching plane vectors (X_1 and X_2) at the conical intersection of 3db-PSB11 model. The conical intersection structure features one highly twisted double bond (about 92°) and involves two electronic configurations, an ionic and a covalent state, that differ for the transfer of one electron between the $C_5 - C_4 - C_3 -$ and $-C_2 - C_1 - N$ fragments.

From the structure of the branching plane it is apparent that in this molecule X_1 and X_2 describe two types of processes. As shown in Scheme 2, motion along the X_1 corresponds to a coupled pyramidalization (wagging) modes at the C_1 and C_4 centers of the π -chain. This motion allows for a widening of the $C_4 - C_3 - C_2 - C_1$ dihedral angle leading to a π -bond breaking process. The X_2 mode is characterized by a stretching deformation (a double bond expansion and single bond contraction mode) of the $N = C_1 - C_2 = C_3 - C_4 = C_5$ chain segment. Thus, motion along the X_2 direction would ultimately yield two structures that be represented by (resonance) formulas with inverted single and double bonds and with the positive charge shifted from the N-terminal to the C_5 -terminal. These

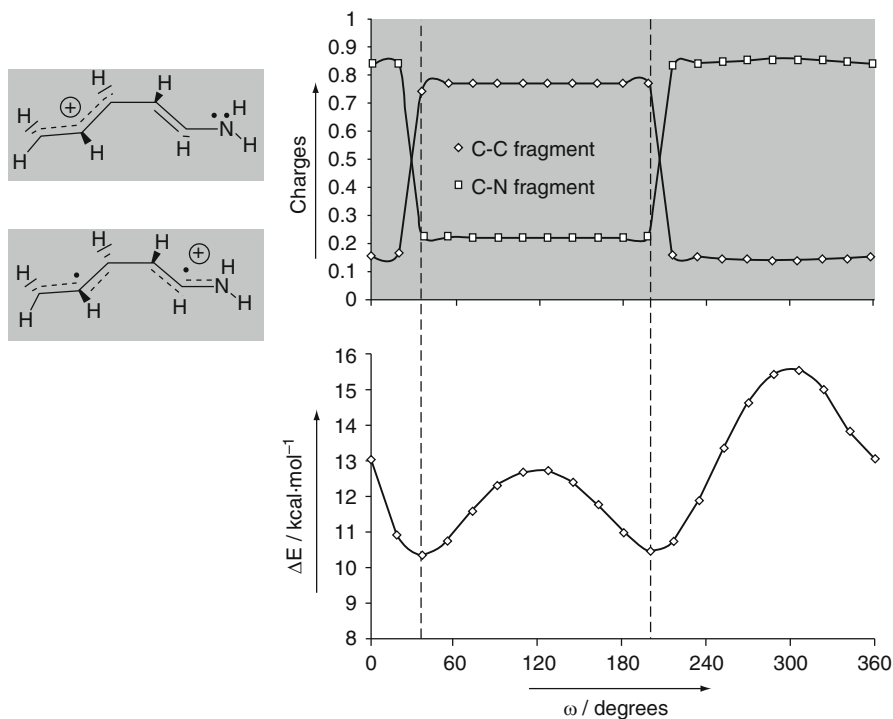


Scheme 2 Deformation of the CI geometry of 3db-PSB11 along the branching plane vectors X_1 and X_2

two 92° twisted structures will be less stable than those generated by displacement along the wagging mode since the deformation along X_1 allows for reconstitution of the central double bond providing strong coupling with the Z/E double bond isomerization coordinate. Thus, structural analysis of the branching plane suggests that upon decay from CI the molecule will generate the Z and E stereoisomers.

The analysis of the wavefunction, taken together with the analysis of the branching plane, provides the basis for the rationalization of the electronic structure of the ground state energy surface comprising the reactant and product valleys (and, eventually, the transition structures connecting them). The result of such an analysis for chromophore model 3db-PSB11 is shown in Scheme 3, where the wavefunction is analyzed in terms of point charges of the $C_5 - C_4 - C_3-$ and $-C_2 - C_1-$ N fragments along a loop centered on the CI (angle ω) and lying on the plane defined by the X_1 and X_2 modes. The charge distribution of the system demonstrates the existence of two different regions. The first region $0^\circ < \omega < 30^\circ$, $200^\circ < \omega < 360^\circ$ is characterized by a structure where the charge is mainly localized on the N-terminal part of the molecule. The second region $30^\circ < \omega < 200^\circ$ is characterized by a structure where the positive charge is mainly located on the C-terminal part of the molecule. The border between the two regions corresponds to the electron transfer events between the two fragments. Notice that the wavefunction changes are associated with the two minima in the energy gap diagram.

As we have previously underlined, the low-lying conical intersections could be only provided through the computation of the photochemical minimum energy path (MEP). However, some cases have been documented where excited state reaction path does not necessarily hit the lowest energy point belonging to the



Scheme 3 Characterization of the topology around the CI of 3db-PSB11 by charge distribution of the model chromophore and the energy profile

intersection space (IS) and the decay may not occur in this region. One of these cases concerns the excited state relaxation path of the 3db-PSB11 chromophore. Indeed, the mapping of the low-lying segments of the IS for this chromophore (see Fig. 8), by means of constrained MEP computations, demonstrated that it ends at a conical intersection with a ca. 70° (CI_{70°) twisted structure. The intersection space remains then coincident with the reaction path up to the lowest energy intersection (CI_{92°) that has a 92° twisted structure (Migani et al. 2003). Notice that in this situation the main locus of excited state (S_1) decay is predicted to be CI_{70° .

In the follow-up study an extended model with four double bonds (4db-PSB11 in Scheme 1) confirmed the topological features found in 3db-PSB11 (Migani et al. 2004). But it also shows expansion of the intersection space segment because of an increased number of double bonds. In the 4db-PSB11 model there are two central double bonds. Twisting each of them at a time and both simultaneously on the excited state can lead to CIs. One can expect a topology of the IS space where these singly twisted CIs are connected by the doubly twisted conical intersections. Indeed, two singly twisted CIs could be identified via IS steepest descent path (ISDP) calculations starting from the doubly twisted CI-geometry T,T- CI (Fig. 9). For nomenclature of these structures, the first letter describes the C2–C3 and the

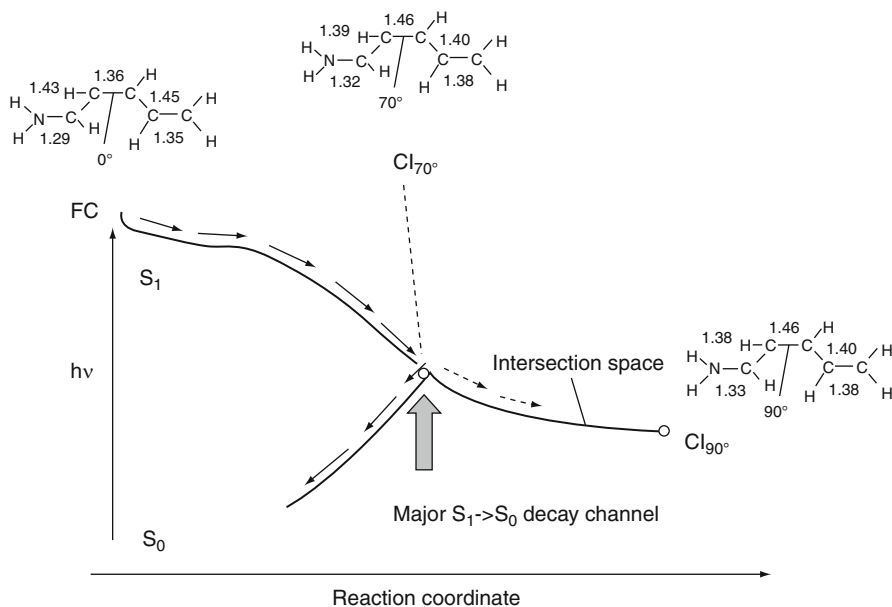


Fig. 8 The excited state reaction path of the cation 3db-PSB11 intercepts the conical intersection point CI_{70° located ca. $5 \text{ kcal} \cdot \text{mol}^{-1}$ above the minimum energy conical intersection CI_{90° . The values of the relevant structural parameters are given in Å and degrees (Redrawn with permission from ref. Migani and Olivucci 2004)

second the C4–C5 double bond conformation, with T being a nearly 90° twisted dihedral. It shows that one can smoothly change the amount of twisting about the C2–C3 and C4–C5 bonds without leaving the IS.

In principle, two additional paths should also connect the complementary isomers of the singly-twisted geometries with the doubly twisted structure. Starting at the same T,T- CI point these two CIs were found (Fig. 10) pointing to T,T- CI serving as a saddle point of index 2 on the IS and not as a transition state. In Fig. 10 we show that a 3D plot of the four ISDPs suggests the existence of a low-lying, two-dimensional cross section (i.e., a surface) of the IS space spanned by CI structures characterized by different degrees of twisting of the two central bonds. This 4db-PSB11 model extends the picture of the S_1/S_0 intersection due to the fact that the bottom of the S_1 energy surface of Rh is spanned by a fairly extended IS segment along the C10–C11–C12–C13 torsion from 80° to 110° . It should be noticed that certain substituents or a suitably designed molecular environment can change the energy profile favoring one isomerization channel over the other or even stabilizing the central part of it.

A more realistic chromophore model including all five double bonds of the polyene chain of retinal (5db-PSB11 in Scheme 1) was studied by Gonzalez Luque et al. (2000). A minimum energy path (MEP) was computed at the CASSCF level

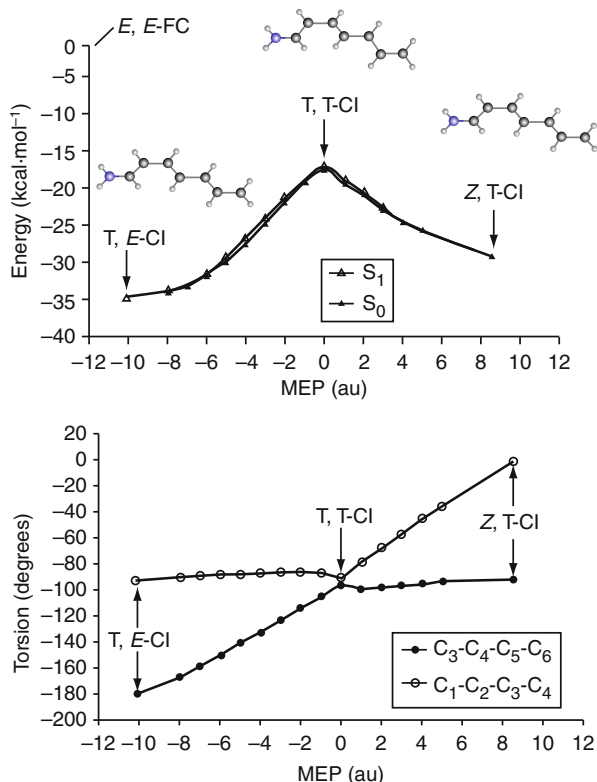


Fig. 9 Left: CASSCF energy profile along the IS coordinate connecting the chemically distinct CI minima T, E and Z,T. The energy values are relative to E, E-4db-PSB11, the most stable isomer. Right: IS coordinate analysis along the corresponding path

of theory. Ten π -electrons included in π -type orbitals with 6-31G* basis set were employed. It should provide an unambiguous cis/trans photoisomerization coordinate for PSB11 (and PSBT) under isolated conditions. This coordinate suggests the dynamic behavior illustrated in Fig. 11, where a metastable species performs many skeletal oscillations along an energy plateau before the reactive torsional modes get fully populated. The plateau may be assigned to the picosecond “fluorescent state” observed in solution and therefore drastically reduces the efficiency of the photoisomerization compared to the protein environment. A second basic feature of the computed path is the positive-charge translocation along the chromophore framework. The positive charge was found to shift partially from the N end to the C end of the models upon $S_0 \rightarrow S_1$ photoexcitation and S_1 in-plane relaxation. The following twisting deformation leads to a complete translocation of the charge that localizes on the fragment containing the C end as already demonstrated for the smaller model 3db-PSB11.

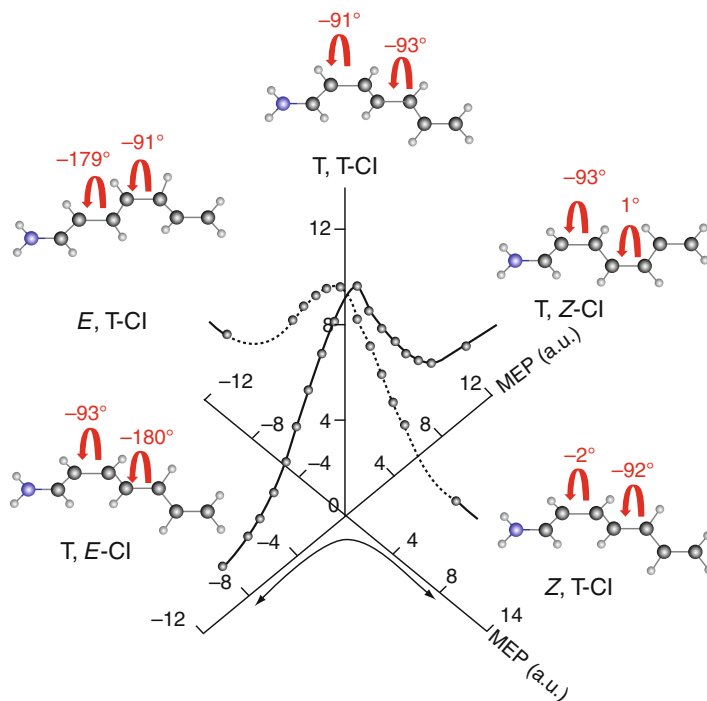


Fig. 10 A 3D representation of the ISDPs connecting T, T-CI to the four IS local minima of 4db-PSB11. The curved arrow marks the T, E-CI \leftrightarrow T, T-CI \leftrightarrow Z, T-CI profile shown in Fig. 9 left

Photofragmentation Through a Conical Intersection: The Photodenitrogenation of a Bicyclic Azoalkane

Photo-induced denitrogenation of 2,3-diazabicyclo[2.2.1]hept-2-ene (DBH) has been the subject of debate ever since it was first reported by Solomon et al. (1968). The observed stereoselectivity in the nitrogen extrusion and housane formation processes was of particular interest. The conical intersection structure which mediates the C–N α -cleavage is characterized by a linear-axial arrangement at the NNC moiety where one of the two CN bonds is still intact (1.48 Å) (Page and Olivucci 2003; Yamamoto et al. 1998). The branching plane vectors X_1 and X_2 correspond to two orthogonal bending motions of the NNC angle, as illustrated in Fig. 12.

When the CI structure is displaced along X_1 and X_2 , the CNN bending motions prompt the formation of the ground state diazenyl diradical (^1DZ in the Scheme 4) either in the exo, endo, or endo-exo forms, as illustrated in Scheme 5.

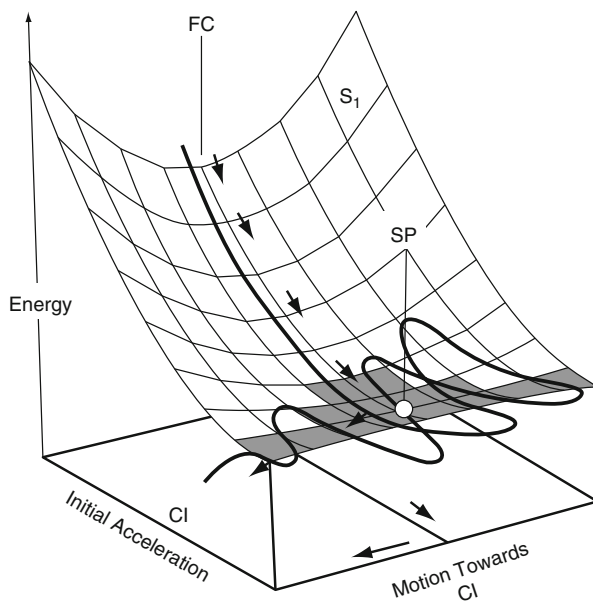


Fig. 11 Shape of the FC → SP region of the S₁ energy surface of 5db-pSB11 model

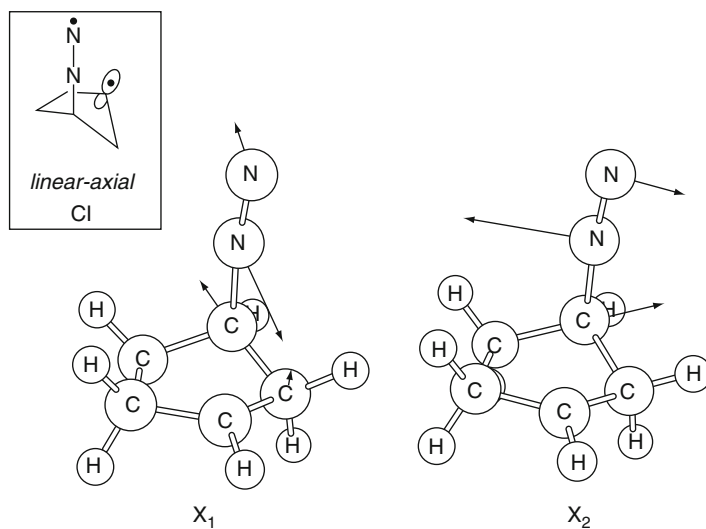
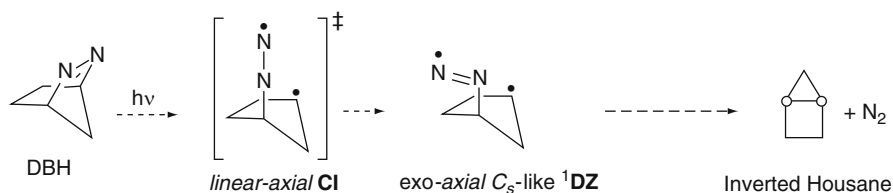
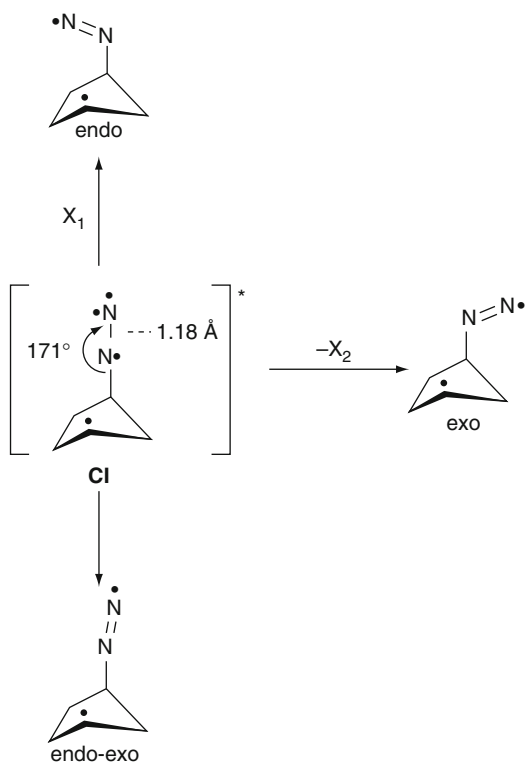


Fig. 12 Gradient difference (X₁) and derivative coupling (X₂) vectors for the CI structure of DBH (illustrated schematically in the inset). The vectors correspond to two orthogonal bending motions of the NNC angle (Redrawn with permission from ref. Sinicropi et al. 2003)



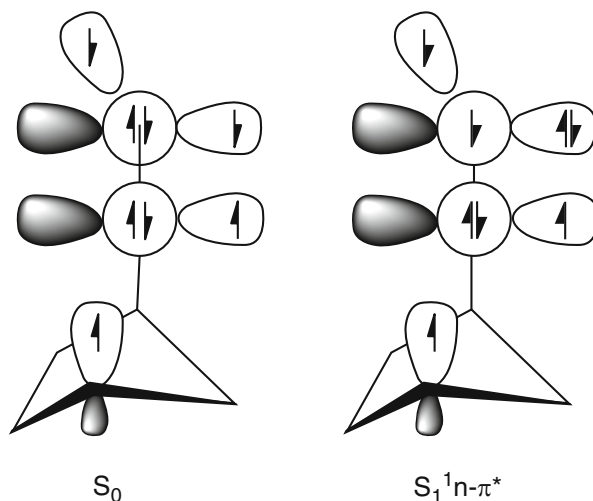
Scheme 4 Photofragmentation of DBH to the Inverted Housane

Scheme 5 Formation of different isomers starting from the diradical CI structure



The electronic structure of the intersecting states of the α -CN bond cleavage of DBH is illustrated in Scheme 6. The electronic configuration of the S₁ state of the azoalkane corresponds to a tetraradical configuration where two electrons are localized in each α -CN σ -orbital, one electron in the excited nitrogen lone pair, and one electron residing in the p-orbital nitrogen bound to the α -C. On the other hand, the electronic configuration of the ground state azoalkane consists of a biradical in which the two α -CN σ -orbitals are singly occupied.

Reaction path calculations explain the observed inversion of stereoselectivity. It is ascribed to an impulsive population of a vibrational mode that triggers axial-to-equatorial ring inversion. This idea is supported by classical molecular dynamics



Scheme 6 Electronic structure of the S_0 and S_1 state

simulations. Figure 13 reveals that after a single oscillation (on a 40 fs timescale) in the direction of a transient bicyclic intermediate, the molecule reaches a strained structure that it can only relax following the initial direction of motion. After 60 fs, the exo-axial 1DZ configuration is reached, followed by the axial to equatorial transition structure on an 80 fs timescale. Within 100 fs, the inverted configuration is observed.

Charge Transfer and Intermolecular Hydrogen Transfer Mediated by a Conical Intersection: Quenching the Fluorescence of Bicyclic Azoalkanes

In contrast to DBH, the so-called “reluctant” azoalkanes are inert to photochemical denitrogenation. One representative system is the 2,3-diazabicyclo[2.2.2]oct-2-ene (DBO, Scheme 7). DBO exhibits a long-lived excited n, π^* -singlet state (on the order of μ s).

Using the 1,2-diazacyclopent-1-ene (pyrazoline) as a model of DBO allows the use of accurate yet computationally demanding (expensive) methods. In combination with experimental evidence, two mechanisms for quenching $^1n, \pi^*$ states were proposed (Sinicropi 2002; Sinicropi et al. 2000, 2001, 2002). DBO is efficiently quenched by hydrogen donors (for instance by solvents such as chloroform, methanol, and benzene) either via a concerted or a stepwise process. DBO luminescence can be quenched by electron donors (e.g., with triethylamine), however, only through a concerted process. Calculations have been carried out using CH_2Cl_2 and methanol to model hydrogen donors, and trimethylamine/dimethyl ether as

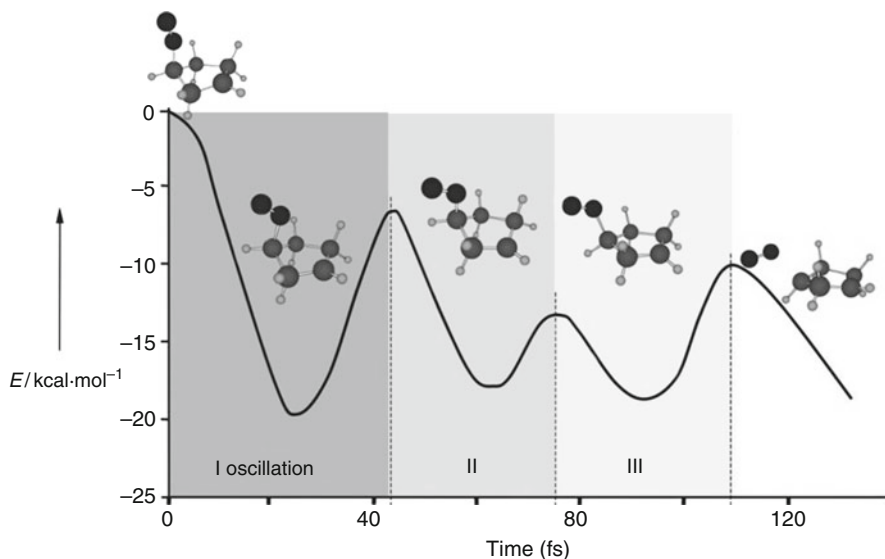
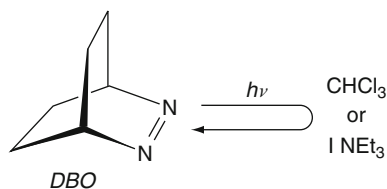


Fig. 13 Triplet DFT energy profiles along the molecular dynamics simulations started in the vicinity of the CI structure of DBH. The snapshots demonstrate the molecular changes along the trajectory (This figure is redrawn with permission from ref. Sinicropi et al. 2003)

Scheme 7 Internal conversion of DBO



strong/weak electron donor solvents on the other hand. As shown in Fig. 14, both quenching routes involve bimolecular photochemical reactions. Deactivation occurs through an S_1/S_0 CI, located halfway along the reaction coordinate, and prompting bifurcation. The first branch (represented by full arrows) resembles a photophysical transformation and has been termed an “aborted” photochemical transformation. The second branch (represented by light arrows) is associated with production of a transient species that reverts to the original starting reactants by passage through a low-lying transition state located on the ground state reaction coordinate (depicted by a dashed curve). Thus, both routes correspond to a photochemical transformation, which was initiated but never achieved.

The conical intersection for the hydrogen abstraction mechanism (in the presence of CH_2Cl_2) is characterized, with respect to the reactant pair, by shortening the H–N distance (from 2.22 to 1.02 Å) and simultaneously lengthening the C–H distance (from 1.07 to 2.01 Å). The structure on the conical intersection for this bimolecular

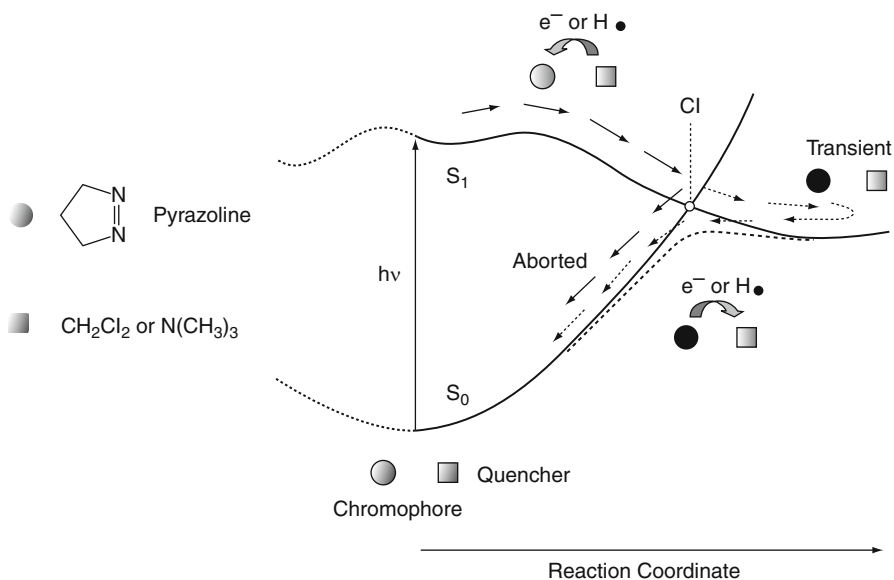


Fig. 14 Potential energy diagram showing the interplay between ground (S_0) and excited (S_1) state surfaces in the fluorescence quenching of ${}^1n, \pi^*$ state chromophores due to an hydrogen donor or electron donor species. The full and light arrows describe the concerted and stepwise energy wastage route

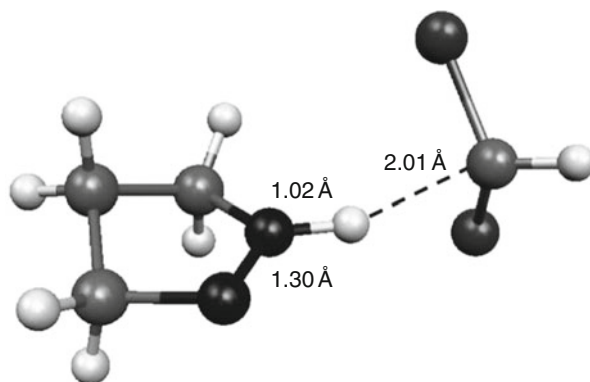
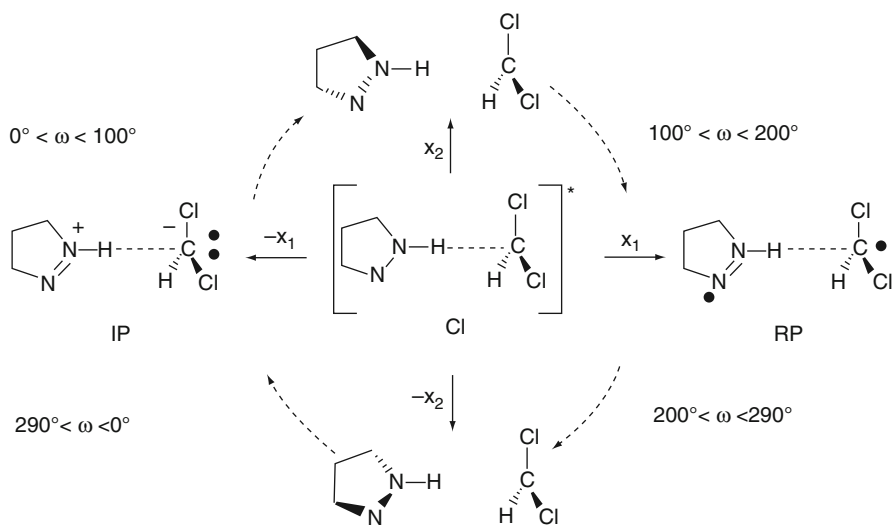
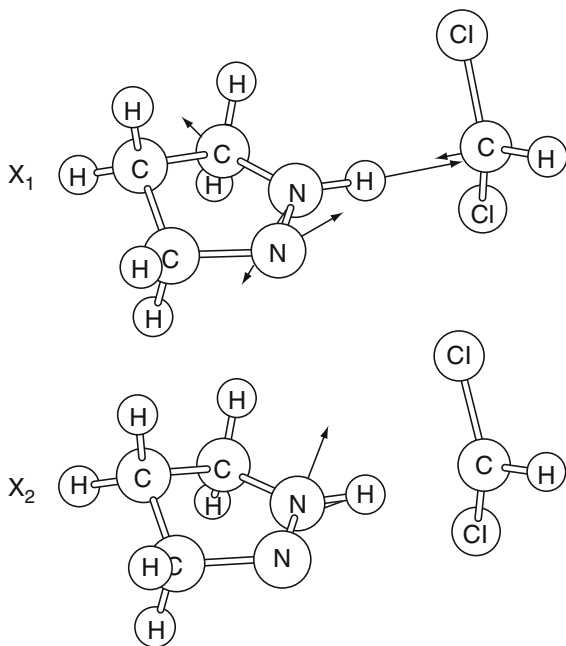


Fig. 15 The molecular structure of the characterized conical intersection that mediates hydrogen abstraction in the presence of CH_2Cl_2

reaction is given in Fig. 15, whereas the degeneracy-lifting vectors are illustrated in Fig. 16.

Motion along X_1 stretches the intermolecular N-H distance whereas displacement along X_2 causes an out-of-plane pyrazoline ring distortion, both translations lifting the degeneracy between the two states. Structural evolution along both displacements is illustrated in Scheme 8. Translation along X_1 leads to the radical

Fig. 16 X_1 and X_2 vectors for the CI structure of pyrazoline + CH_2Cl_2 involved in the hydrogen abstraction mechanism (Redrawn with permission from ref. Olivucci and Sinicropi 2005)



Scheme 8 Characterization of the topology around the CI of associated pyrazoline-dichloromethane model pair

pair (RP) while displacement along $-X_1$ leads to production of an unstable ion pair (IP). Evolution in the X_2 and $-X_2$ directions leads to two equivalent ground state transition structures, featuring a distorted pyrazoline ring.

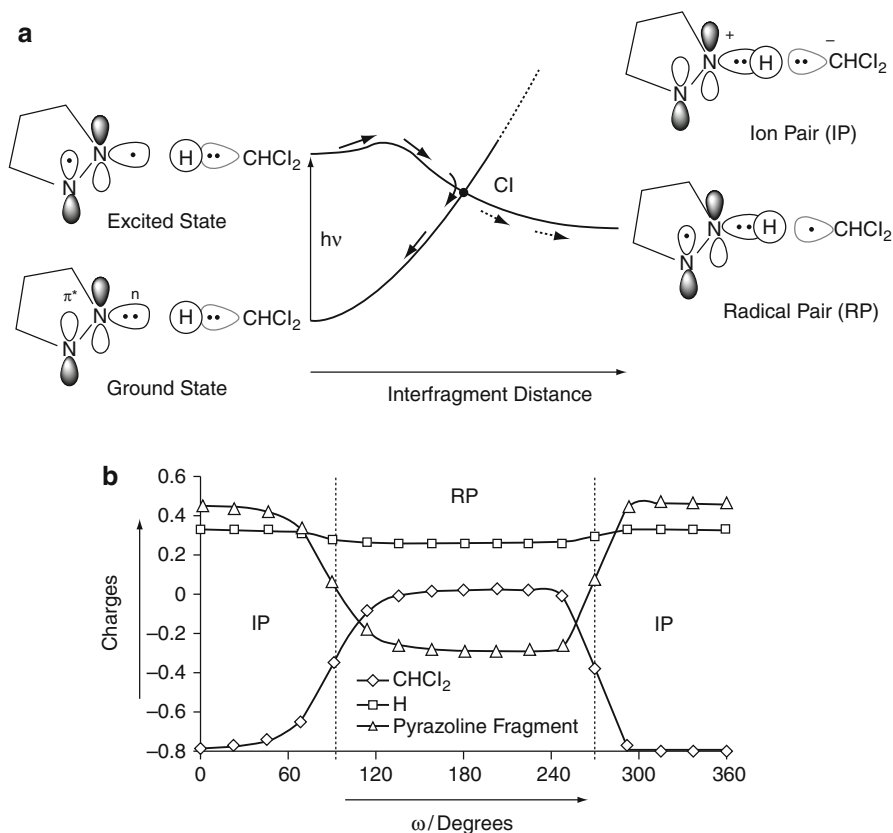


Fig. 17 (a) Modified state correlation diagram of the n, π^* -excited state (ES) of pyrazoline + CH_2Cl_2 correlating with the radical pair (RP) derived from hydrogen atom abstraction and the ground state (GS) correlating with the ion pair (IP) derived from proton abstraction. (b) S_0 fragment charges [au] along a circular cross section centered at CI and spanning the branching plane (pyrazoline fragment, open triangles; hydrogen atom, open squares; CHCl_2 fragment, open circles) (Redrawn with permission from ref. Sinicropi et al. 2001)

An analysis of the wavefunction indicates that the n, π^* -excited state correlates with the radical pair structure (derived from a complete hydrogen atom abstraction), whereas the ground state correlates with the ion pair (derived from proton abstraction), see Fig. 17a. Although this case study deals with bimolecular photochemical processes, it is in a sense similar to the retinal PSB models as the electronic configurations describing the intersecting states are interchanged by charge transfer. The charge distribution of the system along a circular cross section – spanning the plane defined by the X_1 and X_2 vectors and centered around CI – is displayed in Fig. 17b with ω expressing the rotation. It reveals the existence of two distinct regions. The first region ($0^\circ < \omega < 100^\circ$ and $290^\circ < \omega < 360^\circ$) is characterized by an ion pair structure where the charge on the pyrazoline is positive while it is

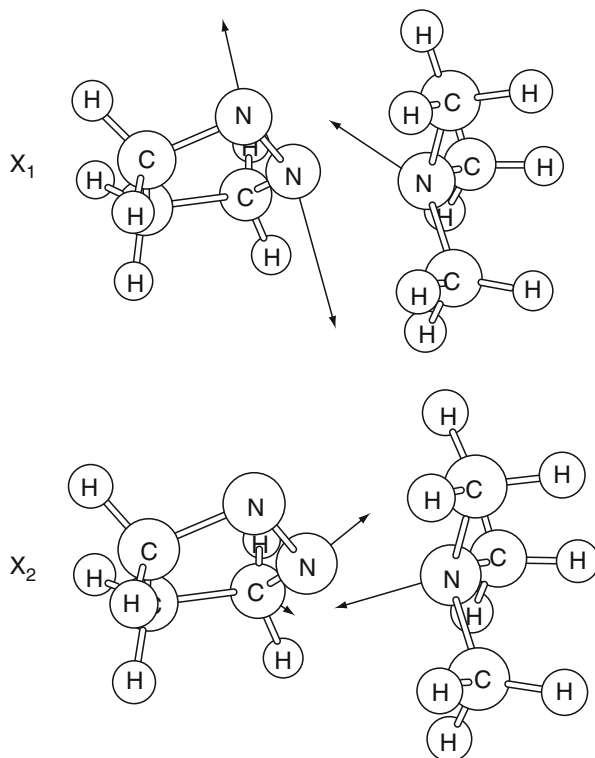


Fig. 18 X_1 and X_2 vectors for the CI structure of pyrazoline + trimethylamine involved in the charge transfer process (Redrawn with permission from ref. Olivucci and Sinicropi 2005)

negative on CHCl_2 . The second region $100^\circ < \omega < 290^\circ$ is characterized by a covalent structure. The border between the two regions corresponds to two sudden electron transfer events, one from the CHCl_2 anion to the pyrazoline cation yielding the RP configuration ($\omega = 90^\circ$) and the second in opposite direction ($\omega = 270^\circ$).

The molecular structure of the conical intersection along with the degeneracy-lifting vectors of the branching plane for the charge transfer process is given in Fig. 18. Notice that X_1 is dominated by an out of plane deformation of the pyrazoline ring whereas X_2 is dominated by the interfragment distance. The computed photochemical reaction path demonstrates that the excited state branch of the photochemical reaction path is dominated by an decrease in the distance between the pyrazoline and $\text{N}(\text{CH}_3)_3$ fragments. After a low-lying excited state barrier, the evolution along the path leads to the formation of an exciplex located in the vicinity of a conical intersection. The intersection is accessed when the distance between the pyrazoline and nitrogen atoms of the amine is ca. 2\AA . In the exciplex, the computed amount of charge transfer from the trimethylamine lone pair to the excited state half-vacant nonbonding orbital of one pyrazoline nitrogen atom is about 0.3 electrons.

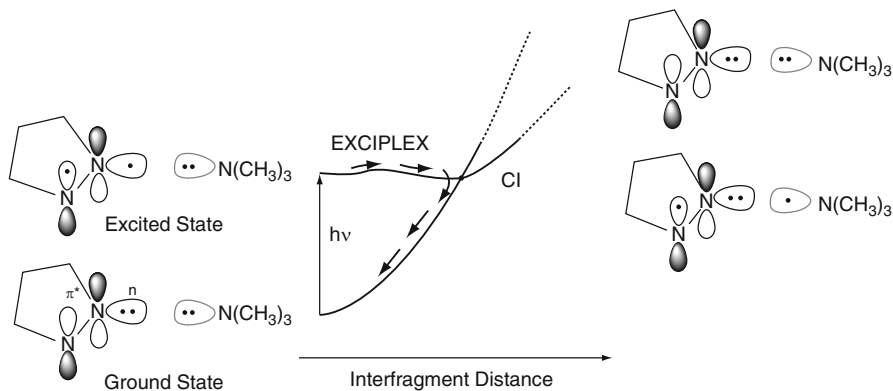


Fig. 19 Modified correlation diagram for the interaction of the n, π^* -excited state of pyrazoline with an electron donor such as the trimethylamine reflecting the occurrence of an exciplexes and a conical intersection along the reaction pathway (Redrawn with permission from ref. Sinicropi et al. 2000)

The exciplex $N \dots N$ two-orbital/three-electron bond can be viewed as a mixture of a covalent ($N = N^{\bullet} \dots : NMe_3$) and an ionic ($N = N^{\bullet-} \dots ^+ NMe_3$) electronic configurations. A steep rise of the ground state energy surface towards the conical intersection is due to a destabilizing two-orbital/four-electron repulsive interaction ($N = N : \dots : NMe_3$) (see Fig. 19).

Case Studies of Chromophores Embedded in a Molecular Environment

Green Fluorescent Proteins (GFP)

In 1961 the Green Fluorescent Protein (GFP) was discovered in the jellyfish *Aequorea Victoria* and became the first representative of the same named protein family (Shimomura 2009). Its chromophore is p-hydroxybenzylideneimidazolinone (HBDI), enclosed in a β -barrel structure. It emits a green fluorescence light if exposed to blue light. The corresponding fluorescence spectrum shows a single peak at 508 nm attributed to the anionic form of the chromophore. In contrast, in water solution the HBDI anion shows a blue-shifted fluorescence that has a very short lifetime.

The GFP chromophore has been studied extensively, both by means of experimental and theoretical methods. Sinicropi et al. (2005) performed the first QM/MM simulation of GFP suitable for the quantitative evaluation of λ_{\max}^a and λ_{\max}^f . Calculations were done at the CASPT2//CASSCF/CHARMM level of theory with the 6-31G* basis set because a previous study (Martin et al. 2004) suggested that dynamic electron correlation is important to reproduce the observed quantities. The

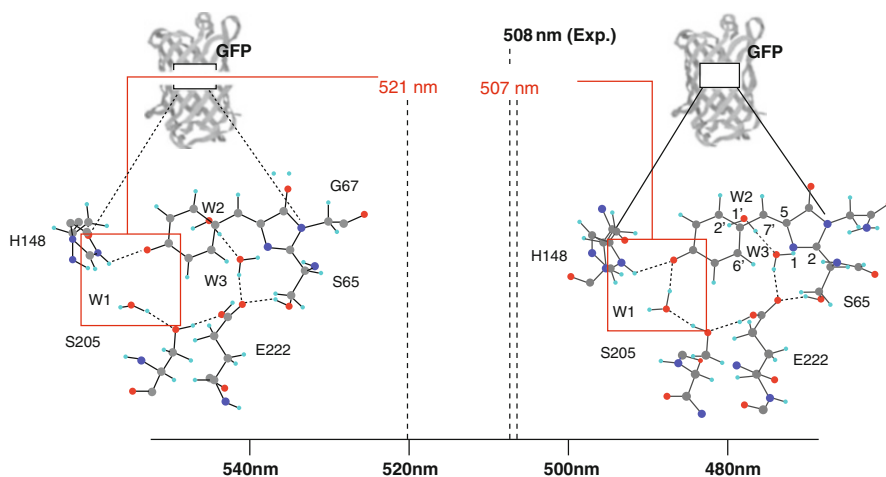


Fig. 20 Right: Chromophore cavity of the model for the GFP fluorescent state I displaying the hydrogen bond network and the values of the computed λ_{\max}^f value. Left: The change in the hydrogen bond network by disruption of the W1–chromophore bond leads to a red shift in the predicted λ_{\max}^f

λ_{\max}^a was evaluated for the anionic ground state form in the emitting state (state I) and in a modified I state with a disrupted hydrogen bond between the chromophore and the closest water molecule (see Fig. 20).

For the full protein model of the I state the calculated peaks for absorption and emission are 468 and 507 nm, respectively, compared with the experimental values of 495 and 508 nm (Brejc et al. 1997; Chatteraj et al. 1996) resulting in an error of less than 4 kcal · mol⁻¹. A somewhat better agreement with experimental data is achieved by Altoe et al. (2007) using a similar computational protocol. The calculations were performed at the CASPT2//CASSCF/AMBER level for the chromophore with the same basis set. Prior the QM/MM excited state calculation the model was equilibrated by MD simulation. The authors calculated the absorption of 504 nm for the I state (experimentally 495 nm).

The GFP model (Sinicropi et al. 2005) shows that there is a charge translocation towards the imidazolinone moiety in the I state, indicating that the excited state is a charge transfer state. It is shown that the hydrogen bond between the proximal water molecule and the phenoxy oxygen of the chromophore localizes the negative charge on that oxygen atom. In fact, as reported in Fig. 20, removing that water molecule has a great effect on structure and energy, which leads to a less stable structure (20 kcal · mol⁻¹ higher in energy). A different effect is found when comparing a protein model with the positively charged counter-ion Arg96 to a protein model without it: the protein environment balances the counter-ion effect to the same extent in S₀ and S₁. A similar quenching effect is also observed for Rh (Strambi et al. 2007) (see below) despite its very different protein and chromophore structure.

Another member of the GFP protein family is asFP595, which was discovered in *Anemonia sulcata* (Quillin et al. 2005). It fluoresces brightly at 595 nm after intense illumination at 568 nm at room temperature. The X-ray crystal structure was obtained in 2005 for the dark state at 1.38 Å resolution. The chromophore was discovered to be the same as in GFP, and the initial hypothesis for explaining the fluorescence was a Z/E photoisomerization because the chromophore cavity can easily accommodate both the isomers, as was seen from the crystal structure. The first QM/MM study on asFP595 was carried out by Grigorenko et al. (2006). The ground state equilibrium structure was determined at the RHF/6-31G^{*} and B3LYP/6-31 + G^{*} level while the excited state properties were studied at the CASSCF/AMBER level with the 6-31 + G^{*} basis set (with a 10 electron/10 orbital active space). The QM region included the chromophore unit and the sidechains of the His197, Glu215, and Arg92 residues, all of which are involved in the hydrogen bonding to the chromophore. For the HBDI moiety three options were considered: the E anion, the Z anion, and the E zwitterion. The protein environment was shown to stabilize the zwitterion with respect to the gas phase, which confirmed the importance of evaluating the role of the protein matrix. The CASSCF vertical transition energies give support to the experimentally based hypothesis that the dark state features an anionic or zwitterionic E conformation, while the emitting state features an anionic Z-isomer formed upon light absorption by the E ground state. However, it is known that the CASSCF vertical excitation energies are far from being quantitative and may change dramatically for transitions to charge transfer states upon CASPT2 correction.

A 2007 study by Schäfer and coworkers (Schaefer et al. 2007) shed light on the photochemical reaction path in asFP595, finding many similarities with GFP. They used TDDFT to establish the general features of the photocycle: after photon absorption the zwitterionic E dark state chromophore is converted into the anionic Z form by proton transfer from the imidazolinone ring to the Glu215 side chain and E → Z isomerization. A proton wire connects the chromophore cavity to the external solution in order to achieve protonation of the phenolate oxygen, which leads to the neutral Z form. Another proton wire is able to release protons from His197 to the solution along five residues and a buried water molecule. The proton transfer seems to be important in preventing the reverse photochemical reaction that would result from absorption of a second photon. Photoisomerization leading to changes in protonation states seems to be a general principle exploited by evolution: both bR and PYP rely on the same mechanism. Subsequently, the CASSCF method (Schaefer et al. 2008) was employed to gain deeper insight into the mechanism and its features at the atomic level. A QM/MM MD simulation was performed, evaluating energies and forces of S₁ and S₀ at the CASSCF/3-21G level with a six electron/six orbital active space to reduce the computational cost. The QM region included only the chromophore, and the OPLS force field was used for the rest of the system. The results confirmed the conclusions of the previous paper in a more detailed way: the proton distribution in the active site controls the possible photochemical pathways of the protein-embedded chromophore. There are six different states, arising from the combination of E and Z conformations with a

neutral, anionic, or zwitterionic chromophore. To switch from the non-fluorescent to the fluorescent state an $E \rightarrow Z$ isomerization is required. It has been observed to happen only for the neutral chromophore which is poorly populated. The fluorescent state is the zwitterionic Z form reached upon protonation of Glu215, but this state is also sparsely populated. These two findings are in agreement with the observed low overall photoisomerization quantum yield of asFP595. Moreover, excitation of the neutral E chromophore should be performed with green light otherwise the blue-shifted Z photoproduct would isomerize back. The neutral Z state is more populated than the other Z forms, so the efficiency of the back reaction, i.e., the switching off, is very high.

Understanding the Spectral Tuning in Retinal Proteins

Rhodopsin, the dim-light photoreceptor in vertebrate retina, belongs to the G-protein-coupled receptor family. The 11- *cis*-retinal (PSB11) that is covalently bound to Lysine 296 via a Schiff-Base link serves as the chromophore (Fig. 2). Upon light absorption PSB11 undergoes a stereoselective Z/E isomerization to the all- *trans* isomer, which leads to a protein-wide conformational change that initiates light perception. As summarized in Fig. 21, this is an ultrafast photochemical reaction featuring a ~ 100 fs lifetime of the excited state and a 200 fs appearance-time of the vibrationally hot intermediate photorhodopsin (photo-Rh). The first isolable intermediate, bathorhodopsin (batho-Rh), is detected on a 2 ps time scale. Below we review the results of recent research efforts to investigate the mechanism of the excited state isomerization in Rh using a QM/MM model.

More than 30 years ago Warshel proposed, on the basis of semiempirical simulations, an isomerization mechanism that could explain how this process can occur in the restricted space of the Rh binding pocket (Warshel 1976). Since two adjacent double bonds were found to isomerize simultaneously the mechanism reveals a so-called bicycle pedal motion. Due to the concerted rotation of two double bonds in opposite directions the overall conformational change is minimized and hence this mechanism was found to be space-saving. The empirical valence bond (EVB) method (Warshel and Levitt 1976) was used to compute the excited state potential energy surface of the chromophore during a trajectory calculation where the steric effects of the protein matrix were modeled by specific restraints on the retinal atoms. Since then, Warshel and his coworkers have improved the model employing better structural data and new computational developments (Warshel and Barboy 1982; Warshel and Chu 2001; Warshel et al. 1991). The main refinement of the bicycle pedal mechanism was that the simultaneous rotation of the adjacent double bonds is aborted at a twist of 40° and leads to the isomerization of only one bond (Warshel and Barboy 1982).

Since the crystal structure of the rhodopsin protein became available in 2000, a more accurate description of the protein environment was available for computational investigation and for the interpretation of the experimental data. In 2003 Ferré and co-workers performed the first QM/MM simulations using the

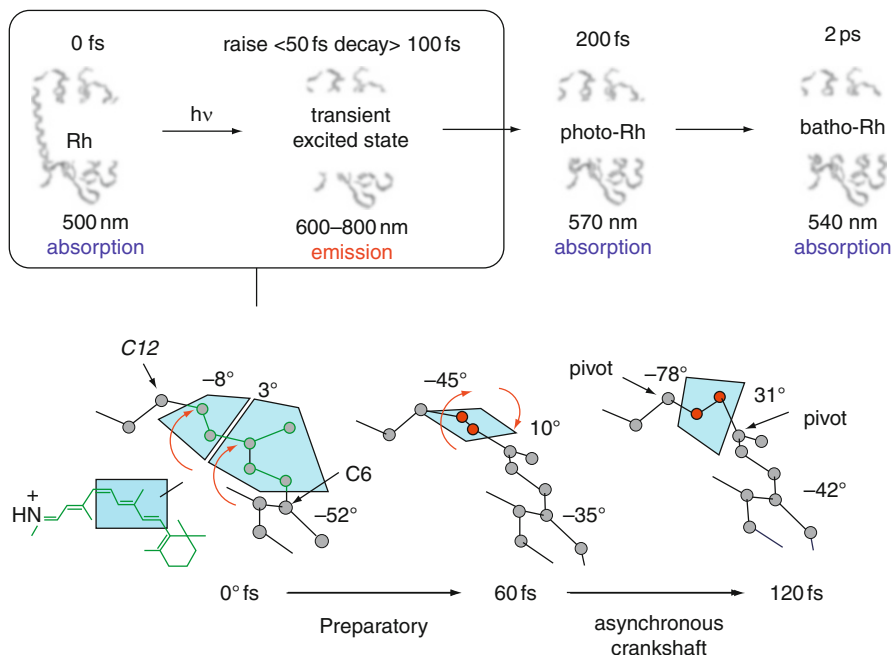


Fig. 21 Top: Observed λ_{\max}^a values and excited state lifetimes associated with the initial steps of the photoisomerization of Rh. Bottom: Schematic representation of the excited state isomerization motion of PSB11 dominated by an asynchronous crankshaft structure deformation (Frutos et al. 2007) and documented for a CASSCF/AMBER model of Rh

ab initio CASPT2//CASSCF/AMBER protocol applied to the protein-embedded retinal chromophore in the excited state (Ferré and Olivucci 2003b). In this study the authors have truncated the retinal chromophore at the QM level to make the calculation feasible. The authors used a five double-bond N to C6 retinal fragment. The CASSCF level of theory was used to determine the ground state equilibrium structure with all 10 π -electrons in 10 π -type orbitals comprising the active space. The protein was modeled with the AMBER force field.

Starting from the 1HZX (Palczewski et al. 2000) crystallographic structure, it was shown that the CASPT2//CASSCF/AMBER protocol can be used to construct a Rh model featuring the full retinal chromophore treated at the QM level. This model could reproduce the stationary spectroscopic features (see Fig. 22 top left for a comparison of the observed and computed values, given in italics and plain text, respectively) with an error similar to the one found for GFP systems described above. For instance, the λ_{\max}^a is reproduced with only 3 kcal \cdot mol $^{-1}$ error (476 versus 498 nm). But also the computation of the corresponding λ_{\max}^a in solution employing the same methodology led to an error within 2 kcal \cdot mol $^{-1}$ of the so-called opsin-shift (the 445 nm λ_{\max} observed for PSB11 in methanol is red-shifted to 498 nm in Rh).

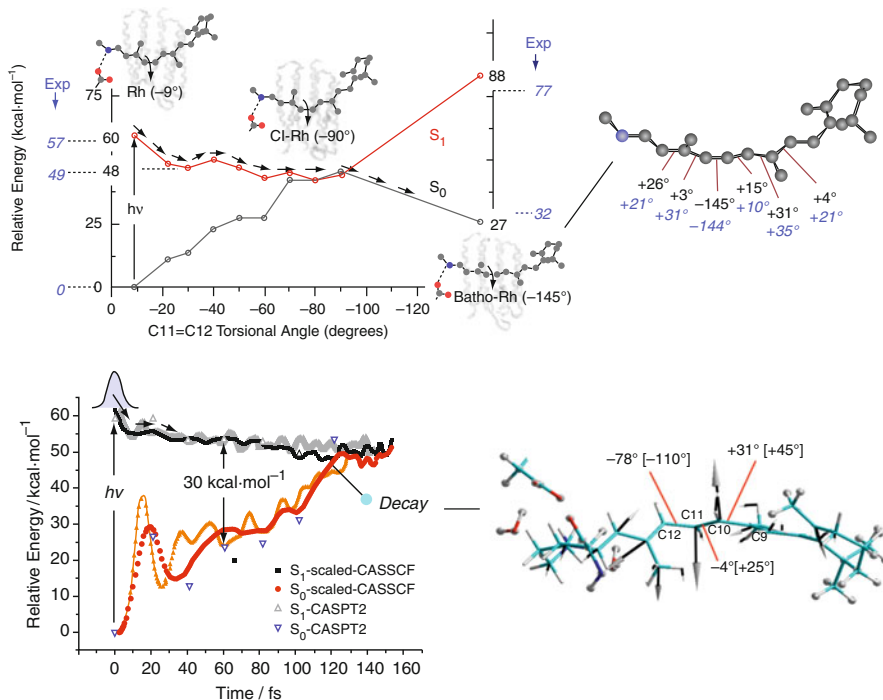


Fig. 22 Top left: Energy profiles along the S₁ photoisomerization coordinate of Rh computed at the CASPT2//CASSCF/AMBER level. The PSB11 structure on the right corresponds to the first isolable intermediate batho-Rh. The values in italics correspond to observed values (See ref. Kukura et al. 2005). Bottom: Scaled-CASSCF S₁ and S₀ energy profiles and CASPT2 points along the S₁ trajectory of Rh. The S₁ and S₀ energy profiles are reported for two Rh models based on the 1HZX (red and black lines) and 1U19 (orange and grey lines) crystallographic structures (M. M. Huntress, M. Olivucci, unpublished data). The structure on the right represents the point of decay (conical intersection structure) reached after ~ 120 fs time evolution. The vectors represent the velocities at the decay point and clearly show a crankshaft-like motion at the center of the PSB11 backbone

Furthermore we have been able to investigate the effect of water relocation inside the Rh cavity (Strambi et al. 2007). A second water (W2) had to be placed into the protein binding pocket according to qualitative space filling and electrostatic considerations since the original 1HZX crystallographic structure only showed the presence of one water molecule (W1). However, the more recent and better resolved crystallographic structure (PDB Code: 1U19) (Okada et al. 2004) points to a different location and different hydrogen bonding for W2. The sensitivity of S₀ → S₁ excitation energies with respect to the position of W2 has been investigated by constructing a new Rh model with the revised water placement. The results from the refined structure indicate a fairly low sensitivity of the λ_{\max}^a to the significant geometrical change of the salt bridge region, which suggests that there is a compensation mechanism that counteracts the effects of this change on the λ_{\max}^a . Our analysis points to a displacement of the chromophore to a cavity region

with a larger positive electrostatic potential as a means to offset the effect of the decrease in NH (+)–O1 (–) salt bridge distance. We have also compared the effect of a reduction of the full opsin model to a model with only the 27 amino acids of Palczewski's cavity (Palczewski et al. 2000), those that surround the retinal. It was found that the value predicted for the $S_0 \rightarrow S_1$ excitation energy is substantially the same indicating that for Rh the charges of distant residues are of minor importance for the optical absorption features of PSB11. Of course, we do not expect this to be a general result. Highly truncated protein models should always be applied carefully until a systematic analysis is carried out to determine their limitations and reliability.

In Fig. 22 top left, we report the S_1 branch of the photochemical reaction path of Rh computed in terms of a relaxed scan driven by the torsional deformation of the reacting bond. Using this data we have located and assigned the S_1 structure that corresponds to transient fluorescent state I. We then located the structure of the lowest lying S_1/S_0 CI (CI-Rh) that also corresponds to the local S_1 minimum. This structure displays a $\sim 90^\circ$ twisted C11–C12 bond. Starting from the CI-Rh geometry, using standard optimization, the first stable ground state intermediate bathorhodopsin (batho-Rh) has been located. It exhibits absorption at 540 nm and features an all- trans-like chromophore structure ($\sim 150^\circ$) dihedral angle around the C11–C12 double bond). The computed structure of batho-Rh can be compared with the structure that was experimentally derived by Mathies and coworkers (Kukura et al. 2005) by femtosecond resolved resonance Raman spectroscopy (see Fig. 22 top right). The comparison indicates that it is possible to predict the structure of an intermediate. Concerning the photon energy of $\sim 32 \text{ kcal} \cdot \text{mol}^{-1}$ that is efficiently stored in batho-Rh, the results of our CASPT2//CASSCF/AMBER energy profile show that we reproduce this quantity with an error of $5 \text{ kcal} \cdot \text{mol}^{-1}$.

A mechanistic picture of the space-saving isomerization mechanism is derived from the reaction path of the Rh QM/MM model and the associated structural deformation. The predominant change in the retinal geometry that occurs immediately after excitation leading to the fluorescent state I is a bond length alteration (BLA) in the $-C9 = C10-C11 = C12-C13 = C14-$ moiety. A complete inversion between single and double bonds is found at the geometry characterizing the state I. The CI-Rh displays a highly helical structure compared with Rh and FS-Rh and is mainly characterized by a large structural change in the $-C9 = C10-C11 = C12-C13 = C14-$ moiety. Thus, the motion driving the $S_1 \rightarrow S_0$ decay is mainly torsional with a rotation of $\sim 60^\circ$ ($20^\circ \rightarrow 90^\circ$) around the C11 = C12 bond, and 37° and 15° twisting around the C9 = C10 and C13 = C14 bonds, respectively. Therefore, from a general point of view, during photoisomerization the structural changes do not occur exclusively at the central double bond but also involve the other two adjacent double bonds, which lead to a global change in the helicity of the chromophore. A mechanism can be derived considering the largest changes. These involve the torsion about the reactive bond C11 = C12 and about the adjacent C9 = C10 bond. As highlighted in Fig. 21 (bottom) these twisting deformations occur in opposite directions and results in the rotation of the $-CH-CH-$ moiety with respect to the remaining framework. The mechanism confirms that the space-saving motion imposed by the tight Rh cavity is of the

crankshaft type. However, since the progression about the C9 = C10 bond is more limited, we can talk of an asynchronous crankshaft mechanism (Frutos et al. 2007). Furthermore, a comparison of the CI-Rh structure with the batho-Rh structure establishes (consistent with the experimental data) that the less twisted C9 = C10 reverts to its original stereochemistry after ground state relaxation, so we can talk of an aborted asynchronous crankshaft mechanism.

Another widely studied retinal protein is bacteriorhodopsin (bR), a light-driven ion pump discovered in *Halobacterium salinarum* in 1971 (Oesterhelt and Stoeckenius 1971). It is a membrane protein, which upon illumination generates and maintains a proton gradient across the cell membrane (Oesterhelt and Stoeckenius 1973). The gradient can be used as a source of energy, for example, by adenosine triphosphate synthase. bR shares some similarities with the visual rhodopsin: seven transmembrane α -helices and a protonated Schiff base with lysine-bound retinal as the chromophore. The X-ray crystal structure was first obtained at 3.50 Å resolution in 1996 (Grigorieff et al. 1996). By 1999, the resolution had improved to 1.55 Å (Luecke et al. 1999). The bR photocycle consists of several intermediates labeled K, L, M, N, and O, which were characterized by spectroscopy and trapped at low temperature in order to be studied by crystallography (Balashov and Ebrey 2001; Lanyi 2000). Light absorption triggers the Z/E isomerization of the C13 = C14 double bond of the PSBAT chromophore. The bond isomerization is completed at the K intermediate, yielding a 13- cis chromophore (PSB13); the following thermal relaxation leads to Schiff base proton transfer to the Asp85 counter-ion, in the L and M steps. The proton flux is completed during the remaining part of the photocycle, along with other accompanying proton transfers and protein structural changes.

The first ab initio QM/MM calculation on bR was reported by Hayashi et al. using multiconfigurational methods to evaluate the excited state properties (Hayashi and Ohmine 2000). The protein model was based on the crystal structure determined by Luecke et al. (1999) where the missing residues were taken from the PDB structures 2BRD (Grigorieff et al. 1996), 1C3W, (Luecke et al. 1999) and 1QHJ (Belrhali et al. 1999). The geometry was optimized at the HF level, using DZV and 3-21G basis sets with polarization functions on oxygen atoms. CASSCF(12,9) was employed to calculate retinal excitation energies. Five different models were built, each with different atoms included in the QM region. The results demonstrated that the protein environment affects the absorption properties of the chromophore. The side chains and the water molecules in the binding pocket affect the geometry of the retinal, forcing a twisting of the C13 = C14 double bond. The protein environment is anionic near the Schiff base, which stabilizes the ground state with respect to the excited state, causing a blue shift of the $S_0 \rightarrow S_1$ excitation energy.

Later the Schulten group (Hayashi et al. 2001) performed a further analysis to elucidate the physical mechanisms of the observed spectroscopic tuning in the rhodopsin family. A comparison between bR and the sensory rhodopsin II (sRII) was done to understand the origin of the change in λ_{\max}^a from 568 nm in bR to 497 nm in sRII. This difference was observed experimentally, in spite of the

same chromophore in both proteins and the similar protein environments. As in the previous work, a QM/MM methodology was employed, with AMBER as the MM force field. The equilibrium geometries of the QM regions were computed at the HF/DZV level. The excitation energies were evaluated with a CASSCF(10,11) wave function. The QM region included the retinal, the retinal-bound lysine (starting from C γ), and parts of the binding pocket: two aspartic acid residues (starting from C β) and three water molecules. The comparison between bR and SRII was performed by calculating the contributions to the $S_0 - S_1$ excitation energy: the total energy was rewritten as a sum of the energy of the isolated chromophore, the electronic reorganization energy arising from the modification of the wave function in the presence of the protein, and the electrostatic interaction energy between protein and chromophore. The results showed that the main contribution is electronic reorganization in the retinal and a closer inspection of the structures showed a reduced counter-ion–Schiff base distance in sRII compared with bR. This is the reason for the blue shift, consistent with the picture given in a previous paper (Hayashi et al. 2001).

Another spectroscopic feature of the bR photocycle was addressed by the same group in 2002 (Hayashi et al. 2002). The early K and L intermediate models were built using MD and refined with QM/MM optimization, following the same method as described above. The geometry optimization showed an increased chromophore distortion on passing from the ground state to the K intermediate. Moreover, the local hydrogen bond network is perturbed: a water molecule is displaced and the hydrogen bond between Thr89 and Asp85 is broken because the latter group undergoes a conformational change. Another result of that change is the reduction of the distance between the Asp85 carboxyl group and the Schiff base, which is thought to prompt the proton transfer in the following L-to-M process. Regarding the optical properties, K and L intermediates are red-shifted with respect to the initial absorption of bR. The excitation energies were evaluated at the CASSCF(12,11)//HF level, with a three state averaged calculation followed by state specific computation for S_0 and S_1 . The theoretical results are in good agreement with the experimental red shift, even though the absolute values are quite different and the spectroscopic shifts are overestimated. The reason for the shortcomings of the computational theory in this case is the failure of the CASSCF method to account for dynamic correlation. An analysis of the individual contributions to the total excitation energy allows one to pinpoint the main factors affecting the spectroscopic shift. Chromophore distortion around the C15 = N and C13 = C14 double bonds was found to be the main factor. It destabilizes S_0 more than S_1 , therefore reducing the energy gap. Another factor is the reduced electrostatic interaction between protein and chromophore, which increases the S_0 energy without affecting S_1 , where the charge is more delocalized towards the β -ionone ring. Nevertheless the geometry is optimized at the HF level, which is known to overestimate bonding character, compared with post-SCF methods, which leads to a likely underestimation of bond torsion in the chromophore.

Deactivation Mechanism in Cytosine-Guanine DNA Base Pair

Deoxyribonucleic acid (DNA) is a common biomolecule in all living organisms that contains the genetic information used for their development and functioning. It is composed of nucleotides, with backbones made of sugars and phosphate groups joined by ester bonds, forming long polymer chains. Attached to each sugar is one of only four bases: adenine (A), cytosine (C), guanine (G), and thymine (T). These bases belong to two types of organic molecules – adenine and guanine are purines that are fused five- and six-membered heterocyclic compounds, while cytosine and thymine are pyrimidines, which are six-membered rings. A pair of strands is usually arranged in the shape of a double helix by strong hydrogen bonds between the bases that stack upon each other. Due to the aromatic character of the nucleotide bases DNA absorbs in the harmful ultraviolet (UV) region of the spectrum. Hence, UV irradiation is one sort of mutagen that can damage DNA and subsequently lead to cancer. It is not surprising that DNA and its building blocks have been extensively investigated by computational quantum chemistry. The full coverage of the applications goes beyond the scope of this book chapter. Instead we are reporting an investigation using multiconfigurational *ab initio* calculations combined with AMBER force field in a QM/MM approach to study ultrafast radiationless deactivation mechanism of cytosine-guanine inter-strand base pair (Groenhof et al. 2007).

Groenhof et al. (2007) used a multiconfigurational approach as part of the QM/MM setup to study the inter-strand excited state proton transfer. One of the fastest deactivation processes in DNA was related to the cytosine-guanine base pair occurring on sub-picosecond time-scale. Therefore excited state molecular dynamics were employed to track this reaction. CASSCF with a reduced active space of eight electrons in eight orbitals and a 3-21G basis set was employed to model the molecular dynamics of the photoactivated C–G base pair. A surface hopping algorithm was used to detect a transition from the excited to the ground state. The crystal structure of the human DNA/topoisomerase I complex provided the initial coordinates for the QM/MM simulations. In total, 22 base pairs of B-DNA molecule were used. The partitioning of the system was done in a way that the cytosine-guanine base pair in the center of the molecule was described at the QM level. The remainder of the system was modeled with AMBER force field (Case et al. 2002). The chemical bonds between the bases and the deoxyribose sugar rings connecting the QM and the MM subsystems were replaced by constraints and the QM part was capped with two hydrogen link atoms. To equilibrate the DNA and the solvent prior to the QM/MM simulations, the system was equilibrated classically for 1,000 ps. The initial conditions for the excited-state simulations were obtained by taking 20 frames at equal time intervals from an additional 2 ps ground state trajectory at the CASSCF(8,8) level. In the excited state MD simulations, a time step of 0.5 fs was used.

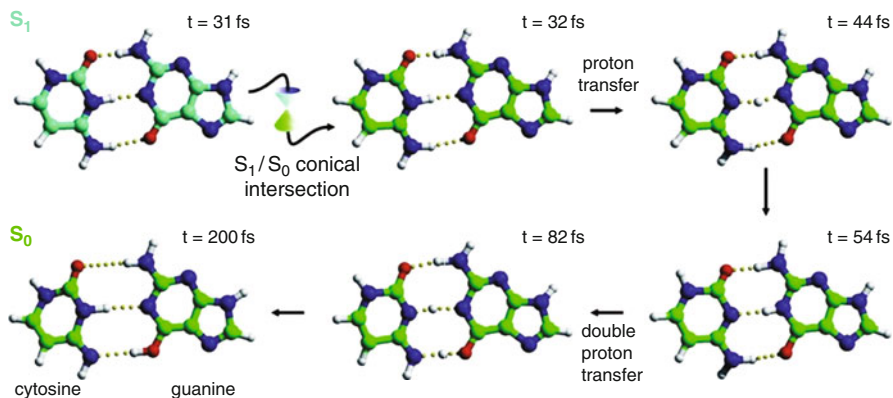


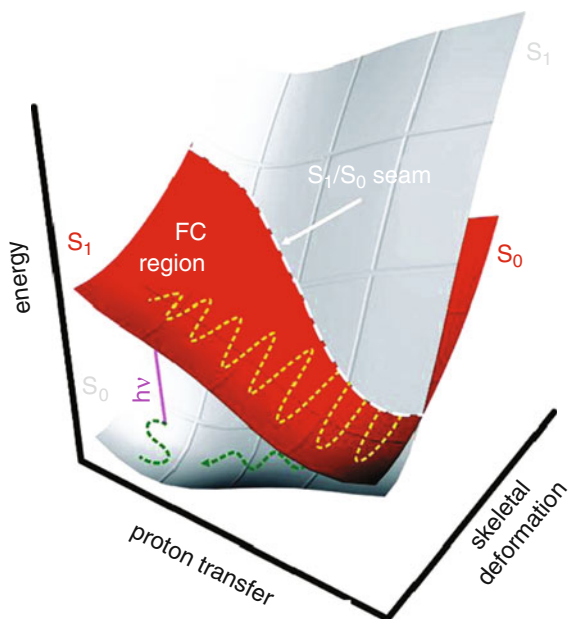
Fig. 23 Sequence of events after the excitation. A proton is abstracted from guanine and accepted by cytosine. After the decay to the ground state via a conical intersection the proton returns to the guanine base. The excess thermal energy that is released upon returning to S_0 is responsible for spontaneous double proton transfer (at 82 fs) which leads to the formation of a different tautomeric state of the base pair (200 fs) (Redrawn with permission from ref. Groenhof et al. 2007)

In consequence, the ultrafast photodeactivation mechanism of the cytosine-guanine base pair was uncovered by the molecular dynamics simulations (see Fig. 23). The excitation to the charge transfer state induces a transfer of a proton from guanine to cytosine within a few femtoseconds. After the proton transfer, the system approaches the conical intersection seam and returns to the ground state. However, within a few femtoseconds the seam is closed again, and a second hop takes the system back to S_1 where it stays until another hop occurs. Most trajectories showed several hopping events between the S_1 and S_0 surfaces, with an average excited state lifetime of 200 fs for the entire process. This can be rationalized by the topology of the $S_1 - S_0$ intersection space (see Fig. 24). The proton transfer coordinate in the vicinity of the S_1 minimum is parallel to the extended hyperline and allows multiple transitions between the crossing states.

Absorption Spectra of a Coumarin in Solution

The calculation of absorption spectra is of great interest to chemists, since these spectra provide a wealth of information about the molecule, its environment and its properties. As a result of solvent effects and solute-solvent interactions the spectra can be quite complex. In addition, some of the information can be convoluted under the broad bands. Nowadays, computational photochemistry serves as a standard tool for spectroscopists, used to assign the experimental spectra. Until recently the calculations yield simple stick absorption spectra, which show the excitation

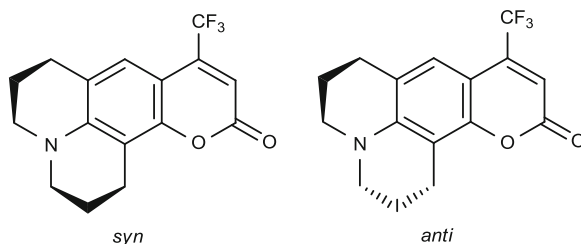
Fig. 24 Excited and ground state potential energy surface crossing of the cytosine-guanine base pair. The two reaction coordinates are proton transfer and skeletal deformation of the bonds. The dashed yellow and green lines schematically represent a path sampled in a typical trajectory on the S_1 and the S_0 potential, respectively. Along the proton-transfer coordinate the system moves out of Franck–Condon region to the minimum of the excited state. Due to the oscillations of the second reaction coordinate the trajectory hits the seam more than once (Redrawn with permission from ref. Groenhof et al. 2007)



energies and their relative intensities, without taking environmental effects into account. Recent advances in quantum chemistry allow researchers to treat solvents and to calculate vibrational contributions to the spectra. Here we present a case study of coumarin C153 (Scheme 9) computed by TD-DFT including solvent effects considered by polarized continuum model (PCM). Improta, Barone, and Santoro have studied the spectra in two different solvents: cyclohexane and dimethylsulfoxide (DMSO) (Improta et al. 2007). Coumarin C153 was selected to evaluate this novel methodology because it exhibits significantly different polarizabilities on its ground and electronically excited states. Therefore solvents with different polarities lead to large solvatochromic shifts that ought to be accounted for by the simulations.

In DMSO the spectrum consisting of an asymmetric broad band, extended towards the blue wing of the maximum, is correctly reproduced by the simulation (Fig. 25 right). This proves the reliability of the calculated vibrational progression hidden within the band. However, similar to the case of cyclohexane, the spectrum is found to be red-shifted by 400 cm^{-1} . The authors ascribe the latter to the effect of dynamic solvent fluctuations which are much more pronounced in polar solvents. Finally, the DMSO \rightarrow cyclohexane solvatochromic shift is estimated to be only 150 cm^{-1} smaller than its experimental counterpart.

Using this novel protocol Improta, Barone, and Santoro were able to simulate the spectrum of coumarin C153 in different solution environments with a high degree of accuracy. Complex band structures were reproduced for solvents of varying polarities with a computational procedure feasible for large molecular systems.



Scheme 9 The calculated stick absorption spectrum of the anti isomer of coumarin at 0 K in cyclohexane is shown in Fig. 25 left. The authors employ an algorithm for automatic selection of transitions between vibrational states since the full set is too computationally demanding. After harmonic analysis the algorithm was used to select $27 \cdot 10^6$ transitions out of a total of 10^{36} states. For the convolution of the spectrum the full width at half maximum of a Gaussian was chosen to match the main experimental bands. The spectrum was found to be unaltered when calculated for 300 K and also to be practically indistinguishable for the anti and syn isomers. The line shapes of the calculated spectrum resemble its experimental counterpart as shown for anti coumarin in cyclohexane (Fig. 25 middle). The two-peak structure with a spacing of $1,400 \text{ cm}^{-1}$ is qualitatively reproduced. The first band is mainly due to the 0–0 transition and a contribution from transitions to the single vibrationally excited state (C1). However, the computed spectrum is found to be blue-shifted by 400 cm^{-1} . The second band is composed of contributions from transitions to the first (C1), second (C2), and third (C3) vibrational states. The only deviation from the experimental spectrum is manifested in the relative height of these two bands

Excited State Molecular Dynamics

Tracking the Photoisomerization of Retinal in Different Environments

The first attempts to study the photoisomerization of the retinal chromophore by using ab initio excited state molecular dynamics were reported by Robb and coworkers (Vreven et al. 1997; Weingart et al. 2004) focusing on a model chromophore in the gas phase at the CASSCF level of theory. A more recent work by Weingart, again for an isolated retinal model, has shown that the quantum yield of the photoisomerization process depends critically on the initial retinal configuration (Weingart 2007). In the resting state of bovine rhodopsin, the chromophore deviates from planarity (Okada et al. 2004; Palczewski et al. 2000). By taking this retinal structure as a starting point for a series of excited state CASSCF semi-classical trajectories the author found that 70 % of the trajectories ended at the all-trans photoproduct. This result suggests that selectivity and quantum efficiency might originate from strain imposed by steric interactions between the chromophore and the protein before photon absorption.

Recently we have reported the first S_1 trajectory computation for Rh carried out with a scaled CASSCF force field that reproduces the static and transient spectroscopic parameters (Frutos et al. 2007). Because of the immense cost of the CASPT2 gradients, their application in trajectory computation is impossible

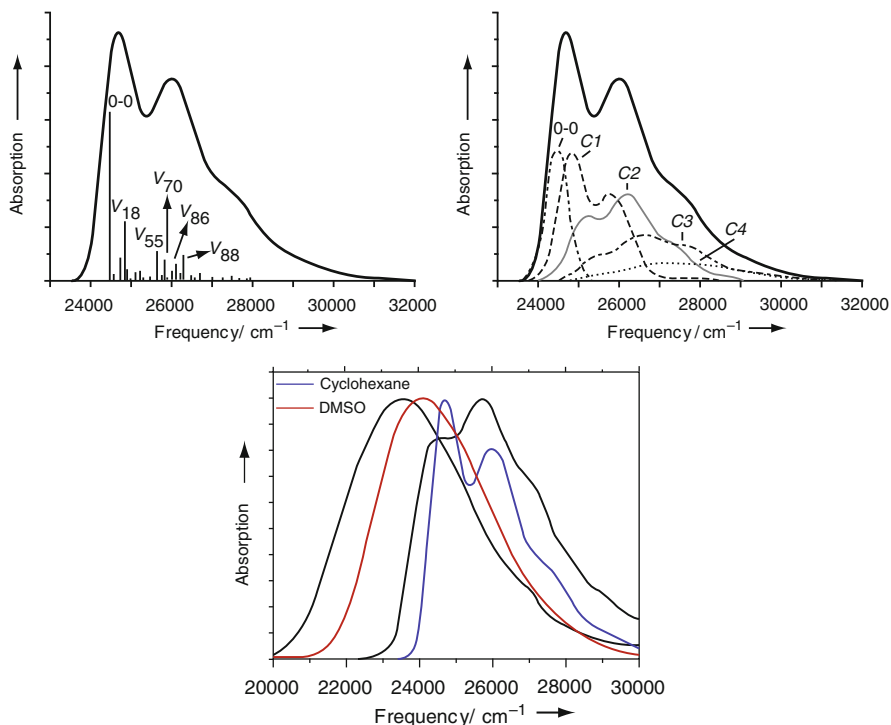


Fig. 25 Calculated spectra of C153 in cyclohexane. Left: 0 K stick spectrum and its assignment including transitions with one quantum in a single oscillator. Middle: A decomposition of the spectrum (thick line) in its components, where C_n dominates collection of transitions to vibrational states with the same number n of simultaneously excited oscillators. Right: Calculated spectra in cyclohexane (blue) and DMSO (red) compared to the experimental spectra (black) (Redrawn with permission from ref. Improta et al. 2007)

for the retinal chromophore even if such a calculation is limited to a few hundred femtoseconds. However, CASSCF energies not only fail to reproduce the observed spectroscopic properties of Rh but also yield an energy profile that is too steep. In order to overcome this problem and evaluate a realistic trajectory we noticed that scaling the CASSCF isomerization energy profile along the S_1 branch of the reaction path of Fig. 22 top left, yields a curve overlapping with the corresponding CASPT2 curve (Frutos et al. 2007). As displayed in Fig. 22 bottom left, the results show that this potential drives the chromophore to the CI on a 120 fs time scale. This time scale is in line with the observed excited state Rh lifetimes, which confirms the suitability of the scaled force field.

The assertion that a scaled CASSCF potential can approximate CASPT2 accuracy has been assessed by single point CASPT2//CASSCF/AMBER calculations along the trajectory. In Fig. 22 bottom-left we report the S_1 potential energy along the computed trajectory. Seven single point CASPT2/6-31G*/AMBER

computations have been performed to validate the scaled energy profile and to compute the correct $S_1 - S_0$ energy gap. The scaled CASSCF energies remain close to the CASPT2 energies all along the trajectory supporting the accuracy of our procedure.

Within 10 fs the S_1 system undergoes a $\sim 8 \text{ kcal} \cdot \text{mol}^{-1}$ energy decrease. Following this event the potential energy decreases slowly and monotonically until, after 110 fs, a region of degeneracy located $\sim 15 \text{ kcal} \cdot \text{mol}^{-1}$ below the Franck-Condon point is reached. Since the initial motion is dominated by simultaneous double bond expansion and single bond contraction, $\sim 8 \text{ kcal} \cdot \text{mol}^{-1}$ of vibrational energy must be initially located along this mode, which leads to complete inversion of the double bond/single bond character centered in the $-C9 = C10-C11 = C12-C13 = C14-$ moiety. The analysis of the entire 140 fs S_1 trajectory together with the resulting backbone deformation depicted in Fig. 22, points to a space saving isomerization mechanism that includes the previously proposed bicycle pedal or crankshaft coordinate. The $-C7 = C8-C9 = C10-C11 =$ segment of PSB11 twists with respect to the two remaining fragments, namely the $-NH = C15-C14 = C13-C12 =$ and the β -ionone ring, during the first 60–70 fs. Such a motion is mainly characterized by a negative twist of the reactive $-C11 = C12-$ bond and a positive twist of the $-C6-C7-$ bond. At the critical 60 fs mark the nature of the motion changes. The $-C6-C7-$ twist stops and the $-C9 = C10$ bond adjacent to the reactive $-C11 = C12-$ bond starts to twist in the positive direction. In other words, the $= C10-C11 =$ fragment rotates with respect to the backbone, leading to an 80° twisted $-C11 = C12-$ bond and to a moderately twisted ($\sim 30^\circ$) $-C9 = C10-$ bond after 110 fs. The nature of this motion is confirmed by plotting the linear momentum vectors in the 65–110 fs region. Further studies have revealed that this motion may be an intrinsic property of the retinal chromophore (Schapiro et al. 2009).

A more recent CASSCF/AMBER trajectory study of the photoisomerization of Rh was reported by Hayashi et al. (2009). The protein model was constructed based on the 1U19 crystal structure. It was equilibrated using classical MD and provided, after a 2 ns equilibration, a sample of 14 different starting points for the trajectory calculations. The polyene chain of the retinal, including the double bond of the β -ionone ring, was treated at the CASSCF level using a DZV basis set. The active space was composed of 12 π -electrons and 12 π -type orbitals. The first two roots were averaged in the CASSCF wave function. In order to describe the interaction between the QM part and the MM part, mechanical embedding was employed. This is an important difference from the excited state trajectory presented above (Frutos et al. 2007) that uses electrostatic embedding instead. In mechanical embedding the interaction between the QM and the MM region is treated at the MM level, which is less accurate. However, the transition from the excited to the ground state was initiated using the energy difference as criterion, namely when it was less than $75 \text{ kcal} \cdot \text{mol}^{-1}$. In addition, a trajectory of isolated retinal was calculated using a structure taken from the equilibrated rhodopsin.

It was found that all 14 CASSCF/AMBER trajectories decayed to the ground state within 100 fs (the shorter time could be attributed to a lack of scaling and

the unrealistically steep CASSCF energy surface), which is of the same order of magnitude as the trajectory by Frutos et al. (2007). Batho-Rh is formed in 13 cases and in one case a 9- cis isomer (isorhodopsin) is produced by a bicycle pedal isomerization of C9–C10 and C11–C12 bond. However, an accompanying rotation of 60°, on average, is found around C9–C10 in the 13 trajectories showing a crankshaft or aborted bicycle pedal-type mechanism, which is consistent with the previous findings.

The first QM/MM trajectory computation describing the photoisomerization of bR was reported by Hayashi et al. (2003). However, the selected QM part corresponded to a highly reduced PSBAT chromophore featuring only three double bonds. This was described by a three state averaged CASSCF wave function with 6 π -electrons in a 6 π -type orbitals active space with a DZV basis set. The MM part was described using AMBER for the protein and the TIP3P model for the water molecules. An ensemble of starting conditions was generated by a classical MD simulation of the resting state of bR at constant temperature (300 K). To compare the effect of the protein on the photoisomerization, trajectories of the same retinal model were calculated under vacuum. In total, 11 trajectories were performed to see if the high selectivity of the isomerization could be reproduced in the absence of the protein environment. It was found that the initial stretching relaxation is followed by a selective isomerization exclusively around the C13–C14 bond. It should be mentioned that this double bond is the central one in the investigated model. Hence, its preference for rotation compared with the terminal bonds is expected. The analysis of the isomerization dynamics of the six isolated chromophore trajectories shows that four trajectories isomerize around C13–C14 and two around the terminal C15–N16 bond. The authors concluded that the high selectivity in bR is a result of the protein environment.

Conclusion

In this chapter we have presented a few selected computational photochemistry case studies of model systems with increasing complexity, starting from isolated molecules that served as models in early works. The photochemical reaction paths of these systems were examined in mechanistic investigations, resulting in a general understanding of light-induced processes. Towards the end of the chapter we have given more recent examples of steady state spectra, minimum energy path calculations, and molecular dynamics simulations in different environments. More realistic models and rigorous simulation not only provided an explanation for the photochemical events, but also allowed a comparison with experimental data: the shape of spectra in solution can be convoluted, ab initio dynamics can predict excited state lifetimes and quantum yields, and conical intersections of chromophores in proteins can be located. All this reveals that the current stage of the tools of computational photochemistry are already providing results with accuracy comparable to experimental observables.

Nevertheless, we have also shown that some questions remain open, needing more accurate treatments which are still not feasible to date. Therefore further breakthroughs in computational photochemistry are anticipated. On one hand the development of quantum chemical methods and algorithms for effective computations is expected in the near future, whereas an increase of computational resources and computing power on the other hand ought to lead to another bloom in this research field.

Bibliography

- Adams, T. R., Adamson, R. D., Austin, B., Baker, J., Bell, A. T., Beran, G. J. O., Baer, R., Besley, N. A., Brooks, B., Brown, S. T., Byrd, E. C., Casanova, D., Challacombe, M., Chai, J., Chakraborty, A. K., Cheng, C., Chien, S., Chipman, D. M., Cramer, C., Crittenden, D., Dachsels, H., Jr., R. A. D., Doerksen, R. J., Dombroski, J. P., Dreuw, A., Dunietz, B. D., Dutoi, A. D., Florian, J., Furlani, T. R., Epifanovsky, E., Gan, Z., Ghysels, A., Gilbert, A. T. B., Gill, P. M. W., Head-Gordon, M., Gwaltney, S. R., Hawkins, G., Hehre, W. J., Herbert, J. M., Heyden, A., Hohenstein, E. G., Hirata, S., Hsu, C., Ishikawa, N., Johnson, B. G., Jung, Y., Kedziora, G., Keil, F. J., Kelly, C., Khaliullin, R. Z., Klunzinger, P., Korambath, P. P., Krylov, A. I., Kurlancheek, W., Kong, J., Kussmann, J., Lee, A. M., Lee, M., Levchenko, S. V., Liang, W., Lin, C. Y., Liotard, D., Livshits, E., Lochan, R. C., Lotan, I., Manohar, P., Marenich, A., Maslen, P. E., Maurice, D. R., Fusti-Molnar, L., Nair, N., Oana, M., O'Neill, D. P., Neuhauser, D., Ochsenfeld, C., Olson, R., Oumi, M., Peters, B., Pieniazek, P. A., Proynov, E. I., Rassolov, V. A., Rhee, Y., Ritchie, J. P., Rohrdanz, M. A., Rosta, E., Schaefer, H. F., III, Shao, Y., Sherrill, C. D., Schultz, N. E., Simmonett, A. C., Slipchenko, L. V., Sodt, A., Steele, R. P., Subotnik, J. E., Thanthiriwatt, K., Thom, A. J. W., Truhlar, D. G., Vanovschi, V., Voorhis, T. V., Wang, T., Warshel, A., White, C., Williams, C. F., Lee Woodcock, H., Wu, Q., You, Z., Zhang, W., & Zhao, Y. (2009). Q-Chem (3.2 ed.). Pittsburgh: Q-Chem, Inc.
- Altoe, P., Stenta, M., & Garavelli, M. (2007). Rhodopsin and GFP chromophores: QM/MM absorption spectra in solvent and protein. *AIP Conference Proceedings*, 963(2), 689–692.
- Andersson, K., Malmqvist, P. A., Roos, B. O., Sadlej, A. J., & Wolinski, K. (1990). 2nd Order perturbation-theory with a CASSCF reference function. *Journal of Physical Chemistry*, 94(14), 5483–5488.
- Andersson, K., Aquilante, F., Barysz, M., Bednarz, E., Bernhardsson, A., Blomberg, M. R. A., Carissan, Y., Cooper, D. L., Cossi, M., Devarajan, A., Vico, L. D., Ferré, N., Fülischer, M. P., Gaenko, A., Gagliardi, L., Ghigo, G., de Graaf, C., Hess, B. A., Hagberg, D., Holt, A., Karlström, G., Krogh, J. W., Lindh, R., Malmqvist, P.-Å., Nakajima, T., Neogrády, P., Olsen, J., Pedersen, T. B., Raab, J., Reiher, M., Roos, B. O., Ryde, U., Schimmelpfennig, B., Schütz, M., Seijo, L., Serrano-Andrés, L., Siegbahn, P. E. M., Stålring, J., Thorsteinsson, T., Veryazov, V., Widmark, P.-O., & Wolf, A. (2009) MOLCAS (Version 7.4. ed.). Lund: University of Lund.
- Andruniow, T., & Olivucci, M. PSBAT absorption spectra in vacuo [Unpublished results].
- Andruniow, T., & Pawlikowski, M. (1998). On the Franck-Condon effects in the absorption spectrum of C10H8 anion. The analysis based on the ab initio MCSCF method. *Acta Physica Polonica A*, 93, 707–715.
- Anglada, J. M., & Bofill, J. M. (1997). A reduced-restricted-quasi-Newton–Raphson method for locating and optimizing energy crossing points between two potential energy surfaces. *Journal of Computational Chemistry*, 18(8), 992–1003.
- Aquilante, F., De Vico, L., Ferré, N., Ghigo, G., Malmqvist, P.-Å., Neogrády, P., Pedersen, T. B., Pitoák, M., Reiher, M., Roos, B. O., Serrano-Andrés, L., Urban, M., Veryazov, V., & Lindh, R. (2009). MOLCAS 7: The next generation. *Journal of Computational Chemistry*, 31(1), 224–247. doi:10.1002/jcc.21318.

- Balashov, S. P., & Ebrey, T. G. (2001). Trapping and spectroscopic identification of the photointermediates of bacteriorhodopsin at low temperatures. *Photochemistry and Photobiology*, *73*(5), 453–462.
- Barbatti, M., & Lischka, H. (2008). Nonadiabatic deactivation of 9H-adenine: A comprehensive picture based on mixed quantum-classical dynamics. *Journal of the American Chemical Society*, *130*(21), 6831–6839. doi:10.1021/ja800589p.
- Barbatti, M., Ruckebauer, M., Szymczak, J. J., Aquino, A. J. A., & Lischka, H. (2008). Nonadiabatic excited-state dynamics of polar pi-systems and related model compounds of biological relevance. *Physical Chemistry Chemical Physics*, *10*(4), 482–494. doi:10.1039/b709315m.
- Bearpark, M. J., Robb, M. A., & Schlegel, H. B. (1994). A direct method for the location of the lowest energy point on a potential surface crossing. *Chemical Physics Letters*, *223*(3), 269–274.
- Belrhali, H., Nollert, P., Royant, A., Menzel, C., Rosenbusch, J. P., Landau, E. M., & Pebay-Peyroula, E. (1999). Protein, lipid and water organization in bacteriorhodopsin crystals: A molecular view of the purple membrane at 1.9 Å resolution. *Structure (London, England: 1993)*, *7*(8), 909–917.
- Ben-Nun, M., & Martinez, T. J. (1998). Nonadiabatic molecular dynamics: Validation of the multiple spawning method for a multidimensional problem. *Journal of Chemical Physics*, *108*(17), 7244–7257.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., & Bourne, P. E. (2000). The protein data bank. *Nucleic Acids Research*, *28*(1), 235–242.
- Bernardi, F., De, S., Olivucci, M., & Robb, M. A. (1990). The mechanism of ground-state-forbidden photochemical pericyclic reactions: Evidence for real conical intersections. *Journal of the American Chemical Society*, *112*(5), 1737–1744.
- Bernardi, F., Olivucci, M., Michl, J., & Robb, M. A. (1996a). Conical intersections in the theory of organic singlet photochemistry. *Spectrum*, *9*, 1–6.
- Bernardi, F., Olivucci, M., & Robb, M. A. (1996b). Potential energy surface crossings in organic photochemistry. *Chemical Society Reviews*, *25*(5), 321–328.
- Bernardi, F., Olivucci, M., & Schegel, H. B. (2006). Michael Alfred Robb: A short biography. *Molecular Physics*, *104*(5), 655–659. doi:10.1080/00268970600595129.
- Blomgren, F., & Larsson, S. (2005). Exploring the potential energy surface of retinal, a comparison of the performance of different methods. *Journal of Computational Chemistry*, *26*(7), 738–742.
- Braslavsky, S. E. (2007). Glossary of terms used in photochemistry, 3rd edition. *Pure and Applied Chemistry*, *79*, 293–465.
- Bravaya, K., Bochenkova, A., Granovsky, A., & Nemukhin, A. (2007). An opsin shift in rhodopsin: Retinal S0-S1 excitation in protein, in solution, and in the gas phase. *Journal of the American Chemical Society*, *129*(43), 13035–13042. doi:10.1021/ja0732126.
- Brejč, K., Sixma, T. K., Kitts, P. A., Kain, S. R., Tsien, R. Y., Ormö, M., & Remington, S. J. (1997). Structural basis for dual excitation and photoisomerization of the *Aequorea victoria* green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(6), 2306–2311.
- Brooks, R., Brucoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, *4*, 187–217.
- Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Wang, J., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Tsui, V., Gohlke, H., Radmer, R. J., Duan, Y., Pitera, J., Massova, I., Seibel, G. L., Singh, U. C., Weiner, P. K., & Kollman, P. A. (2002). *Amber* (7th ed.). San Francisco: University of California.
- Cembran, A., Bernardi, F., Olivucci, M., & Garavelli, M. (2005). The retinal chromophore/chloride ion pair: Structure of the photoisomerization path and interplay of charge transfer and covalent states. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(18), 6255–6260. doi:10.1073/pnas.0408723102.

- Chatteraj, M., King, B. A., Bublitz, G. U., & Boxer, S. G. (1996). Ultra-fast excited state dynamics in green fluorescent protein: Multiple states and proton transfer. *Proceedings of the National Academy of Sciences of the United States of America*, 93(16), 8362–8367.
- De Vico, L., Olivucci, M., & Lindh, R. (2005). New general tools for constrained geometry optimizations. *Journal of Chemical Theory and Computation*, 1(5), 1029–1037. doi:10.1021/ct0500949.
- Dreuw, A., & Head-Gordon, M. (2005). Single-reference ab initio methods for the calculation of excited states of large molecules. *Chemical Reviews*, 105(11), 4009–4037.
- Fantacci, S., Migani, A., & Olivucci, M. (2004). CASPT2//CASSCF and TDDFT//CASSCF mapping of the excited state isomerization path of a minimal model of the retinal chromophore. *The Journal of Physical Chemistry B*, 108(7), 1208–1213. doi:10.1021/jp0362335.
- Farazdel, A., & Dupuis, M. (1991). On the determination of the minimum on the crossing seam of two potential energy surfaces. *Journal of Computational Chemistry*, 12(2), 276–282.
- Ferré, N., & Olivucci, M. (2003a). The amide bond: Pitfalls and drawbacks of the link atom scheme. *THEOCHEM Journal of Molecular Structure*, 632, 71–82. doi:10.1016/s0166-1280(03)00289-6.
- Ferré, N., & Olivucci, M. (2003b). Probing the rhodopsin cavity with reduced retinal models at the CASPT2//CASSCF/AMBER level of theory. *Journal of the American Chemical Society*, 125(23), 6868–6869. doi:10.1021/ja035087d.
- Ferré, N., Cembran, A., Garavelli, M., & Olivucci, M. (2004). Complete-active-space self-consistent-field/Amber parameterization of the Lys296-retinal-Glu113 rhodopsin chromophore-counterion system. *Theoretical Chemistry Accounts*, 112(4), 335–341. doi:10.1007/s00214-004-0593-0.
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, J. M., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, O., Foresman, J. B., Ortiz, J. V., Cioslowski, J., & Fox, D. J. (2009). *GAUSSIAN 09 (Vol. Revision A.1)*. Wallingford CT: Gaussian Inc.
- Frutos, L. M., Andruniów, T., Santoro, F., Ferré, N., & Olivucci, M. (2007). Tracking the excited-state time evolution of the visual pigment with multiconfigurational quantum chemistry. *Proceedings of the National Academy of Sciences of the United States of America*, 104(19), 7764–7769. doi:10.1073/pnas.0701732104.
- Gonzalez-Luque, R., Garavelli, M., Bernardi, F., Merchan, M., Robb, M. A., & Olivucci, M. (2000). Computational evidence in favor of a two-state, two-mode model of the retinal chromophore photoisomerization. *Proceedings of the National Academy of Sciences of the United States of America*, 97(17), 9379–9384.
- Gordon, M. S., Schmidt, M. W., Frenking, G., Kim, K. S., & Scuseria, G. E. (2005). Advances in electronic structure theory: GAMESS a decade later. In C. Dykstra, G. Frenking, K. S. Kim, & G. E. Scuseria (Eds.), *Theory and applications of computational chemistry: The first forty years* (pp. 1167–1189). Amsterdam: Elsevier.
- Grabowski, Z. R., Rotkiewicz, K., & Rettig, W. (2003). Structural changes accompanying intramolecular electron transfer: Focus on twisted intramolecular charge-transfer states and structures. *Chemical Reviews*, 103(10), 3899–4032. doi:10.1021/cr940745l.
- Grigorenko, B., Savitsky, A., Topol, I., Burt, S., & Nemukhin, A. (2006). Ground-state structures and vertical excitations for the kindling fluorescent protein asFP595. *The Journal of Physical Chemistry B*, 110(37), 18635–18640.

- Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M., & Henderson, R. (1996). Electron-crystallographic refinement of the structure of bacteriorhodopsin. *Journal of Molecular Biology*, 259, 393–241.
- Groenhof, G., Bouxin-Cademartory, M., Hess, B., deVisser, S. P., Berendsen, H. J. C., Olivucci, M., Mark, A. E., & Robb, M. A. (2004). Photoactivation of the photoactive yellow protein: Why photon absorption triggers a trans-to-cis isomerization of the chromophore in the protein. *Journal of the American Chemical Society*, 126(13), 4228–4233.
- Groenhof, G., Schäfer, L. V., Boggio-Pasqua, M., Goette, M., Grubmüller, H., & Robb, M. A. (2007). Ultrafast deactivation of an excited cytosine-guanine base pair in DNA. *Journal of the American Chemical Society*, 129(21), 6812–6819. doi:10.1021/ja069176c.
- Groenhof, G., Schäfer, L. V., Boggio-Pasqua, M., Grubmüller, H., & Robb, M. A. (2008). Arginine52 controls the photoisomerization process in photoactive yellow protein. *Journal of the American Chemical Society*, 130, 3250–3251.
- Hayashi, S., & Ohmine, I. (2000). Proton transfer in bacteriorhodopsin: Structure, excitation, IR spectra, and potential energy surface analyses by an ab initio QM/MM method. *The Journal of Physical Chemistry B*, 104(45), 10678–10691.
- Hayashi, S., Tajkhorshid, E., Pebay-Peyroula, E., Royant, A., Landau, E. M., Navarro, J., & Schulten, K. (2001). Structural determinants of spectral tuning in retinal proteins-bacteriorhodopsin vs sensory rhodopsin II. *The Journal of Physical Chemistry B*, 105(41), 10124–10131.
- Hayashi, S., Tajkhorshid, E., & Schulten, K. (2002). Structural changes during the formation of early intermediates in the bacteriorhodopsin photocycle. *Biophysical Journal*, 83(3), 1281–1297.
- Hayashi, S., Tajkhorshid, E., & Schulten, K. (2003). Molecular dynamics simulation of bacteriorhodopsin's photoisomerization using ab initio forces for the excited chromophore. *Biophysical Journal*, 85(3), 1440–1449.
- Hayashi, S., Tajkhorshid, E., & Schulten, K. (2009). Photochemical reaction dynamics of the primary event of vision studied by means of a hybrid molecular simulation. *Biophysical Journal*, 96(2), 403–416.
- Hoffmann, M., Wanko, M., Strodel, P., König, P. H., Frauenheim, T., Schulten, K., Thiel, W., Tajkhorshid, E., & Elstner, M. (2006). Color tuning in rhodopsins: The mechanism for the spectral shift between bacteriorhodopsin and sensory rhodopsin II. *Journal of the American Chemical Society*, 128(33), 10808–10818.
- Improta, R., Barone, V., & Santoro, F. (2007). Ab initio calculations of absorption spectra of large molecules in solution: Coumarin C15313. *Angewandte Chemie International Edition*, 46(3), 405–408.
- Jorgensen, W. L., & Tirado-Rives, J. (1988). The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6), 1657–1666. doi:10.1021/ja00214a001.
- Kandori, H., Schichida, Y., & Yoshisawa, T. (2001). Photoisomerization in rhodopsin. *Biochemistry (Moscow)*, 66, 1197–1209.
- Koga, N., & Morokuma, K. (1985). Determination of the lowest energy point on the crossing seam between two potential surfaces using the energy gradient. *Chemical Physics Letters*, 119(5), 371–374.
- Kukura, P., McCamant, D. W., Yoon, S., Wandschneider, D. B., & Mathies, R. A. (2005). Structural observation of the primary isomerization in vision with femtosecond-stimulated Raman. *Science*, 310(5750), 1006–1009. doi:10.1126/science.1118379.
- Kutateladze, A. G. (2005). *Computational methods in photochemistry*. Boca Raton: Taylor & Francis.
- Laino, T., & Passerone, D. (2004). Pseudo-dynamics and band optimizations: Shedding light into conical intersection seams. *Chemical Physics Letters*, 389(1–3), 1–6. doi:10.1016/j.cplett.2004.02.110.

- Lanyi, J. K. (2000). Molecular mechanism of ion transport in bacteriorhodopsin: Insights from crystallographic, spectroscopic, kinetic, and mutational studies. *The Journal of Physical Chemistry B*, *104*, 11441–11448.
- Levine, B. G., Ko, C., Quenneville, J., & Martinez, T. J. (2006). Conical intersections and double excitations in time-dependent density functional theory. *Molecular Physics*, *104*(5–7), 1039–1051. doi:10.1080/00268970500417762.
- Lischka, H., Shepard, R., Shavitt, I., Pitzer, R. M., Dallos, M., Müller, T., Szalay, P. G., Brown, F. B., Ahlrichs, R., Böhm, H. J., Chang, A., Comeau, D. C., Gdanitz, R., Dachsel, H., Ehrhardt, C., Ernzerhof, M., Höchtel, P., Irle, S., Kedziora, G., Kovar, T., Parasuk, V., Pepper, M. J. M., Scharf, P., Schiffer, H., Schindler, M., Schüler, M., Seth, M., Stahlberg, E. A., Zhao, J.-G., Yabushita, S., Zhang, Z., Barbatti, M., Matsika, S., Schuurmann, M., Yarkony, D. R., Brozell, S. R., Beck, E. V., & Blaudeau, J.-P. (2006). COLUMBUS, an ab initio electronic structure program. (release 5.9.1 ed.). Vienna: University of Vienna.
- Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P., & Lanyi, J. K. (1999). Structure of bacteriorhodopsin at 1.55 Å resolution. *Journal of Molecular Biology*, *291*, 899–911.
- Martin, M. E., Negri, F., & Olivucci, M. (2004). Origin, nature, and fate of the fluorescent state of the green fluorescent protein chromophore at the CASPT2/CASSCF resolution. *Journal of the American Chemical Society*, *126*(17), 5452–5464.
- Mathies, R. A., & Lugtenburg, J. (2000). The primary photoreaction of rhodopsin. In D. G. de Stavenga, W. J. Grip, & E. N. Pugh (Eds.), *Handbook of biological physics* (Vol. 3, pp. 55–90). Amsterdam: Elsevier Science Press.
- Migani, A., & Olivucci, M. (2004). A computational strategy for organic photochemistry. In W. Domcke, D. R. Yarkony, & H. Köppel (Eds.), *Conical intersections: Electronic structure, dynamics and spectroscopy* (pp. 271–320). Singapore: World Scientific.
- Migani, A., Robb, M. A., & Olivucci, M. (2003). Relationship between photoisomerization path and intersection space in a retinal chromophore model. *Journal of the American Chemical Society*, *125*, 2804–2808.
- Migani, A., Sinicropi, A., Ferré, N., Cembran, A., Garavelli, M., & Olivucci, M. (2004). Structure of the intersection space associated with Z/E photoisomerization of retinal in rhodopsin proteins. *Faraday Discussions*, *127*, 179–191.
- Nielsen, I. B., Lammich, L., & Andersen, L. H. (2006). S-1 and S-2 excited states of gas-phase Schiff-base retinal chromophores. *Physical Review Letters*, *96*(1), 1830418304.
- Oesterhelt, D., & Stoekenius, W. (1971). Rhodopsin-like protein from the purple membrane of halobacterium halobium. *Nature: New Biology*, *233*, 149–152.
- Oesterhelt, D., & Stoekenius, W. (1973). Functions of a new photoreceptor membrane. *Proceedings of the National Academy of Sciences of the United States of America*, *70*, 2853–2857.
- Okada, T., Sugihara, M., Bondar, A.-N., Elstner, M., Entel, P., & Buss, V. (2004). The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *Journal of Molecular Biology*, *342*, 571–583.
- Olivucci, M., & Sinicropi, A. (2005). Computational photochemistry. In M. Olivucci (Ed.), *Computational photochemistry. Theoretical and computational chemistry* (Vol. 16, pp. 1–33). Amsterdam: Elsevier.
- Page, C. S., & Olivucci, M. (2003). Ground and excited state CASPT2 geometry optimizations of small organic molecules. *Journal of Computational Chemistry*, *24*(3), 298–309. doi:10.1002/jcc.10145.
- Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Trong, I. L., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., & Miyano, M. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, *289*, 739.
- Pistolessi, S., Sinicropi, A., Pogni, R., Basosi, R., Ferré, N., & Olivucci, M. (2009). Modeling the fluorescence of protein-embedded tryptophans with ab initio multiconfigurational quantum chemistry: The limiting cases of parvalbumin and monellin. *The Journal of Physical Chemistry B*, *113*(49), 16082–16090.

- Ponder, J. W., & Case, D. A. (2001). Force fields for protein simulations. *Advances in Protein Chemistry*, 66, 27–85.
- Quillin, M. L., Anstrom, D. A., Shu, X. K., O'Leary, S., Kallio, K., Chudakov, D. A., & Remington, S. J. (2005). Kindling fluorescent protein from *Anemonia sulcata*: Dark-state structure at 1.38 angstrom resolution. *Biochemistry*, 44(15), 5774–5787.
- Ragazos, I. N., Robb, M. A., Bernardi, F., & Olivucci, M. (1992). Optimization and characterization of the lowest energy point on a conical intersection using an MC-SCF Lagrangian. *Chemical Physics Letters*, 197(3), 217–223.
- Rajput, J., Rahbek, D. B., Andersen, L. H., Hirshfeld, A., Sheves, M., Altoe, P., Orlandi, G., & Garavelli, M. (2010). Probing and modeling the absorption of retinal protein chromophores in vacuo. *Angewandte Chemie International Edition*, 49(10), 1790–1793. doi:10.1002/anie.200905061.
- Robb, M. A., Bernardi, F., & Olivucci, M. (1995). Conical intersection as a mechanistic feature of organic-photochemistry. *Pure and Applied Chemistry*, 67(5), 783–789.
- Roos, B. O. (1987). The complete active space self-consistent field method and its applications in electronic structure calculations. In K. P. Lawley (Ed.), *Advances in chemical physics; Ab initio methods in quantum chemistry - II* (pp. 399–445). Chincester: Wiley.
- Roos, B. O., & Taylor, P. R. (1980). A complete active space SCF (CASSCF) method using a density-matrix formulated super-CI approach. *Chemical Physics*, 48(2), 157–173.
- Schäfer, L. V., Groenhof, G., Kligen, A. R., Ullmann, G. M., Boggio-Pasqua, M., Robb, M. A., & Grubmueller, H. (2007). Photoswitching of the fluorescent protein asFP595: Mechanism, proton pathways, and absorption spectra. *Angewandte Chemie International Edition*, 46(4), 530–536.
- Schäfer, L. V., Groenhof, G., Boggio-Pasqua, M., Robb, M. A., & Grubmuller, H. (2008). Chromophore protonation state controls photoswitching of the fluoroprotein asFP595. *PLoS Computational Biology*, 4(3), e1000034.
- Schapiro, I., Weingart, O., & Buss, V. (2009). Bicycle-pedal isomerization in a rhodopsin chromophore model. *Journal of the American Chemical Society*, 131(1), 16–17. doi:10.1021/ja805586z.
- Schlegel, H. B., & Robb, M. A. (1982). MC SCF gradient optimization of the $\text{H}_2\text{CO} \rightarrow \text{H}_2 + \text{CO}$ transition structure. *Chemical Physics Letters*, 93, 43–46.
- Schmidt, M. W., & Gordon, M. S. (1998). The construction and interpretation of MCSCF wavefunctions. *Annual Review of Physical Chemistry*, 49, 233–266.
- Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S. J., Windus, T. L., Dupuis, M., & Montgomery, J. A. (1993). General atomic and molecular electronic-structure system. *Journal of Computational Chemistry*, 14(11), 1347–1363.
- Sekharan, S., Weingart, O., & Buss, V. (2006). Ground and excited states of retinal Schiff base chromophores by multiconfigurational perturbation theory. *Biophysical Journal*, 91(1), L7–L9.
- Senn, H. M., & Thiel, W. (2009). QM/MM methods for biomolecular systems. *Angewandte Chemie International Edition*, 48(7), 1198–1229.
- Shao, Y., Molnar, L. F., Jung, Y., Kussmann, J., Ochsenfeld, C., Brown, S. T., Gilbert, A. T. B., Slipchenko, L. V., Levchenko, S. V., O'Neill, D. P., DiStasio, R. A., Lochan, R. C., Wang, T., Beran, G. J. O., Besley, N. A., Herbert, J. M., Lin, C. Y., Van Voorhis, T., Chien, S. H., Sodt, A., Steele, R. P., Rassolov, V. A., Maslen, P. E., Korambath, P. P., Adamson, R. D., Austin, B., Baker, J., Byrd, E. F. C., Dachsels, H., Doerksen, R. J., Dreuw, A., Dunietz, B. D., Dutoi, A. D., Furlani, T. R., Gwaltney, S. R., Heyden, A., Hirata, S., Hsu, C. P., Kedziora, G., Khalliulin, R. Z., Klunzinger, P., Lee, A. M., Lee, M. S., Liang, W., Lotan, I., Nair, N., Peters, B., Proynov, E. I., Pieniazek, P. A., Rhee, Y. M., Ritchie, J., Rosta, E., Sherrill, C. D., Simmonett, A. C., Subotnik, J. E., Woodcock, H. L., Zhang, W., Bell, A. T., Chakraborty, A. K., Chipman, D. M., Keil, F. J., Warshel, A., Hehre, W. J., Schaefer, H. F., Kong, J., Krylov, A. I., Gill, P. M. W., & Head-Gordon, M. (2006). Advances in methods and algorithms in a modern quantum chemistry program package. *Physical Chemistry Chemical Physics*, 8(27), 3172–3191.
- Shimomura, O. (2009). Discovery of green fluorescent protein (GFP) (nobel lecture). *Angewandte Chemie International Edition*, 48(31), 5590–5602. doi:10.1002/anie.200902240.

- Sicilia, F., Blancafort, L., Bearpark, M. J., & Robb, M. A. (2007). Quadratic description of conical intersections: Characterization of critical points on the extended seam. *The Journal of Physical Chemistry, A*, *111*(11), 2182–2192. doi:[10.1021/jp067614w](https://doi.org/10.1021/jp067614w).
- Sicilia, F., Blancafort, L., Bearpark, M. J., & Robb, M. A. (2008). New algorithms for optimizing and linking conical intersection points. *Journal of Chemical Theory and Computation*, *4*(2), 257–266. doi:[10.1021/ct7002435](https://doi.org/10.1021/ct7002435).
- Siegbahn, P. E. M., Almlöf, J., Heiberg, A., & Roos, B. O. (1981). The complete active space SCF (CASSCF) method in a Newton–Raphson formulation with application to the HNO molecule. *Journal of Chemical Physics*, *74*(4), 2384–2396.
- Simons, J. (1983). *Energetic principles of chemical reactions*. Boston: Jones and Bartlett.
- Sinicropi, A. (2002). *The mechanism of light energy wastage and exploitation in In,p*chromophores*. Ph.D., University of Siena, Siena.
- Sinicropi, A., Pischel, U., Basosi, R., Nau, W. M., & Olivucci, M. (2000). Conical intersections in charge-transfer induced quenching. *Angewandte Chemie International Edition*, *39*(39), 4582–4586.
- Sinicropi, A., Pogni, R., Basosi, R., Robb, M. A., Gramlich, G., Nau, W. M., & Olivucci, M. (2001). Fluorescence quenching via sequential hydrogen, electron, and proton transfer in the proximity of a conical intersection. *Angewandte Chemie International Edition*, *40*(22), 4185–4189.
- Sinicropi, A., Nau, W. M., & Olivucci, M. (2002). Excited state quenching via “unsuccessful” chemical reactions. *Photochemical and Photobiological Sciences*, *1*(8), 537–546.
- Sinicropi, A., Page, C. S., Adam, W., & Olivucci, M. (2003). Computational study on the origin of the stereoselectivity for the photochemical denitrogenation of diazabicycloheptene. *Journal of the American Chemical Society*, *125*(36), 10947–10959. doi:[10.1021/ja0263137](https://doi.org/10.1021/ja0263137).
- Sinicropi, A., Andruniow, T., Ferre, N., Basosi, R., & Olivucci, M. (2005). Properties of the emitting state of the green fluorescent protein resolved at the CASPT2//CASSCF/CHARMM level. *Journal of the American Chemical Society*, *127*(33), 11534–11535.
- Solomon, B. S., Thomas, T. F., & Steel, C. (1968). Primary processes in the photochemistry of bicyclic azo compounds. *Journal of the American Chemical Society*, *90*, 2249–2258.
- Strambi, A., Coto, P. B., Ferré, N., & Olivucci, M. (2007). Effects of water re-location and cavity trimming on the CASPT2//CASSCF/AMBER excitation energy of rhodopsin. *Theoretical Chemistry Accounts*, *118*(1), 185–191. doi:[10.1007/s00214-007-0273-y](https://doi.org/10.1007/s00214-007-0273-y).
- Teller, D. C., Okada, T., Behnke, C. A., Palczewski, K., & Stenkamp, R. E. (2001). Advances in determination of a high-resolution three-dimensional structure of rhodopsin, a model of G-protein-coupled receptors (GPCRs). *Biochemistry*, *40*, 7761–7772.
- Turro, N. J., Ramamurthy, V., & Scaiano, J. C. (2009). *Principles of molecular photochemistry: An introduction*. Sausalito: University Science.
- Valsson, O., & Filippi, C. (2010). Photoisomerization of model retinal chromophores: Insight from quantum monte carlo and multiconfigurational perturbation theory. *Journal of Chemical Theory and Computation*, *6*(4), 1275–1292. doi:[10.1021/ct900692y](https://doi.org/10.1021/ct900692y).
- Virshup, A. M., Punwong, C., Pogorelov, T. V., Lindquist, B. A., Ko, C., & Martinez, T. J. (2009). Photodynamics in complex environments: Ab initio multiple spawning quantum mechanical/molecular mechanical dynamics. *The Journal of Physical Chemistry B*, *113*(11), 3280–3291. doi:[10.1021/jp8073464](https://doi.org/10.1021/jp8073464).
- Vreven, T., Bernardi, F., Garavelli, M., Olivucci, M., Robb, M. A., & Schlegel, H. B. (1997). Ab initio photoisomerization dynamics of a simple retinal chromophore model. *Journal of the American Chemical Society*, *119*(51), 12687–12688.
- Wanko, M., Garavelli, M., Bernardi, F., Niehaus, T. A., Frauenheim, T., & Elstner, M. (2004). A global investigation of excited state surfaces within time-dependent density-functional response theory. *Journal of Chemical Physics*, *120*(4), 1674–1692.
- Wanko, M., Hoffmann, M., Strodel, P., Koslowski, A., Thiel, W., Neese, F., Frauenheim, T., & Elstner, M. (2005). Calculating absorption shifts for retinal proteins: Computational challenges. *The Journal of Physical Chemistry B*, *109*(8), 3606–3615. doi:[10.1021/jp0463060](https://doi.org/10.1021/jp0463060).
- Warshel, A. (1976). Bicycle-pedal model for the first step in the vision process. *Nature*, *260*(5553), 679–683.

- Warshel, A., & Barboy, N. (1982). Energy storage and reaction pathways in the first step of the vision process. *Journal of the American Chemical Society*, *104*(6), 1469–1476.
- Warshel, A., & Chu, Z. T. (2001). Nature of the Surface Crossing Process in bacteriorhodopsin: Computer simulations of the quantum dynamics of the primary photochemical event. *The Journal of Physical Chemistry B*, *105*, 9857–9871.
- Warshel, A., & Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology*, *103*(2), 227–249.
- Warshel, A., Chu, Z. T., & Hwang, J.-K. (1991). The dynamics of the primary event in rhodopsins revisited. *Chemical Physics*, *158*, 303–314.
- Weingart, O. (2007). The twisted C11 = C12 bond of the rhodopsin chromophore—a photochemical hot spot. *Journal of the American Chemical Society*, *129*(35), 10618–10619. doi:[10.1021/ja071793t](https://doi.org/10.1021/ja071793t).
- Weingart, O., Migani, A., Olivucci, M., Robb, M. A., Buss, V., & Hunt, P. (2004). Probing the photochemical funnel of a retinal chromophore model via zero-point energy sampling semiclassical dynamics. *The Journal of Physical Chemistry. A*, *108*(21), 4685–4693. doi:[10.1021/jp049140b](https://doi.org/10.1021/jp049140b).
- Werner, H.-J., Knowles, P. J., Lindh, R., Manby, F. R., Schuetz, M., Celani, P., Korona, T., Mitrushenkov, A., Rauhut, G., Adler, T. B., Amos, R. D., Bernhardsson, A., Berning, A., Cooper, D. L., Deegan, M. J. O., Dobbyn, A. J., Eckert, F., Goll, E., Hampel, C., Hetzer, G., Hrenar, T., Knizia, G., Koepl, C., Liu, Y., Lloyd, A. W., Mata, R. A., May, A. J., McNicholas, S. J., Meyer, W., Mura, M. E., Nicklass, A., Palmieri, P., Pflueger, K., Pitzer, R., Reiher, M., Schumann, U., Stoll, H., Stone, A. J., Tarroni, R., Thorsteinsson, T., Wang, M., & Wolf, A. (2008). MOLPRO, a package of ab initio programs (Version 2008.3 ed.). Cardiff: University of Cardiff.
- Yamamoto, N., Olivucci, M., Celani, P., Bernardi, F., & Robb, M. A. (1998). An MC-SCF/MP2 study of the photochemistry of 2,3-diazabicyclo [2.2.1] hept-2-ene: Production and fate of diazenyl and diazabicyclo biradicals. *Journal of the American Chemical Society*, *120*(10), 2391–2407.
- Yarkony, D. R. (1993). Systematic determination of intersections of potential energy surfaces using a Lagrange multiplier constrained procedure. *Journal of Physical Chemistry*, *97*(17), 4407–4412.

Part V

Chemoinformatics

Tomasz Puzyn

Jaroslav Polanski and Johann Gasteiger

Contents

Introduction	1998
History of Chemoinformatics	1999
Scope of Chemoinformatics	2000
From Molecules to Substances and Compounds	2002
Chemical Space	2004
Storing of Chemical Structure Information in the Computer	2007
Mapping Chemical Graphs into 0D Descriptors	2008
Mapping Chemical Graphs into 1D Descriptors	2009
Mapping Chemical Graphs into 2D Descriptors	2010
Mapping Chemical Graphs into 3D Descriptors	2011
Databases and Database Searching	2014
Prediction of Properties	2018
Representations for Coding Molecules	2021
Mapping Chemical Graphs into 0D (Global) Descriptors	2022
Mapping 2D Molecular Graphs into 1D and 2D Descriptors	2023
Representation of 3D Structures and Their Mapping into Lower Dimensionality	2024
Representation of Molecular Surfaces	2026
Handling of Conformational Flexibility	2027
Representation of Chemical Reactions	2028
Trends in Structure Representation	2028
Applications of Molecular Descriptors for Property Prediction	2029
Drug Design	2030
Structure Elucidation	2034
Organic Synthesis Design	2035

J. Polanski (✉)

Organic Chemistry and Chemoinformatics, University of Silesia Institute of Chemistry,
Katowice, Poland

e-mail: jaroslav.polanski@us.edu.pl

J. Gasteiger

Computer-Chemie-Centrum, Universität Erlangen-Nürnberg, Erlangen, Germany

e-mail: johann.gasteiger@fau.de

Chemical Compounds Beyond Structure Representation	2035
Conclusions	2036
Bibliography	2036

Abstract

Molecular descriptors and properties are two basic forms of chemical information representing compounds in the computer. Chemical space is a structure for the mapping of these information types. It is not always easy to distinguish between descriptors and properties which are the rare annotations of chemical compounds and often need to be predicted. However, this distinction helps in understanding structure-property approaches and the possible origins for their failures. A variety of molecular descriptors is available both for storing molecules and for their processing in computers and we do not have here a sharp differentiation. The trend in property prediction goes away from pure modeling and prediction endeavors to the development of models that can also be interpreted and thus help us in increasing chemical insight and knowledge.

Introduction

Chemoinformatics is a term that has been coined to describe a discipline organizing and coordinating the application of informatics methods in chemistry. The definition of chemoinformatics evolved with the changing landscapes of informatics (computer sciences) and chemistry (Gasteiger and Engel 2003). According to Greg Paris, this “encompasses the design, creation, organization, storage, management, retrieval, analysis, dissemination, visualisation and use of chemical information, not only in its own right, but as a surrogate or index for other data, information and knowledge” (Paris 1999). Frank K. Brown describes this discipline as “the combination of all the information resources that a scientist needs to optimize the properties of a ligand to become a drug” (Brown 1998). Latest definitions broadened the scope of the area to the “the application of informatics methods to solve chemical problems” or simply to *in silico chemistry* (Gasteiger and Engel 2003; Polanski 2009).

At the semantic level, two elements constructing the designation *chemoinformatics* clearly relate this discipline to chemistry and informatics. The designation originated in the 1990s, when a new name was sought for the discipline that uses computers in drug design. This reflected the increasing potential of informatics and a large dose of the early optimism of computer applications in the field of drug design. Chemistry which originated this subdiscipline is the science of **chemical compounds and their transformations**. The maze of molecules, chemical entities, compounds, substances, and species (Woody et al. 2012) that was created by nature and chemists was not easy to be translated into machine understandable data. The human chemist is flexible enough to be aware of the problematic definitions, while a single uncertainty is already a problem for the computer. Computer understandable

chemistry is required for the machine to process the data. At the same time, it is also required to enable an interaction between chemist and computer to translate structure data of the molecular objects into a machine-readable and processable system that would be clear enough and unambiguous. Teaching computer chemistry appeared to be an enormously hard exercise.

A variety of textbooks, handbooks, and handbook chapters (Gasteiger and Engel 2003; Gasteiger 2003; Leach and Gillet 2003; Bajorath 2004; Lavine 2005; Bunin et al. 2007; Foulon and Bender 2010; Polanski 2009; Polanski and Bak 2010; Karthikeyan and Vyas 2014) is available in the field. However, we also need to make a few definitions in order to structure the field of chemoinformatics in a more precise manner. Furthermore, we want to describe still unsolved dilemma and discuss problems and concepts within this discipline.

History of Chemoinformatics

Historically, the field of chemoinformatics goes back to the 1960s although the name chemoinformatics was coined only in 1998. An illustrative historical perspective is drawn by the evolution of the journals publishing research in the discipline. Chemical information formed an important component of chemoinformatics. This was one of the reasons why the *Journal of Chemical Documentation* changed its name to *Journal of Chemical Information and Computer Sciences* in 1975. Then, in turn, this journal changed its name to *Journal of Chemical Information and Modeling* in 2005. “Computer science is now too far from the chemical core and chemists believe that they are generating their own tools for computer chemistry investigations” (Gasteiger and Engel 2003; Polanski 2009). The *Journal of Cheminformatics* has just started to appear in 2009. When the name chemoinformatics was introduced, the term *chemometrics* had already been coined. Chemometrics did not refer to all computer applications in the field of chemistry. Instead, it focused on some computer applications in analytical chemistry. This was probably the first field of computer applications in chemistry organized by a common name. Chemometrics started to deal with objectives less complex than those targeted by chemoinformatics.

In the very beginning of computer science, chemists realized its important potential for storing of chemical information, for molecular modeling, and for structure manipulation. At the same time, bringing molecules into the virtual computer world was a complex problem which had never been attempted by chemometrics. Various roots for chemoinformatics can be identified: In the mid-1960s, work was initiated on building databases on chemical information. The problem of a unique and unambiguous representation of chemical structures was addressed by Morgan (1965) which led to the construction of the Chemical Abstracts Service databases. Hansch and Fujita published their seminal work on quantitatively modeling biological activity which led to the establishment of the field of quantitative structure-activity relationship (QSAR) (Hansch and Fujita 1964). In the late 1960s, Langridge and coworkers visualized 3D molecular models

on cathode ray tubes (Lesk 1977). The DENDRAL project was initiated in 1964 at Stanford University for deriving the structure of organic compounds from their mass spectra (Lindsay et al. 1980). This work is widely considered to be the first application of artificial intelligence to organic chemistry. Other approaches to computer-assisted structure elucidation (CASE) were made in the late 1960s by Sasaki in Japan (Sasaki et al. 1968) and Munk in Arizona (Shelley et al. 1978). In the late 1960s and early 1970s, several groups at Harvard, SUNY, Brandeis, and TUM in Germany (Corey and Wipke 1969; Gelernter et al. 1973; Hendrickson 1971; Blair et al. 1974) took up the problem of computer-assisted synthesis design (CASD). Although none of the systems that were developed found general acceptance by organic chemists, the work on these systems eventually created the foundation for in-house databases on chemical structures and reactions such as MACCS, REACCS, and the Beilstein database.

Chemoinformatics, cheminformatics, molecular informatics, and chemical informatics are related names or even synonyms. In fact, *chemometrics* that uses, for example, the same techniques for analyzing data is evidently covered by chemoinformatics. However, for historical reasons, chemometricians often prefer to preserve their separate identity. For sure, however, seeing these disciplines as a unity will help much in better understanding the world of chemistry.

Scope of Chemoinformatics

Attempting to identify the scope of chemoinformatics, we can try to indicate differences in quantum chemistry and chemoinformatics. Both focus on modeling molecular effects. Hard theories with brute-force approaches are available for small objects. This decides that quantum chemistry uses *deductive learning*, providing a theory from which predictions are made. The complexity of the larger chemical and biological bodies decides that hard physical and mathematical approaches fail. Thus, cheminformatics uses *inductive learning*, trying to learn from experimental observations and data. Surprisingly, such methods are highly successful not only here but generally in science, where real progress is often accompanied by serendipity. Learning from experiments and data has been at the core of understanding chemical phenomenon from the very beginning of chemistry as a science. Cheminformatics now provides tools to do that kind of learning in a more systematic and efficient way. Quantum chemistry and cheminformatics may meet somewhere in the future for the better understanding of complex chemical and biological objects.

Accordingly, chemical data are the core issue in cheminformatics which both learning from data and focusing on data analysis for data prediction. Data prediction can be distinguished into two basic groups. First are molecular predictions, e.g., modeling 3D structures. Second, there are predictions of properties for molecular assemblies forming chemical substances. Both types of predictions are of crucial importance for drug design which is a major field of application of cheminformatics. Although drug design is still a central arena where cheminformatics is played, the scope of the discipline has included, from the very beginning, practically

Table 1 Chemoinformatics from problems to concepts

Problems	More important concepts
Chemical data documentation and searches: database searches and management	SMILES coding
Calculating molecular descriptors: 2D (chemical graphs) and 3D (molecular modeling) representations are mapped into single numbers (0D); vectors, fingerprints(1D); matrixes, surface maps (2D); surface, atomic representations, force fields, virtual or real receptor data (3D); etc.	Connectivity (graph theory) approaches Additivity concept Molecular modeling Force fields
Molecular modeling: molecular structure 3D data generation in silico from their 2D or 1D representations. This includes both predictions of molecular topography and molecular descriptors, i.e., atomic annotations data by calculated (simulated) atomic properties	Molecular mechanics Molecular dynamics Force field Molecular interaction field Partial atomic charges Lipophilic potential
Structure elucidation: (synonym: structure-spectra correlations) property to molecular structure is mapped in FCS when we are attempting to find a structure having a certain spectra or in structure to property version (VCS and FCS) we are simulating a spectra for a given molecule (substance)	RDF structure coding (2D) 2D structure – 2D spectra correlation
Mapping structure to activity (SAR): a series of structures (FCS substances) is needed to study SAR, which are usually synthesized. The real goal of SAR is usually to predict the training series for designing new structures and substances by property predictions in a qualitative or quantitative procedures	SAR QSAR QSAR domain Similarity measures Privileged structures Fragonomics
Property prediction: compounds' series (FCS) are mapped from FCS into VCS; two basic versions are available, i.e., property versus property or structure versus property. However, in the design step for novel compounds in VCS (we can design both compounds VCS or FCS where some properties can be registered in databases and/or literature), we are always to use a structure version (no property is available in VCS)	QSAR QPAR logP versus partition coefficient Fragonomics Virtual screening
Synthesis design: designing organic synthesis in product (target molecule) to reagent (synthon, reagent equivalent) mapping	Synthons Retrosynthetic analysis Synthesis tree

all computer applications in any field of chemistry. In Table 1, we are introducing chemoinformatics by basic problems and concepts developed in this discipline.

On the other hand, the trio, quantum chemistry, chemometrics, and chemoinformatics, nicely illustrates how creative is science forming different approaches, depending upon the encountered complexity of the problems, methods available, and reductionism level needed. We tried to conclude this in Table 2.

Table 2 In silico chemistry from quantum chemistry to chemoinformatics

	Quantum chemistry	Chemometrics	Chemoinformatics
Object	Atoms or small molecules	Soft mathematics for chemistry	Large chemical (bio)systems
Complexity	Low	Moderate	High
Precision	High	Moderate	Low
Reductionism	Low	Moderate	High
Problems	Well-defined molecules (atomic physics)	Ill-defined mathematics in chemistry	Ill-defined chemistry
	Hard mathematics available	Soft mathematics available	The core of the problem is beyond mathematics
Methods	Hard science	Soft mathematics	Soft science
	Deductive learning	Scientific guess	Inductive learning
Tools	Mathematics	Mathematics	Mathematics
	Physics	Physics	Chemistry
			Physics
			Chemistry
		Quantum chemistry chemometrics	

From Molecules to Substances and Compounds

Chemistry focuses on chemical compounds. But what do we mean by a chemical compound? The dictionary meaning of chemical compound is a substance formed from two or more elements chemically united in fixed proportions or in other words composed of identical molecules consisting of atoms of two or more chemical elements. Molecules and substances are precisely defined. Substantial definitions of chemistry were formulated by IUPAC. This includes especially:

- *Chemical element*
- *Chemical species*
- *Chemical substance*
- *Measurement*
- *Mixture*
- *Molecular entity*
- *Molecule*
- *Property*

A reader interested in more details should refer to the IUPAC compendium, an online version is available (IUPAC(a), goldbook.iupac.org), to precisely understand these terms. In particular, a molecule is an atomic scale body. Since the definition refers to a single entity, we are dealing in the laboratory with such objects only very rarely. In the real world, molecules agglomerate forming substances. From the

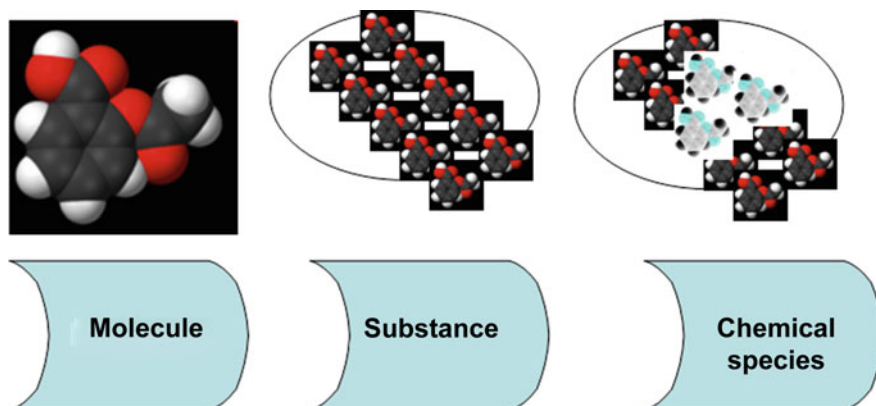


Fig. 1 The increase in information complexity from molecule to chemical compound

viewpoint of information theory, agglomeration brings a substantial complication. This process can significantly influence molecular structure and properties. Further, a single molecule (molecular entity) under experiment can form a substance arranged by an ensemble of chemically identical molecular entities. This effect of differentiation of the initial uniform molecular body is called speciation, and the resulting system is known under the term chemical species. In Fig. 1, we illustrate the increase of the system complexity from molecules to substances. Molecular agglomeration brings the problem of intra- and extramolecular interactions which changes the molecular arrangement and properties. Molecular interactions can include both original molecules and the molecules of different molecular species, and we cannot decide a priori on the importance of these interactions. Furthermore, in many experiments, we are dealing with molecules in solutions where individual molecules are surrounded by solvent molecules. The environment, concentration, pH value, etc., bring further complications influencing substances.

In fact, both chemistry and chemoinformatics deal with chemical compounds. A chemical compound is a historically formed term that covers both molecules and substances including molecular agglomeration to a single and to multiple molecular entities (chemical species). What we mean by chemical compound depends on the context and can denote:

- A *synonym of a molecule* (composed of at least two different atoms) representing a single chemical body for virtual needs (coding, in silico processing, visualization, etc.)
- A *molecule* representing a single chemical body under measurement (e.g., a single DNA strand that can be observed, e.g., under microscopy)
- A *substance* composed of a replicated but single molecular entity under measurement
- A *chemical species* under measurement when replicated is more than one molecule type

In fact, a molecule is a model of a chemical compound available both in the form of a physical representation that shows what this object looks like and as an idealized image of an atomic scale body. Formally, the term molecular descriptor has been coined by Todeschini to describe all numbers relating to molecules and obtained as “a final result of a logic and mathematical procedure transforming chemical information encoded within a symbolic representation of a molecule into a useful number” (Todeschini and Consonni 2000). However, we are still not consequent enough in differentiating molecules from substances. For example, it reads: “Molecular descriptors are divided into two main classes: experimental measurements, such as logP, molar refractivity, dipole moment, polarizability, and, in general, physicochemical properties, and theoretical molecular descriptors” (Consonni and Todeschini 2010). We should distinguish *experimental measurements*, i.e., *properties* (mainly for real substances), from other numbers describing in silico *simulated* molecules or substances, namely, molecular descriptors. By the way, this inconsequence is not a coincidence because historically in chemistry we are determined to think in the category of chemical compounds. To realize the complication of defining molecular bodies forming a single chemical compound, the reader is referred to Bobach et al. (2012).

With the development of chemical knowledge, we realized the need for measuring similarities or, alternatively, the diversity of chemical compounds that helps the systematic probing or screening of objects within the controlled molecular structure span. For example, in QSAR this problem is referred to as defining the model domain (Netzeva et al. 2005).

In sections “[Storing of Chemical Structure Information in the Computer](#)” and “[Representations for Coding Molecules](#),” we will largely concentrate on the representation of molecular structures. In section “[Chemical Compounds Beyond Structure Representation](#),” we will briefly discuss the representation of chemical compounds that cannot be specified with a molecular structure.

Chemical Space

Chemical space (CS) is a concept which appeared in the hope of organizing a whole population of chemical compounds. Chemical space is a structure for the mapping of chemical compounds by descriptors and properties. Despite widespread attention, this term has not been precisely defined. The idea was inspired by the cosmological universe populated by stars, where chemical compounds replace cosmological objects to *navigate chemical space* (Lipinski and Hopkins 2004). This was shaped by multidimensional descriptors (fingerprints) or properties, while the comparison or similarity studies are performed by the comparison with a reference compound (Medina-Franco et al. 2008). The parameters characterizing similarity to a given reference compound can also be annotated in such a space. The distribution of compounds in space and neighborhoods and relations between

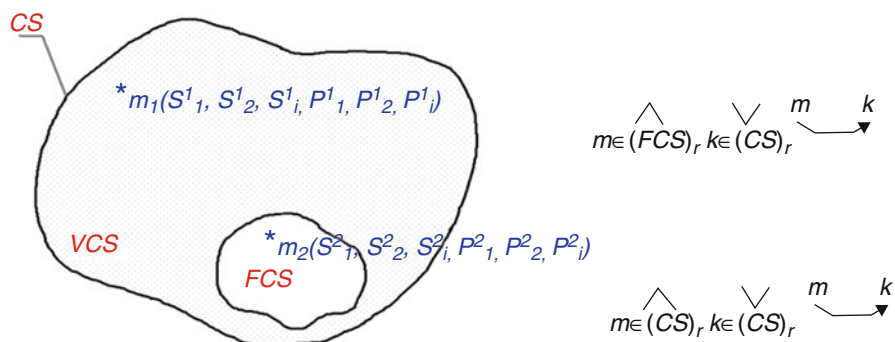


Fig. 2 Chemical space (CS). On the *right-hand side*, two basic problems of in vitro and in silico organic synthesis are defined as the mapping of molecules (chemical compounds) m into products k

the objects populating CS depend upon the molecules selected to populate CS. It was suggested that chemical space can be defined as the entire collection of all meaningful chemical compounds, typically restricted to small organic molecules (Dobson 2004). For a detailed review and discussion of the abovementioned concepts, the reader should compare reference Boehm (2011).

A definition of CS related to mathematics was attempted in Polanski (2009). In Fig. 2, we illustrate this concept. The CS is populated by chemical compounds, i.e., both molecules and substances are covered by this structure. The CS can be divided into two basic moieties, factual chemical space (FCS), described by chemical compounds that have already been obtained and registered, and virtual CS (VCS) mapping hypothetical molecular structures. We can use the mathematic formalism for mapping molecular objects, their properties, or descriptors. Accordingly, we note two common chemical problems, chemical synthesis and reaction predictions, using this structure (Fig. 2). An in vitro reagent should be a known substance of FCS, and its product can be both an FCS and a VCS element; m should be included in the reaction domain r for reagents and k in the reaction domain of products. The difference for in silico operation is that both m and k can be CS elements.

It is interesting to further analyze the structure of chemical space. If we hypothetically visualize CS as a matrix, this will include three main blocks of data. First, these will be the data coding constitution and stereochemistry of the molecules, and then come the molecular descriptors of the molecules, and, finally, the properties measured for the substances formed by the molecules. Since molecules and substances together are chemical compounds, the CS is populated by chemical compounds.

It is worth mentioning here that the differences between VCS and FCS will not be so large as could be expected. By FCS we mean that the compounds were registered in any database or a publication. But even in such cases, we usually have nothing more than melting point and ^1H NMR spectra registered. The FCS

compound collections (libraries) for actual property measurements are still available rather rarely. Thus, the CS matrix resembles a sparse matrix populated by few data even for the compounds that are already registered. On the other hand, *measured* properties are completely unavailable for virtual compounds (VCS). A number of missing values in FCS means that we are working more on virtual molecules with descriptors and predicted properties than with real property values. It means that the data flood in chemistry is a myth. It is a local effect observed for only a rather few chemical compounds and properties. We still do not have substantial information for chemical compounds. Actually, the large number of data results more from the large number of compounds registered (ca. 70 million or 10^8 compounds) than from the completeness of their properties measured. The number of chemical compounds obtained is however nothing if we compare this to the enormous number of potential chemical structures which ranges anywhere from 10^{18} to 10^{180} molecules (Boehm 2011).

Interestingly the libraries of real chemical compounds collected by pharmaceutical companies are probably the largest chemistry deposits available in FCS. It is also just in recent years that such libraries have become available in chemistry.

It is also interesting to enumerate the number of registered chemical compounds in VCS (Boehm 2011):

- Screening libraries of pharma companies 10^6
- Commercially available compounds 10^7
- Patent/literature compounds 10^8

The structure of chemical space helps us in understanding how rare are real property annotations of chemical compounds. Properties decide on substance usability and on its importance in chemistry. They allow compound identifications and comparisons. This seems trivial and chemists were aware of this very early. The FCS was described by *Chemisches Zentralblatt*, which appeared as early as 1830 and by Beilstein's *Handbuch der Organischen Chemie* which contained already in 1881 two volumes, registering 1500 compounds, with more than 2000 pages. Even such small FCS needed a system addressing chemical compounds in the compendium. Chemical formulae, chemical name (nomenclature system), and a special structure of *Systemnummers* helped in finding the compound's entry in the Beilstein Handbook. Despite this, for years, using Beilstein needed expertise in chemistry.

At the very beginning, each new compound obtained was of substantial interest exploring new chemistry areas. However, a relatively large population of 100 million compounds explored by now decides our preferences in our current chemical investigations. Nowadays, we focus not on new compounds but rather on exceptional properties in FCS (Hammond 1968). This decides that we do not care for publication of any property data for any new FCS compound. Currently, published compound properties are stored in databases available online.

Storing of Chemical Structure Information in the Computer

Two essential tasks have to be solved in chemoinformatics: the storing of chemical compound information and the prediction of properties of chemical compounds. Basically, therefore, two types of molecular descriptors can be identified. First are descriptors needed to map the structures into unique and unambiguous codes understandable for the computers. This is important both for experimental chemistry working with substances in FCS and theoretical chemistry dealing with molecules or virtual chemical substances (VCS). This type of mapping is performed especially by the descriptors discussed in sections “[Mapping Chemical Graphs into 0D Descriptors](#),” “[Mapping Chemical Graphs into 1D Descriptors](#),” “[Mapping Chemical Graphs into 2D Descriptors](#),” and “[Mapping Chemical Graphs into 3D Descriptors](#).” Then, we have a variety of molecular descriptors designed to be used for in silico molecular predictions, property prediction, or molecular visualization. In the latter case, the prediction of properties of compounds, it is more important to describe the structure of molecules in a manner appropriate for the prediction of the property of interest than to have a fully reversible mapping between molecular structures and descriptors, i.e., once we transformed the molecule into the descriptors, we often cannot easily and/or unambiguously come back to this specific molecule. We will first discuss in this section the descriptors needed for storing compound information in the computer. The derivation of descriptors for the prediction of properties will be discussed in section “[Representations for Coding Molecules](#).”

In order to control the CS (molecular management and manipulation), we need to have efficient machine-searchable databases registering all compounds that have been synthesized by chemists from the very early days to today. This problem is referred as *structure representation and searching*.

Molecules formed by atoms connected by chemical bonds can formally be represented in a simplified 2D version using graph theory which replaces an atom by a node (vertex) and a bond by an edge. The visualization of molecules by chemical graphs annotated with atomic symbols was such a successful *chemist-friendly* concept that this seems today to be a molecule archetype deciding the perception of chemistry and developing the essence of chemical knowledge. In other words, today we simply cannot imagine understanding the basic concepts of chemistry without 2D chemical graphs. In this sense, chemistry is in the lucky and highly valuable position of having a language that is internationally accepted and understood. The task is then to translate this language into a computer-readable and understandable form. The formal features of graph theory important for mapping molecules into chemical or molecular graphs will not be discussed here further, and the reader is referred to a number of available literature ranging from a very formal to more popular descriptions. A comprehensive overview is provided by Gasteiger (2003).

The storing of information on the structure of chemical compounds has to be unique, i.e., the stored information should for a given compound always be the same, and it has to be unambiguous, i.e., not two different compounds should

receive the same representation. To uniquely and unambiguously code molecules, we need to define their constitution and stereochemistry, where constitution means the connectivity pattern between atoms and stereochemistry defines the spatial arrangement of the atoms in a reduced form.

It is worth mentioning, however, that in the majority of chemical applications, the stereochemical description does not include a precise description of the real 3D molecular structure (which is known relatively rarely), but rather its rough scheme relative to a standard, e.g., two hypothetically possible 3D structures of trans-1,2-dibromocyclohexane are available, but we usually do not focus on this fact. A variety of such uncertainties decided that finding the strict rules for coding molecules for computer applications appeared to be an unexpectedly tedious problem. Moreover, molecular graphs at the early stages of computer technology development were too complex for direct application. The complex nature of molecular graphs can be illustrated by a question asked by Cayley and Schiff: what is the number of all alkanes with molecular formulae $C_{20}H_{42}$, a relatively simple chemical structure (Cayley 1875; Schiff 1875). This can be reformulated into a graph theoretical problem to the question of how many graphs can be arranged to contain 20 C and 42 H nodes with edges fulfilling chemical bonding rules. The answer is 366,319, and it took almost 60 years to answer this question by Henze and Blair (1931) and Reymond et al. (2010). Accordingly, what was done in the early stages of chemoinformatics was a simplification of the chemical graphs.

Mapping Chemical Graphs into 0D Descriptors

The most simple transformation of a chemical graph is a single real number. Mathematically, in this operation, we map chemical information into its simplest form of zero dimensionality (0D), while the calculated number (molecular descriptor) is called a global descriptor. A typical 0D descriptor used for data storage and addressing is:

- Molecular formulae (MF)

An important feature of the mapping of a molecular graph into an MF is that we do not use atom connectivity information for this. We simply sum the atoms of the certain types; in other words, we only annotate the vertexes of the graph. This has the implication that an MF is nonreversible and ambiguous. However, we are often using this notation in the course of documentation and addressing compounds, e.g., in databases or in the literature, e.g., we can use the molecular formula index to search for the appropriate section within the ACS Chemical Abstracts where the compound information can be found.

Mapping Chemical Graphs into 1D Descriptors

Mapping to a 0D representation substantially decreases information on chemical compounds, and 0D descriptors are usually too coarse to code molecules. Instead, 1D descriptors such as:

- Systematic nomenclature
- Line notation

appeared to be enough for the efficient coding of molecules and for database searching.

The concept of a systematic nomenclature scheme (chemical name) by naming rules is human-oriented, which means that easy chemical nameability and name readability by human chemists have been given high priority. However, the systematic nomenclature such as the IUPAC system quite often leads to rather complicated names reflecting the complexity of chemical structures. A human-friendly nomenclature does not necessarily meet the requirements of a computer-oriented system. For example, restricting the name generated for a single structure to a unique value is still a challenge and is to be solved by a preferred name program (PNP) currently under consideration (IUPAC(b), iupac.org). The first system that allowed the input of chemical structures in the form of their chemical names was developed by Beilstein in 1986. This was operated internally at the Beilstein Institute, and the structure input was restricted to the Beilstein notation sub-rules. Currently, the systematic nomenclature is a fully reversible system with software, allowing the mapping of a molecule to a name and a name to a molecule structure.

A line notation is another important system that maps a 2D molecular graph into a linear sequence of letters and numbers. The first systems for this mapping were developed even before computer applications, and a peak of success was in the 1960s and 1970s when this provided a rapid system for coding even large molecules. In a computer system, this was especially suitable due to the compactness of data storage requirements, an important problem at that time. Among historically important line notations:

- Wiswesser (WLN)
- ROSDAL
- SYBYL (SLN)
- SMILES

the latter one, SMILES, evidently predominates today. This is caused by its very simple semantic rules and the intuitive nature of the system. In Fig. 3, we show some illustrative examples. SMILES codes support any chemical entry in Wikipedia, and almost each molecular editor can generate SMILES. Vice versa, SMILES can replace a molecular input into a variety of databases, e.g., Reaxys and/or molecular

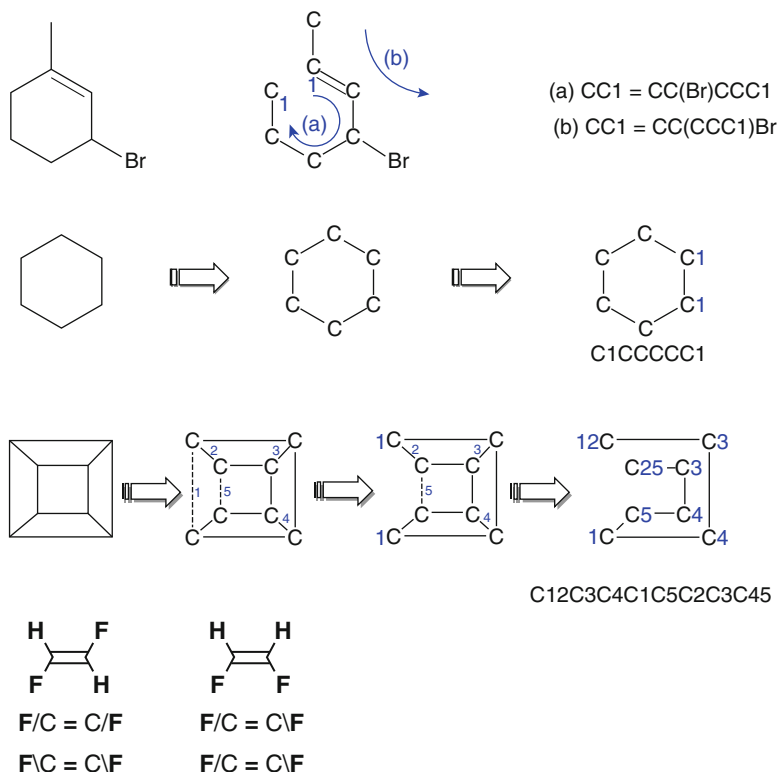


Fig. 3 SMILES, a simple line notation concept transforming a molecular graph (2D) into its 1D representation. In the *upper* part, the notation of some structural elements (double bond and branching at carbon atom) is shown. Below, the formation of chains and stereochemistry in SMILES coding is indicated

editors. A smell of success for this system is in the air if we realize that a Google search can be performed today directly by a SMILES input.

Mathematically, a line notation performs a reduction of information from a 2D into a 1D representation. Dimensionality reduction, while preserving 2D information, explains the success here. Lower dimension means simplification in data manipulation. An important remark is that this dimensionality reduction is fully reversible. So we can both obtain a SIMLES for any molecule and remap this back to reproduce the original molecule.

Mapping Chemical Graphs into 2D Descriptors

Molecular graphs were developed for coding molecules for the needs of human chemists; however, they appeared impractical for feeding computers with molecular

data. Matrix notations or linear codes are the methods developed for mapping molecular graphs into the computer-friendly data. A variety of matrix systems (2D descriptors) were designed, where the numeric representation depends upon the scheme used for coding atoms and bonding. Popular notations include the following matrices:

- Adjacency
- Atom connectivity
- Distance (topological or geometrical)
- Bond
- Bond electrons
- Incidence

Although a matrix system performs well, the number of entries needed to describe a molecule squares with the increase in the number of atoms in the molecule, making it impractical for storing large molecules or large datasets.

A connection table (CT) is an alternative method, allowing us to present molecular graphs in a form of lists of atoms and bonds in a molecule which increase linearly with the number of atoms. A number of CT versions are available, and the reader is referred to Gasteiger (2003) for a detailed introduction.

In Fig. 4, we illustrate an adjacency matrix and in Fig. 5 a connection table and of a simple organic molecule.

A CT notation has now become the predominant form of representing molecular structures. Standard notations such as Molfiles and SDFfiles have been developed that are globally used for information exchange, e.g., in Fig. 6, we illustrate a Molfile coding the 3D structure of salicylic acid by also giving the 3D coordinates of the atoms. In effect, the international language of chemical graphs has been translated into an internationally standardized computer-readable form. (As these representations are coded in ASCII, they can also be read by a human.) An interesting example of coding vitamin C, a chemical compound with all tautomers, by a CT can be found in Bobach et al. (2012).

Today, graphical molecule editors have been developed as tools for the direct creation, modification, or inputting of molecular graph 2D information. Vice versa tools are available for obtaining molecular graphs as outputs from computer-stored CTs.

Mapping Chemical Graphs into 3D Descriptors

A 2D molecular representation defines the molecular topology (atomic constitution and connectivity). However, this in its cleanest form does not code stereochemistry, i.e., the spatial arrangement of the atoms around an atomic center called atomic configuration. This can be coded within chemical graphs by enriching them with solid and dashed wedged bonds, indicating above-the-plane or below-the-plane pointing of the bonds. The configuration is a simplified 3D molecular representation

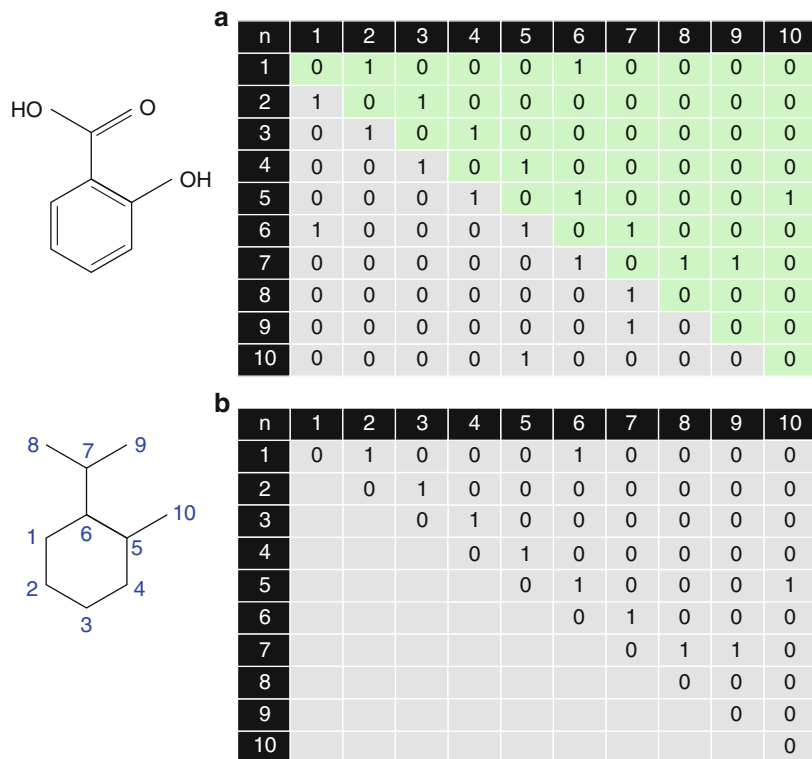
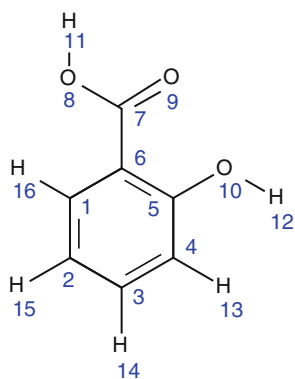


Fig. 4 Two possible forms of an adjacency matrix of salicylic acid: $a_{ij} = 1$, if i and j are bonded or 0 if not bonded

indicating a *version* of the atomic arrangement (stereoisomer) but not the real 3D atomic coordinates.

Stereochemical codes in chemistry are developed to achieve a significant simplification in the visualization of 3D molecules in 2D for the human chemist. Stereochemical information is somewhere in between topology (2D) and topography (3D). In chemistry, beside the solid and dashed wedged bond rule enriching molecular graphs, we illustrate stereochemistry using certain conventions, such as Fischer, sawhorse, or Newman projections. Basically, stereochemistry is a molecular descriptor that defines the arrangement of atoms around an atomic center relative to a standard. The reader should note that this type of molecular descriptor is different from those discussed in section “[Representations for Coding Molecules.](#)” To better understand this, we compared in Fig. 7 the Fischer and Newman projections of stereoisomers of tartaric acid (two enantiomers and one meso compound). A simple compound but a complex problem: a single constitution (chemical compound), two different enantiomers (molecular species, if present in a single substance), and potentially three substances (diastereoisomers). Actually, this is even more



Atom list		Bond list		
1 st atom	2 nd atom	bond order		
1	C	1	2	2
2	C	1	6	1
3	C	1	16	1
4	C	2	3	1
5	C	2	15	1
6	C	3	4	2
7	C	3	14	1
8	C	4	5	1
9	O	4	13	1
10	O	5	6	2
11	O	5	10	1
12	H	6	7	1
13	H	7	8	1
14	H	7	9	2
15	H	8	11	1
16	H	10	12	1

Fig. 5 A connection table of salicylic acid

complicated if we analyze the rotamers in Fig. 7c. Although rotamer I is achiral due to having a center of symmetry, rotamers II and III are in fact chiral, and the achirality of the substance is due to the compensation of chirality contributed by these two chiral conformers. To illustrate how important and complicated stereochemical problems can be if transformed into the computer, a reader can refer to Carbonell et al. (2013).

Basically, two possible representations for the real 3D structure are available. First are 3D atomic coordinates measured as the actual property of compounds (substances), e.g., through crystallography. Accordingly, such data are available in databases (compare section “[Databases and Database Searching](#)”). Second, there are predictions of 3D structures based on molecular modeling. In the second case, this operation can be explained as a mapping of a 2D (CT) or a 1D (SMILES) molecular representation into 3D (see section “[Representation of 3D Structures and Their Mapping into Lower Dimensionality](#)”).

Technically, 3D structures can be noted in several data formats:

- SDfile (Structure Data)
*.sdf
- Molfile (Connection Table Format)
*.mol

salicylic acid

header										
counts line										
16	16	0	0	0	0	0	0	0	0999	v2000
-0.0030	-0.0253	-0.0206	C	0	0	0	0	0	0	0
1.3935	-0.0258	-0.0235	C	0	0	0	0	0	0	0
2.0893	1.1858	0.0002	C	0	0	0	0	0	0	0
1.3954	2.4033	0.0259	C	0	0	0	0	0	0	0
-0.0060	2.4037	0.0315	C	0	0	0	0	0	0	0
-0.6979	1.1866	0.0075	C	0	0	0	0	0	0	0
-0.7803	3.7058	0.0633	C	0	0	0	0	0	0	0
-2.0024	3.6776	0.0628	O	0	0	0	0	0	0	0
-0.1266	4.8729	0.0921	O	0	0	0	0	0	0	0
2.0862	3.6201	0.0462	O	0	0	0	0	0	0	0
3.0329	3.5401	0.0125	H	0	0	0	0	0	0	0
-0.5469	-0.9648	-0.0402	H	0	0	0	0	0	0	0
1.9366	-0.9657	-0.0440	H	0	0	0	0	0	0	0
3.1746	1.1695	-0.0014	H	0	0	0	0	0	0	0
-1.7831	1.1710	0.0105	H	0	0	0	0	0	0	0
-0.6983	5.6335	0.1110	H	0	0	0	0	0	0	0
1	2	2	0	0	0	0	0	0	0	0
1	6	1	0	0	0	0	0	0	0	0
1	12	1	0	0	0	0	0	0	0	0
3	2	1	0	0	0	0	0	0	0	0
2	13	1	0	0	0	0	0	0	0	0
4	3	2	0	0	0	0	0	0	0	0
3	14	1	0	0	0	0	0	0	0	0
5	4	1	0	0	0	0	0	0	0	0
4	10	1	0	0	0	0	0	0	0	0
6	5	2	0	0	0	0	0	0	0	0
5	7	1	0	0	0	0	0	0	0	0
6	15	1	0	0	0	0	0	0	0	0
7	8	2	0	0	0	0	0	0	0	0
7	9	1	0	0	0	0	0	0	0	0
9	16	1	0	0	0	0	0	0	0	0
10	11	1	0	0	0	0	0	0	0	0
M	END									

atom block

bond block

optional properties block

connection table

Fig. 6 Molfile coding of the 3D structure of salicylic acid

- CIF (Crystallographic Information File)
*.cif
- PDB (proteins)
*.pdb

These data files can represent both descriptor and property data and can be converted into each other (compare Chemaxon, chemaxon.com).

Databases and Database Searching

It is quite clear that an overview of the flood of chemical information can nowadays only be managed by databases. With more than 100 million chemical compounds known, it is simply impossible for a human to know from memory for any compound whether it has been made and what its properties are. In this sense, chemoinformatics has delivered indispensable methods for research in chemistry and for the further development of its knowledge.

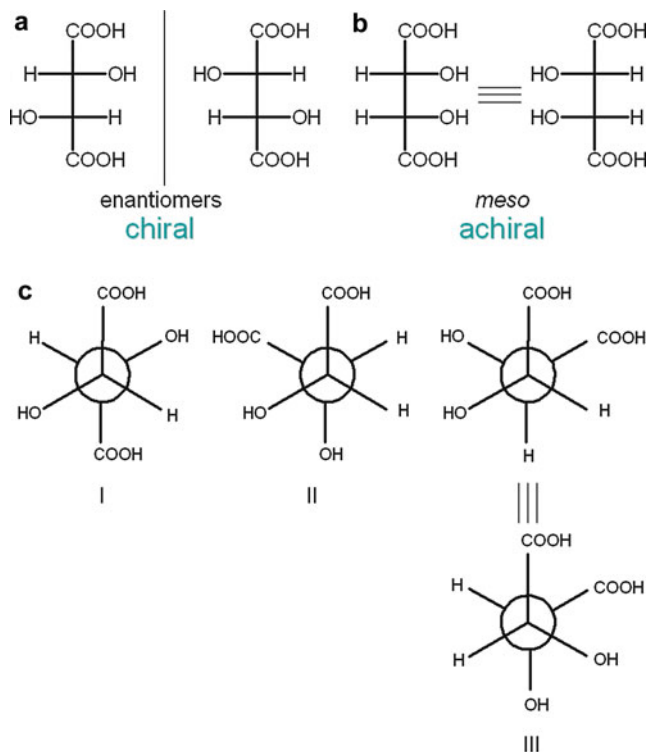


Fig. 7 The stereochemical problem of tartaric acid isomers noted in Fischer (a, b) or Newman (c) projections. Chiral molecules can arrange into chiral substances but also into an achiral one which rotamers I–III are shown in (c)

In section “[Storing of Chemical Structure Information in the Computer,](#)” we discussed the computer representation of molecules (chemical structures). These computer constructions are highly standardized for coding purposes and can be derived from molecules. Otherwise, real properties of substances that are beyond molecules cannot be calculated but are to be measured in experiments. Measured data involving chemical substances are to be eventually organized in the form of searchable and manageable databases. A search engine, the heart of a chemistry database, usually is today an intuitive user friendly system that can be operated by drawing a molecule or by inputting a SMILES code representing a substance, while the searches can be controlled by logical and/or truncation (wildcard) operators (Gasteiger and Engel 2003). In Table 3, we specified some database taxonomies found in literature.

Structure databases, e.g., Gmelin, Beilstein that are now continued by Reaxys (Elsevier) or the Cambridge Structural Database (CSD), illustrate the importance of chemical information for the development of chemical knowledge. Organic or medicinal chemistry or drug design cannot do without database systems. Internet

Table 3 Database taxonomies found in literature

Formal model	Hierarchical
	Network
	Relational
	Object-based
Content	Literature
	Factual
	Numeric
	Meta-databases (entries for other databases)
	Catalogs of chemical compounds
	Research projects databases
	Structure databases where 3D molecular structures are entries
	Small molecules
	Bioinformatics data, i.e., proteins and genetic (DNA and RNA) data
	Chemical reaction databases
	Internet, e.g., web search engine Google
Data type	Biological activity
	Spectroscopy
	Crystallography
	Genetic
	Proteomics

online access to databases is currently a common standard, while both the charged licensed and free systems are widely represented among various organization schemes.

Property data of chemical compounds stored in databases can be simple numerical and factual representations, complex records, e.g., ^{13}C , ^1H NMR, IR, MS spectra or proteomics, protein data. The PDB (protein data bank) structural data format is an example of the data standardization for such data type. The PDB data include more than 100,000 structures (2015/03/03) deposited which are freely available from PDB, www.rcsb.org where also the statistics for the PDB data is given.

The ZINC database is an interesting example of a free database of commercially available compounds for virtual screening. ZINC contains over 35 million FCS compounds that are commercially available from 235 suppliers. The data are prepared in ready-to-dock, 3D formats for a large number of substances: 95,361,026 entries. The chemical compounds (FCS) are annotated in particular with chemical properties and biological activity. The purpose of this database is an easy access to molecular data needed for high-throughput virtual screening. Technically, ZINC stores 3D information in mol2, 3D SDF, and DOCK file formats.

Interestingly, properties are available also for compounds whose molecular structures are not exactly known, e.g., polymers. The CHEMnetBASE, Polymers:

A Property Database (Chemical Databases Online, polymersdatabase.com) contains almost 1000 polymers and 1500 monomers annotated with:

- General description
- Volumetric and calorimetric properties
- Surface properties and solubility
- Transport properties
- Mechanical properties
- Electrical properties
- Optical properties
- Polymer stability
- Applications and commercial products

It is interesting to compare the population of the experimentally measured properties to the registered FCS which is ca. 10^8 (100,000,000):

- Partition coefficient data 10^4 (30,000)
- Structure (crystallographic) data 10^6 (750,000)
- Biological data for screening test (pharma) libraries 10^7

Chemoinformatics has also developed methods that allow an efficient searching for compounds in databases. One can perform searches for full structures, for compounds containing certain substructures, or for similar structures. Various definitions of structural similarity are available. Substructure and similarity searches provide lists of compounds that can be further investigated for the problem at hand.

There is, however, an interesting and important remark to be made concerning the availability of biological data for pharma libraries. The data gathered by pharma industry are not available to the general audience and/or competing companies. This important problem and obstacle in drug design was identified and the so-called *data sharing* concept has been developed to enhance the efficiency of data acquisition and use in pharma. Actually, some procedures and norms for data sharing have been developed (Bird and Frey 2013). Moreover, since we are awaiting significant innovations in medicinal chemistry and in pharma companies, the so-called *collaborative drug design* concept appeared (CDD, collaborativedrug.com). This needs a substantial collaboration and data sharing with pharmaceutical industry. Currently, we believe that this is an absolute requirement for the new era discoveries based on new targets (macromolecular drug receptors) for which personally tailored medicines can be developed in the so-called translational drug discovery approach (Littman and Krishna 2011).

The next problem is that chemical literature publishes only data for positive experiments. This means that chemical journal data are those relating to high-activity compounds. This decides that usually negative data are unavailable and the whole data statistics is skewed. For example, the average potency for the real drugs

on the market is lower than the potency of compounds published in journals (Walters et al. 2011).

Prediction of Properties

Actually measured properties are by definition not available for VCS compounds but are also rare for most FCS compounds. Thus, we often have to replace them with predicted values. Chemistry is a science describing the facts of nature by collecting data in experiments that are performed. The data provide the basis for mathematical models which can be seen as a compressing tool. For example, if the experiment probing two variables provided us with a matrix of parameters of size $10,000 \times 2$ and this can be modeled by a simple linear equation, we are replacing 20,000 numbers by two numbers (*regression*) or even a single number (*ridge regression*).

Regression is a tool developed not only for modeling data but also for predicting data by intra- or extrapolation. We will use here the concept of quantitative structure-activity relationship (QSAR) initially developed by Hansch. QSAR is of fundamental importance in drug design to explain a simple approach to property prediction. In fact, what we are doing in the Hansch approach is modeling drug transport through biological membranes by a partition coefficient (Fig. 8). A partition coefficient is a property measured in an experiment. In drug design, we need, however, to design new molecular objects for which we cannot yet perform experimental measurements. Thus, in a Hansch model, a partition coefficient is replaced by the logP value which can be formally classified as a 0D molecular descriptor. The *additivity concept* provides us with a system in which logP is calculated by summing contributions from structure fragments. These contributions were estimated based on a regression model relating the partition coefficient measured for the substances to fragmental molecular descriptors. The logP value for the whole molecule can then be calculated by simply putting together the increments describing the individual molecular fragments.

Actually, partition coefficients have been measured experimentally only for 30,000 substances (Martel et al. 2013) which is only a tiny fraction of the FCS population. On the basis of this database, we can estimate logP for millions of virtual molecules under design. In fact, even after these substances will be synthesized, the actual partition coefficient values will be measured for them only in rare cases.

This means that the mapping of partition coefficients into logP (both for unmeasured FCS and unavailable VCS compounds) is a substantial extrapolation.

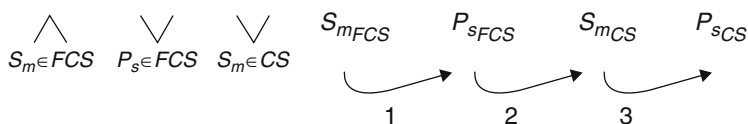


Fig. 8 Schematic notation of the prediction of the partition coefficient by logP calculation. Details in text

However, a real problem of this mapping lies elsewhere. During the design step, we are replacing a substance with an isolated molecule and then come back to the substance after its actual synthesis. Below we illustrate this by the CS operator which helps to understand the problem.

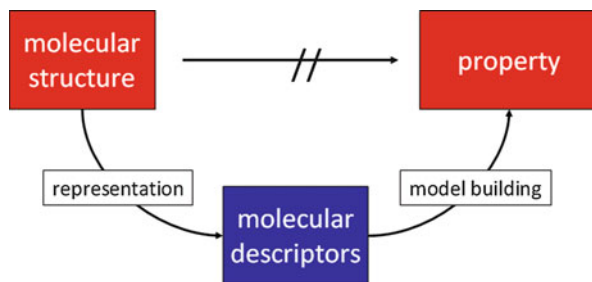
Accordingly, the logP (molecular descriptor, S_{mFCS}) for the training compound series is related to the measured partition coefficient of the related substances (P_{sFCS}) that are available in FCS (mapping 1; physical chemistry in vitro), which provided a basis (regression model) for the calculation of the respective logP (S_{mCS}) for any compound in CS (mapping 2; chemoinformatics in silico). We are usually forgetting about the mapping 3 (hypothesized to be the identity mapping), which means we are using in modeling logP as a replacement for the partition coefficient (P_{sCS}) of the substance under design.

In more or less congeneric series, this replacement usually does not distort the prediction quality, and the mapping can be seen as a reasonable intrapolation, but with an increase in the extrapolation span into the structures which are not well covered by the training substances, we may neglect a variety of real inter- and intramolecular effects encoded in the actual lipophilicity, if measured for substances (Pliska et al. 1996). This effect is not well explored for logP but is very well known for the similar Hammett equation and Hammett constant, where after fascination with linearity, it is a distortion from linearity that provided us really valuable information on chemical reactivity and reaction mechanisms (Schreck 1971). We should remember that logP is not the same as the partition coefficient. Instead, logP is a fragonomics-based molecular descriptor, a correlate of the partition coefficient, aimed to mimic this property.

We can distinguish several basic scenarios for property predictions. Property prediction is based on inter- or extrapolation operation as described above for the partition coefficient. This can, however, be further divided into regression models based directly on another property measured, namely, property-property models, e.g., partition coefficients can be estimated on the substance solubility data (if real data are available for FCS compounds) and alternative models where predictions are made on the basis of molecular descriptors, known as structure-property relationships (compare paragraph 6). A single or multiple property or molecular descriptor representations can be used for this purpose. In most cases, the latter approach will be followed by representing the molecular structures by structure descriptors. Therefore, the entire section “[Representations for Coding Molecules](#)” is devoted to this theme.

From a technical point of view, modeling a dependent variable by a data analysis method usually needs independent variables where each molecule of a dataset is represented by the same number of elements. The application of a mathematical data analysis method requires each object, in this case each molecule, to be represented by the same number of independent variables, i.e., the same number of molecular descriptors. This is a serious requirement for chemistry because the molecules of a dataset usually have different sizes. If a molecule were represented by a single value for each atom, different numbers of variables would result for the various molecules. This would prevent the application of a mathematical data

Fig. 9 The two-step approach to the prediction of properties of compounds that cannot be directly calculated



analysis method such as a multi-linear regression analysis or an artificial neural network. Therefore, the specification of many molecular descriptors must include a mathematical transform of the molecular structure to give for each molecule descriptors of the same number of elements irrespective of its size. This problem is known under the name *uniform-length descriptor* (Baumann 1999). This implies at least one rule to define a fixed number of element in each vector and one rule to fill in the vector elements when the values are missing (Todeschini and Consonni 2000). A variety of mathematical operators, e.g., autocorrelation, can be used to achieve this transform of a molecule. The examples of *uniform-length descriptors* are (Todeschini and Consonni 2000):

- Autocorrelation descriptors
- EVA descriptors
- Topological charge indices
- Molecular profiles
- Spectra descriptors
- Affinity fingerprints
- Radial distribution function (RDF) descriptors

Once we have a dataset of molecules represented by molecular descriptors, a data analysis has to be applied to obtain a model for the prediction of the desired property (see Fig. 9). Thus, in essence, a two-step process is applied: first, the calculation of descriptors for all molecules in a dataset. And then, a mathematical data analysis method such as a statistical or pattern recognition method is used to build a model for property prediction. Note that we talk here about two different types of mathematical methods: one method to make sure that molecules with different number of atoms obtain the same number of descriptors and a second type of mathematical methods for data analysis or building a model. The latter is the substantial domain of chemometrics.

Below we enumerate some of the methods popular in chemoinformatics:

- Statistics, e.g., mean, sum, regression, and correlation analysis
- Neural networks
- Pattern recognition

- Dimensionality reduction (feature selection, feature reduction)
- Principal component analysis (PCA)
- Partial least squares (PLS)
- Fourier analysis
- Wavelet transformations
- Machine learning
- Experimental design

Since currently the trend in computer science goes to open source or freeware projects, also many data analysis methods are available on the Internet. A reader can find a useful online guide for this as well as useful books and manuals, free software or packages, for example, at the www site (StatPages.org, statpages.org), or more specific information on chemical applications at Milano Chemometrics (a) (michem.disat.unimib.it).

Representations for Coding Molecules

Section “[Representations for Coding Molecules](#)” only gives a concise introduction into molecular descriptors. The reader can find a more detailed presentation in the next chapter “Molecular Descriptors” in this book or alternatively in the reference Todeschini and Consonni (2000).

There are many possible molecular descriptor taxonomies. Some of possible descriptor categories are listed here.

Depending upon structure representation type:

- Coding constitution and/or stereochemistry
- Atomic representations
- Fragment codes
- Topology 2D
- Geometry (topography)
- Molecular interaction fields
- Surface annotation (properties)
- Conformational profiles

Calculated for complex molecular aggregates:

- Virtual receptor site (simulated receptor data)
- Ligand-receptor systems (real receptor data)

Substance property correlates:

- Fragment-based substance property correlates, Hansch or Hammett constants

Descriptor data type and structure:

- 0D (single number), 1D (vector), 2D (map, pixel), and 3D (volume, voxel)
- Fingerprints
- Hashed fingerprints
- Structure keys
- Boolean
- Integer number
- Real number
- Tensor
- Scalar field
- Vector field

Several descriptors important for storing chemical structure information in computers were discussed in section “[Storing of Chemical Structure Information in the Computer](#).” Below we present some other methods developed for in silico chemistry, for the prediction of properties, with an emphasis on dimensionality mapping or distortion during these transformations.

Mapping Chemical Graphs into 0D (Global) Descriptors

The simplest transform of a chemical graph is a single real number. Mathematically, in this operation, we map chemical information into its simplest form of 0D dimensionality, while the calculated number (molecular descriptor) is called a global descriptor. Some representative examples are:

- Molecular weight (MW)
- Average molecular weight
- Atom counts
- Bonds counts
- Multiple bonds
- Double bonds
- Triple bonds
- Number of atoms
- Number of rings
- Sum of atomic volumes
- Total polar surface area
- logP
- Single number connectivity indices

LogP is a classical example where we need both connectivity information and the annotation of atoms with fragmental logP (hydrophobicity) descriptors for this calculation. Some global descriptors can be calculated even without connectivity information, e.g., MF (compare section “[Storing of Chemical Structure Information](#)”

in the Computer”) or MW which does not need connectivity information but only atomic weights. Alternatively, geometric bonding information and atomic charges are needed for dipole moment calculations, while the precise 3D atomic arrangement is needed for the calculation of the moment of inertia.

In the context of the property versus descriptor differentiation, we can ask a question if atomic annotations (MW, fragmental logP) are atomic descriptors and not measured atomic properties. Fragmental logPs are property correlates, however, that means clearly a molecular descriptor but not a property. Partial atomic charges are computer-simulated values and not a measured property. We can doubt about atomic weights which were measured in experiments. Actually, the calculation of the MW descriptor is entirely analogous to the logP calculation. The only difference is that MW is obtained by summing atomic contributions so accurate a simulation that the resulting descriptor will give a 100 % correlation with the property (MW) measured for the whole molecule in a potential real-world experiment with a single molecule (MS spectroscopy) or a substance (e.g., cryometry).

Mapping 2D Molecular Graphs into 1D and 2D Descriptors

1D line notations and 2D matrix representations that are used for storing chemical structure information are also popular for *in silico* structural manipulation. In general, transforms to lower dimensions mean also a distortion of information, so we cannot fully recover the information of the original. The exclusion here are descriptors designed for addressing and storing information (paragraph 6) in which a full reversibility of information is a substantial need. In particular, mapping 2D molecular graphs into 1D or 2D can provide us with representations that do not distort information. Note especially the SIMLES (1D) descriptors which are both unambiguous and fully reversible (paragraph 6.2). The availability of these descriptor types decides that they are also used for other *in silico* application.

Technically, some important and popular 1D molecular descriptors can be classified as (Todeschini and Consonni 2000; Gasteiger and Engel 2003):

- Structure keys
 - Augmented atoms
 - Atoms pairs
 - MDL keys
 - CACTUS screen vectors
- Hash codes
 - Topology based
 - Stereochemistry based
 - Atomic charges based
- Molecular fingerprints
 - Daylight fingerprints
 - Unity fingerprints

- Holograms
- MOLPRINT-2D fingerprints
- Extended connectivity fingerprints
- Functional connectivity fingerprints

Structure keys define structure motifs or fragments in the form of a Boolean vector, i.e., whether present (1), otherwise (0). Fingerprints are Boolean vectors that define a set of patterns. A pattern can, for example, be a path of predefined length, where each path is given by the atom and bond annotations along the path. Hashed fingerprints provide fingerprint-type information of all possible paths that traverse a molecule with up to certain number of bonds. Hash structure codes are vectors developed for storing information, where hashing means to cut or chop the input into smaller information pieces of accepted number within a file length (Todeschini and Consonni 2000; Gasteiger and Engel 2003).

Representation of 3D Structures and Their Mapping into Lower Dimensionality

The representation of 3D structures was discussed in paragraph 6. 3D structure data can be both a molecular descriptor (if predicted) and a real property if measured.

The expansion of 2D molecular graphs (with stereochemical information where needed) to structures with 3D atomic coordinates can be obtained by:

- 3D structure generators, e.g., CORINA (data- and rule-driven methods)

Additionally,

- Ab initio quantum chemistry
- Semiempirical quantum chemistry
- Molecular mechanics

are available further to refine the 3D structure (e.g., as given by a Z-matrix).

Basically, two different methods can be used for coding a 3D molecular representation in complex molecular descriptors. Both are based on a Cartesian coordinate system defining atomic positions. In the first, we are supplementing this by a connection table listing the bonding between atoms. The second codes are based on internal coordinates, i.e., bond lengths, bond angles, and torsion angles. An example of such a notation is the so-called Z-matrix. Although a Z-matrix seems to be a less human-friendly option for *in silico* modeling, it offers the advantage of providing an explicit list of internal molecular data which can be directly used for *in silico* simulations.

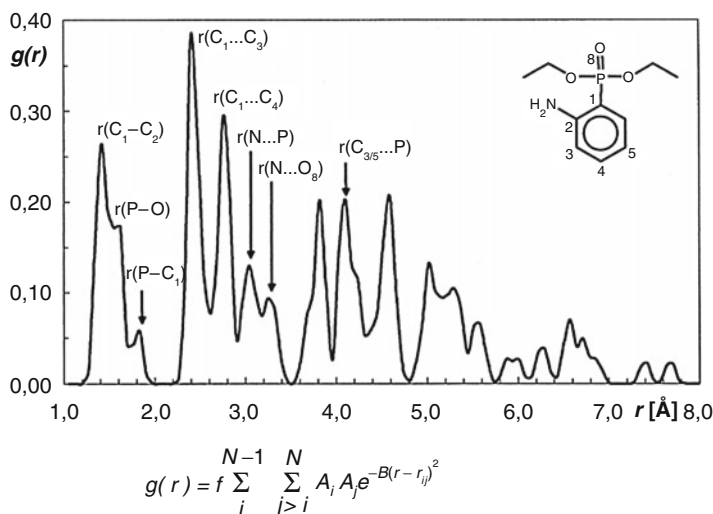


Fig. 10 Radial distribution function (RDF). A 2D molecular descriptor for coding 3D molecular structures with atomic numbers as atom property, where f is a scaling factor, N is the number of atoms, r is the distance between atoms i and j , and B is a smoothing parameter (Hemmer et al. 1999)

Other approaches are molecular transforms of the 3D information into lower dimensionality data. Below we give some illustrative examples:

- **3D to 2D**

- 3D to 2D mapping is a 2D self-organized map of partial atomic charges. This appeared as an efficient model for the comparison of a series of bioactive molecules and a simple simulator of a biological receptor, i.e., the *receptor-like neural network* (Polanski 1997).
- 2D maps of atomic poses (3D structural data) are used in SOM-4D-QSAR mapping (Polanski and Bak 2003).

- **3D to 1D**

- Radial distribution function (RDF) is an interesting method here (Hemmer et al. 1999). This was illustrated in Fig. 10.
- Molecular Shape Spectrum in 4D QSAR (Hopfinger et al. 1997).

- **3D to 0D**

A simple transformation of the 3D structural data into the *principal moments of inertia* about any axis (most often along the X, Y, or Z coordinates) is defined as

$$I = \sum_{i=1}^A m_i \cdot r_i^2$$

where A is the atomic number, m_i is the atomic mass, and r_i is the perpendicular distance from the chosen axis.

An interesting example of a useful molecular descriptor obtained by a 3D to 0D mapping of the structural data is the *RMSD* molecular descriptor (root-mean-square deviation of atomic positions):

$$RMSD = \frac{1}{n} \sqrt{\sum_{i=1}^n (v_{ix} - w_{ix})^2 + (v_{iy} - w_{iy})^2 + (v_{iz} - w_{iz})^2}$$

$v = [v_x, v_y, v_z]$ and $w = [w_x, w_y, w_z]$ structural data for two molecular representations, n = number of atoms.

This is used for the comparison of two differently modeled 3D structures or for the comparison of the modeled structure with the real 3D structure.

Representation of Molecular Surfaces

Clearly, a molecular surface is a molecular descriptor and not a property. As this picture is very illustrative and satisfies our imagination of the molecular world, we use it to explain and understand molecular effects. Actually, this is not necessarily a true copy of a molecular body. For example, the repulsive effect of the atoms shaped into hard sphere-like entities used to explain the energy barrier between the conformers, explaining the energy preference for the staggered conformer over the eclipsed one. The sphere-like ethane can be found in any chemical textbook. However, it has been found just recently that steric repulsion is only of secondary importance for the determination of the ethane folding principles. In fact, quantum mechanics shapes the molecule into a form in which steric repulsion favors the eclipsed conformer. Instead, hyperconjugation explains the rotation of the single C-C bond into the staggered form (Weinhold 2001).

A number of digital descriptors are available for the quantification of shape effects related to the surface (Mezey 1993; Polanski 2003). Shape descriptors are:

- Molecular shape indices
- Minimal steric difference (MSD)

that can be estimated from molecular graphs, i.e., from 2D representations after their expansion into 3D structures.

Alternatively, real 3D structures can be used for the description of molecular shapes by performing direct calculations or by performing complex *in silico* simulations. Some illustrative examples are:

- Van der Waals or Connolly surfaces
- Solvent accessible surface area

- Molecular shape analysis, MSA
- Molecular volumes, STERIMOL
- Quantitative shape-activity relationship, QShAR
- Comparative molecular surface analysis, CoMSA

A molecular surface substantially reframes the digital discontinuous form of the structure view to a continuous landscape-like picture. Thus, digitalization is not always the main point in the description of shape representations illustrated by surfaces. Instead, visualization provides the data for molecular shape analyses. Illustrative examples of such analyses can be found in reference Zupan and Gasteiger (1999) or under the entry topological feature maps in reference Todeschini and Consonni (2000).

Handling of Conformational Flexibility

Molecules are dynamic objects capable of rotation along single bonds or stimulated with higher energies even along multiple bonds. The set of rotamers provides us with the so-called conformational profile space of a molecule. Rotamers is a term coined in organic chemistry which is directly connected to conformational analysis both based on an intuitive analysis or on qualitative measures called *steric effects* or quantitatively probed by precisely calculated energies values. The conformational space of rotamers usually focuses our attention on several poses. These rotamers should easily be imagined and visualized without computer-aided methods with typical values of bonding angles and/or the lowest or highest energy states. A *pose* is a less strictly defined molecular conformation populating a conformational profile space. Typically, the term pose refers to computer-generated conformations, and the number of poses can reach very high numbers. Conformational information is a molecular descriptor much like stereochemical information. However, we can measure the property of a substance composed of certain conformations if they are stable enough.

Represented by poses, a conformational profile space that is typical for each molecule can be transformed into a 3D map in a procedure that resembles the 2D mapping of molecular surfaces. Thus, if we replace a pixel (2D) of the map with a volume unit called voxel (3D), we can easily explain the analogy. Voxels can be annotated with molecular descriptors which refer to a single molecule or describe the difference between a given molecule and a reference molecule. The simplest voxel-counting type is the atom count, i.e., the number of atoms found in a unit cube. Alternatively, partial atomic charges can be summed or averaged in voxels. This method is used for the calculation of the grid cell occupancy descriptors, GCOD, 4D-QSAR (Hopfinger et al. 1997), or their fuzzy neural network equivalents, SOM-4D-QSAR (Polanski and Bak 2003), while in both methods, the pose set has been generated by molecular dynamics simulations.

Representation of Chemical Reactions

As said in the Introduction, chemistry is the science of chemical compounds and their transformations. In this sense, the computer representation of chemical reactions presented in the very beginning an essential challenge for chemoinformatics. It was soon realized that an efficient representation of chemical reactions has to go beyond the storing of starting materials and products of a reaction. To answer important questions about a chemical reaction, it is necessary to store information on the essential changes in a chemical reaction, the breaking and making of bonds in the course of a reaction. These changes are represented by the so-called reaction center, the atoms and bonds directly participating in the transformation of the starting materials into the products.

We will here not go deeper into those tasks and how they were solved. Suffice to say that indeed these tasks were solved and, thus, the basis for reaction databases and for systems on computer-assisted synthesis design has been laid in section “[Applications of Molecular Descriptors for Property Prediction](#).” More details can be found in Gasteiger (2003).

Trends in Structure Representation

It is more and more realized that chemoinformatics can help in increasing our chemical knowledge. In this endeavor, the trend in property prediction shifts from models that are good in predictive capability to models that can also be interpreted.

To this effect, molecular descriptors have to be used in the model building process that is prone to interpretation. Two types of descriptors can be used for this purpose, fragmental descriptors and descriptors that are based on physicochemistry. A variety of fragment-based descriptors are available. Recently, fragments that had been developed for toxicity alerts were extended and made publicly available (Yang et al. 2015; Chemotyper, chemotyper.org). These chemotypes can also be used for analyzing properties other than toxicity. Furthermore, the atoms and bonds of these substructures can be enriched by physicochemical effects such as partial atomic charges.

A range of physicochemistry-based descriptors are available; quite a few have been listed in section “[Representations for Coding Molecules](#).” A typical example is the logP descriptor that models the influence of lipophilicity on a property. In a more global approach, a hierarchy of structure descriptors has been developed that combines geometry with physicochemistry (Gasteiger 2006). In this approach, geometry goes from 2D through 3D to molecular surfaces. And physicochemistry encompasses a variety of effects such as partial atomic charges, electronegativity, or polarizability effects. Mathematical methods are used for combining geometry with physicochemistry including the use of mathematical transforms to obtain the same number of independent variables irrespective of the size of the molecules in the dataset.

The endeavor of obtaining models that can be interpreted also must have consequences on the choice of a data analysis method. Rather than using highly sophisticated modeling techniques that squeeze the ultimate out of a dataset – and thus quite often hide the influence of the molecular descriptors used – methods have to be employed that clearly show the effect of each descriptor used. For example, a simple multi-linear regression analysis might not produce the highest prediction power of a model, but it clearly shows the influence of each descriptor on the property under consideration.

This trend in property prediction is also reflected in publishing policies. More and more will manuscripts on QSAR studies only be accepted for publication if they also provide new data or if they lead to more insight into the effects influencing the modeled property.

From a practical point of view, a variety of descriptors can be calculated today in an online mode or by freeware packages both for individual molecules and in batch mode for a series of molecules. Just to cite as an example of mi-logP values calculated online, visit directly the web page of Molinspiration, molinspiration.com. The method is robust enough to *process* practically all organic and most organometallic molecules. The user can input a SMILES or draw a 2D molecular graph. Other molecular descriptors called here *molecular physicochemical properties* that are available from the site are:

- 3D molecular structure data
- Molecular polar surface area TPSA
- Molecular volume
- “Rule of 5” parameters
- Number of rotatable bonds

An interesting new option are databases with calculated descriptors, e.g., we can find a database of 1124 descriptors calculated for 234,773 molecules (Milano Chemometrics (a), michem.disat.unimib.it).

Applications of Molecular Descriptors for Property Prediction

Formally, property-to-property relationships are limited to FCS compounds with measured properties. The solubility or chromatographic behavior versus lipophilicity (quantitative structure-retention relationships) is an illustrative example here. As, however, properties are rare, we are often replacing a real property-property systems with property versus property mapping which are then extended to the extrapolated values (compare $\log P$ calculation discussed in section “[Prediction of Properties](#)”).

However, the most popular method here is a structure-property strategy based on using molecular descriptors as independent variables. For example, a discussion of molecular descriptors available for the prediction of oral bioavailability and their performances can be found in reference Ahmed et al. (2012). Drug design is an

example of the structure-property method aimed at predicting the biological activity of chemical compounds.

Drug Design

Drug design is a central challenge to chemistry and pharmacy. Accordingly, chemoinformatics has found many applications to this problem. The majority of drugs are molecules capable of interaction with macromolecular targets controlling processes in living organisms. This paradigm defines two basic strategies for searches for new drugs. First are structure (receptor dependent)-based methods, and second are ligand-based (receptor independent) ones.

Molecular design is often used as a synonym for *drug design*. This interchange illustrates the complex nature of the problem showing a hesitation if we are to focus on a molecule or a substance (drug). The so-called lead structure, namely, an active compound, is an essential need for starting any development project for a new drug. A lead structure can be found as a consequence of:

- Serendipitous discovery
- Natural substance screening
- Screening chemical compounds resulting from synthetic chemistry
- Computer-aided structure-based design
- Ligand-based screening of chemical libraries (both of real or virtual compounds)
- Combinatorial chemistry
- Receptor-based dynamic combinatorial chemistry (property-oriented combinatorial chemistry)
- Receptor-based chemical methods, e.g., NMR methods

A general disagreement between the lead structure-activity versus its complexity (Zartler and Shapiro 2005) is a substantial dichotomy in drug design which was realized just recently (Fig. 11). Since the development of chemical, physical, and biological technologies allowed the identification of compounds with increasingly lower activities, the population of potential lead structures significantly increased. However, the activity level drastically decreases for low molecular compounds. Just recently, this paradigm has founded fragonomics, i.e., the strategy of the identification of active small fragments (low MW molecules) and the rules for connecting them into compounds of higher complexity and activity. The potential of this method is broadly investigated now. For example, we realized that the activity and complexity of real drugs are lower, *slim pharma* than these under R&D development, *molecular obesity* (Hann 2011). What should be the advantageous range of logP values, the preferred ADMET range, etc., for fragments, are examples of problems under current diagnosis.

Once a lead structure is found, it needs to be optimized and/or modified for the requirements of an individual receptor or target system. Due to the needs for substantial simplification, we are usually using here the model of a single

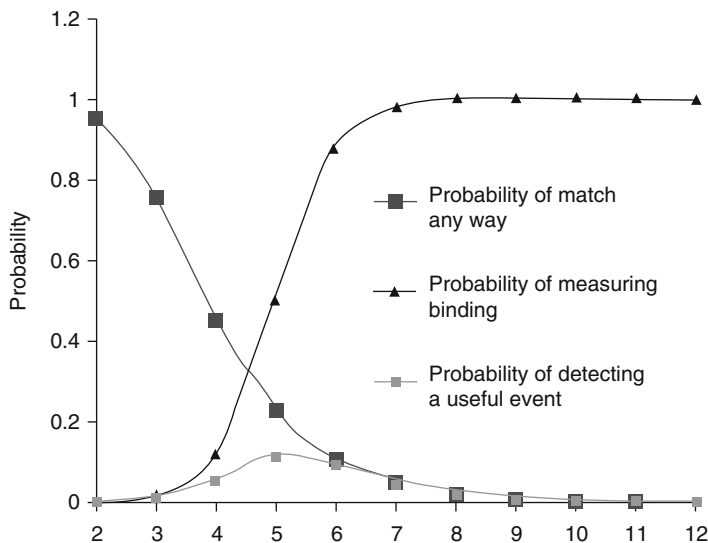


Fig. 11 The complexity versus binding affinity dichotomy. Low MW ligands are more probable to be identified due to statistics, but their affinities are lower than for chemical compounds with larger complexity and higher MW (Adapted from Zartler and Shapiro 2005)

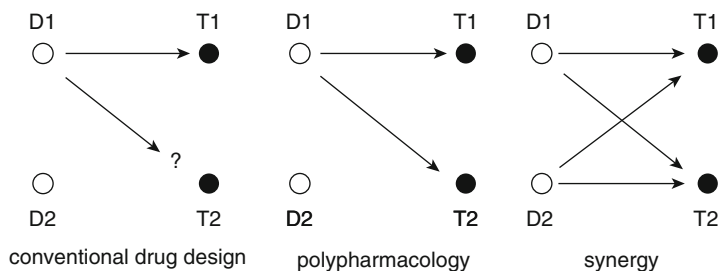


Fig. 12 The increase in complexity from conventional drug design to the synergy by replacing the single-drug to single-target interactions with the multidrug D1, D2 (mD) to multi-target T1, T2 (mT) system (Adapted from Musiol et al. 2014)

drug-single target. The challenge of information complexity increases when replacing a single by a multidrug target system as explained in Fig. 12.

The requirement for optimization is, however, not a single but a series of structures. Accordingly, the question arises: how should we modify a lead structure? There are a number of concepts and rules starting from the technical ones, e.g., Topliss schemes, through chemical ones:

- Changing of substitution type and pattern
- Homologation

- Ring and/or chain transformations (enlargements, shortening, branching, transformation, increasing system rigidity, or vice versa increasing system flexibility)
- (Bio)isosterism
- Peptidomimetics

to the chemoinformatics-based chemical information metrics targeted at measuring the design domain in CS covered by chemical compounds under the R&D project. Accordingly, we can use for this the measures of compounds':

- Similarity
or
- Diversity

QSAR is an example of the optimistic strategy aimed at drug design in which we are attempting to quantitatively predict the activity of the new substance, a potential drug. In its simplest form, this strategy can be noted as follows:

$$A = F(S)$$

where A is the activity and S is the molecular descriptors.

A Free Wilson analysis (synonyms are *additivity model* and *de novo approach*) is a property modeling method based on a truly structure-property scenario. In this method, we are searching for a regression model where the independent variable block is the so-called group contribution of two substituent series mounted at two positions of the reference compound. The resulting regression coefficients of the indicator variables are the biological activity contributions of the corresponding structural elements. The efficiency of this approach shows that, in fact, the mapping to the partition coefficient is not necessarily a requirement. An interesting comparison of the methods can be found in reference Kubiny (1993). Recently, it has been shown that structure information in structure-property scenario can be coded directly by SMILES codes (Worachartcheewan et al. 2014).

The last decades brought the idea of applying in QSAR studies rather complex molecular descriptors based on molecular interaction fields, conformational compound profiles, or even describing ligand-receptor systems (receptor-dependent) both with simulated and real receptor data. In Table 4 we present a QSAR classification according to the dimensionality that can be found in the literature (Polanski 2009b). This classification is not fully consistent because $m = 0$ and $m = 1$ refer to the data structure of molecular descriptor, i.e., one element (univariate data) vector in 0D and a vector (multivariate)-type molecular descriptor where each element can describe another global property in 1D; while in 2D to 6D, the dimensionality indicates the type of molecular representation tested. Vector type data (dimension = 0 or 1) is always needed in 0÷6D QSAR modeling. For example, a SMILES line notation is a 1D representation, but it was designed to code 2D matrix molecular graphs. The term receptor-independent or receptor-dependent notation is sometimes used to indicate if we include real receptor data, e.g., 3D structure,

Table 4 Classification of the receptor-independent QSAR, 0÷6D (Polanski 2009)

QSAR dimension m^a	Molecular representation used for molecular descriptor calculation	Information dimension	Example of S variable
0D	Single number (univariate X data) parameter referring to global descriptors	0D	logP
1D	A multivariate combination of global descriptors or 1D descriptors	1D	logP and Hammett constants, Free Wilson analysis, SMILES-based QSAR
2D	Structural patterns relating to 2D molecular representations	2D	Connectivity, 2D pharmacophores, SMILES-based QSAR
3D	Single-state 3D ligand representations	3D	Molecular interaction field (MIF) descriptor in CoMFA
4D	Multiple 3D ligand representations (conformations, orientations, protonation)	3D	Grid cell occupancy descriptor (GCOD; 4D QSAR)
5D	3D QSAR + induced fit scenarios	3D	Quasar (quasi-atomistic receptor surrogate) in 5D QSAR
6D	5D QSAR + multiple solvation scenarios	3D	Quasar in 6D QSAR

into the analysis. Below we present a possible coherent version of this approach organized according to the information complexity where:

Dimension indicates dimensionality of the molecular descriptor

- 0D Single number, global 0D property or descriptor
- 1D 1D molecular descriptor

Dimension indicates molecular representation used for molecular descriptor calculation

- 2D 2D information
- 3D Single static 3D ligand states
- 4D Multiple ligand states
- 5D Virtual or pseudoreceptor constructed from the active ligand series
- 6D 5D + multiple solvation
- 7D Real receptor data

Such an m-QSAR classification resembles somehow to the way chemists refer to the protein structures. The amino acid sequence defines the primary structure, the local organization defines the secondary structure, the overall protein shape describes the tertiary structure, and the effects resulting from the interactions with

another protein molecule are explained in the quaternary protein structure. Thus, dimension refers here more to data complexity than to real topology (topography).

There are a variety of properties that are of substantial importance in chemistry and drug design. Examples are absorption, distribution, metabolism, excretion, and toxicity (ADMET) in pharmacokinetics. Aqueous solubility can be a useful property that can be used in property-property models (Gasteiger and Engel 2003), and Caco 2 cell line membrane transport behavior has been recently applied to the prediction of substance-cell absorption (property-property approach) in pharma industry contributing to our understanding of drug efflux transporters such as P-glycoprotein (ABCB1) and BCRP (ABCG2). This proved to be an unexpectedly successful approach. Accordingly, Caco 2 behavior was simulated *in silico* in the so-called membrane interaction QSAR with 4D fingerprint descriptors to provide this property prediction (Iyer et al. 2007).

Toxicity is another of the abovementioned substance properties of key importance not only in drug design but also generally in chemistry and everyday life which can be illustrated by the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) regulations (REACH, ec.europa.eu). The complex software based on multidimensional Quasar descriptors, including receptor dependent models, is available from Vedani group as a freeware OpenVirtualToxLab™ license for toxicity predictions (VirtualToxLab, biograf.ch).

Structure Elucidation

Currently a variety of physical methods are available for structure determination and elucidation:

- Electron diffraction
- X-ray diffraction
- ¹H and ¹³C NMR
- Mass spectrometry
- Infrared spectroscopy

For each compound synthesized in the laboratory or isolated from natural products, routinely IR, NMR, and mass spectra are taken. Interpretation of the data generated during these experiments and finding structure-spectra correlations are central problem in organic chemistry. Computer-assisted structure elucidation (CASE) is a generic name for the computer techniques used for this purpose. Generally, the strategy under CASE is a property (spectra) to structure (2D or 3D molecular descriptor data) mapping. The CASE problem was addressed by chemoinformatics already in the 1960s. However, due to the complexity of the problem, no general-purpose system is yet available.

The general problem of structure elucidation will not be discussed here any further, and the reader should consult the available literature. An interesting example

of CASE is a counter-propagation neural network that we input with the IR spectra coded by Fourier or Hadamard coefficients (2D data) and radial distribution function (RDF) for code structure representations (3D data). Due to the full reversibility of the RDF code, the IR spectrum in this method could be directly mapped into the 3D molecular structure (Hemmer et al. 1999).

Organic Synthesis Design

Several groups in the USA, Germany, England, and Japan started with the development of systems for computer-assisted synthesis design (CASD) already in the late 1960s and early 1970s (Corey and Wipke 1969; Gelernter et al. 1973; Hendrickson 1971; Blair et al. 1974). At that time, many basic problems in chemoinformatics such as storing chemical structures, unique and unambiguous coding of molecules, ring and aromaticity perception, and representation of chemical reactions had to be solved. This work laid the foundation to the development of databases on chemical structures and reactions such as MACCS, REACCS, and the Beilstein Database which can all be traced back to efforts made in the CASD groups.

On the other hand, no CASD system has yet found general acceptance by organic chemists. The reasons are manifold, not the least that organic chemists want to personally take the challenge of designing their syntheses intellectually without assistance by a computer. However, what Herbert Gelernter has said in 1973 is still valid: “The amount of information to be processed and the decisions between many alternatives suggests the use of computers in organic synthesis design.” In this sense, we can expect that some time in the future, CASD systems of global usefulness will appear and be used.

Chemical Compounds Beyond Structure Representation

There are many compounds and substances where an exact molecular structure is not known or cannot be given. Examples are plastic materials, mixtures of compounds (see Fig. 1: chemical species), many inorganic materials, etc. Nevertheless, there is definitely interest to also store such species in databases or to predict their properties. In such cases, representations other than molecular descriptors have to be chosen. These representations can be representations of the compounds that they were made of and their respective ratio and experimental conditions of their preparation. Other important descriptors are physical or chemical data for the characterization of such substances. Of particular value for property prediction are spectral data because they give important indirect information on the composition and structure of those substances. Not much work has been done in this direction, but for sure many studies will come.

Conclusions

Cheminformatics has substantially changed the availability of information on chemical compounds in databases. The present level of chemical investigations would be inconceivable without access to databases on chemical compounds and chemical reactions. Efficient database mining and knowledge discovery methods changed the way in which we are designing chemical substances and their reactions. Today there is no problem for finding selective or specific reactions among a variety of possible options. There is also no problem for finding equivalent reagents for various possible synthons and available commercial chemicals for them.

Predicting properties is the second important area of cheminformatics applications. We still need to improve our efficiency here since the available FCS data on real properties are still rare for chemical compounds. From the data analysis point of view, we have developed a variety of methods efficient in this area.

Molecular descriptors and properties are two basic forms of chemical information representing compounds in the computer. Although it is not always easy to distinguish between these information types because we often need to predict properties, this helps in understanding structure-property approaches and the possible origins for their failures. A variety of molecular descriptors is available both for storing molecules and for their processing in computers, and we do not have here a sharp differentiation.

The trend in property prediction goes away from pure modeling and prediction endeavors to the development of models that can also be interpreted and thus help us in increasing chemical insight and knowledge. However, probably the most important aspect of cheminformatics is that it is allowing us for a better understanding of chemistry. In this context, we must first clearly define a chemical problem, avoiding any possible uncertainties to be able to teach this to the computer. Secondly, the availability of chemical information at our desks improves our learning ability in this area and increases our chemical knowledge.

Bibliography

- Ahmed, S. S. S. J., & Ramakrishnan, V. (2012). Systems biological approach of molecular descriptors connectivity: Optimal descriptors for oral bioavailability prediction. *Plos One*, 7(7), e40654.
- Bajorath, J. (Ed.). (2004). *Cheminformatics: Concepts, methods, and tools for drug discovery*. Totawa: Humana Press.
- Baumann, K. (1999). Uniform-length molecular descriptors for quantitative structure-property relationships (QSPR) and quantitative structure-activity relationships (QSAR): Classification studies and similarity searching. *TRAC-Trends in Analytical Chemistry*, 18(1), 36–46.
- Bird, C. L., & Frey, J. G. (2013). Chemical information matters: An e-research perspective on information and data sharing in the chemical sciences. *Chemical Society Reviews*, 42(16), 6754–6776.
- Blair, J., Gasteiger, J., Gillespie, C., Gillespie, P. D., & Ugi, I. (1974). Representation of constitutional and stereochemical features of chemical systems in computer-assisted design of syntheses. *Tetrahedron*, 30(13), 1845–1859.

- Bobach, C., Boehme, T., Laube, U., Pueschel, A., & Weber, L. (2012). Automated compound classification using a chemical ontology. *Journal of Cheminformatics*, 4, 40.
- Boehm, M. (2011). Virtual screening of chemical space: From generic compound collections to tailored screening libraries. In C. Sotriffer (Ed.), *Virtual screening*. Weinheim: WILEY-VCH.
- Brown, F. K. (1998). Chemoinformatics: What is it and how does it impact drug discovery. *Annual Reports Medicinal Chemistry*, 33, 375–384.
- Bunin, B. A., Bajorath, J., Siesel, B., & Morales, G. (2007). *Chemoinformatics: Theory, practice, and products*. New York: Springer.
- Carbonell, P., Carlsson, L., & Faulon, J.-L. (2013). Stereo signature molecular descriptor. *Journal of Chemical Information and Modeling*, 53(4), 887–897.
- Cayley, E. (1875). Ueber die analytischen Figuren, welche in der Mathematik Baume genannt werden und ihre Anwendung auf die Theorie chemischer Verbindungen. *Chemische Berichte*, 8, 1056–1059.
- CDD, Collaborative Drug Discovery. (2016). <https://www.collaborativedrug.com>
- Chemaxon, MolConverter. (2016). <https://www.chemaxon.com/products/marvin/molconverter>
- Chemical Databases Online. Polymers: A properties database. (2016). <http://www.polymersdatabase.com>
- Chemotyper, The ChemoTyper application. (2016). <https://chemotyper.org>
- Consonni, V., Todeschini, R. (2010). Molecular descriptors. In T. Puzyn et al. (Eds.), *Recent advances in QSAR studies* (pp. 29–102). Springer, Amsterdam, The Netherlands
- Corey, E. J., & Wipke, W. T. (1969). Computer-assisted design of complex organic syntheses. *Science*, 166(3902), 178–192.
- Dobson, C. M. (2004). Chemical space and biology. *Nature*, 432(7019), 824–828.
- Foulon, J.-L., & Bender, A. (2010). *Handbook of chemoinformatics algorithms*. London: CRC Press.
- Gasteiger, J. (Ed.). (2003). *Handbook of chemoinformatics: From data to knowledge, 4 volumes*. Weinheim: Wiley-VCH.
- Gasteiger, J. (2006). 2006 American chemical society award for computers in chemical and pharmaceutical research – Of molecules and humans. *Journal of Medicinal Chemistry*, 49(22), 6429–6434.
- Gasteiger, J., & Engel, T. (Eds.). (2003). *Chemoinformatics: A textbook*. Weinheim: Wiley-VCH.
- Gelernter, H. L., Sridharan, N. S., Hart, A. J., & Yen, S. C. (1973). The discovery of organic synthetic routes by computer. *Topics in Current Chemistry*, 41, 113–150.
- Hammond, G. S. (1968). Norris award lecture, private communication.
- Hann, M. M. (2011). Molecular obesity, potency and other addictions in drug discovery. *Med-ChemComm*, 2(5), 349–355.
- Hansch, C., & Fujita, T. (1964). Rho-sigma-pi analysis. Method for correlation of biological activity + chemical structure. *Journal of the American Chemical Society*, 86(8), 1616.
- Hemmer, M. C., Steinhauer, V., & Gasteiger, J. (1999). Deriving the 3D structure of organic molecules from their infrared spectra. *Vibrational Spectroscopy*, 19(1), 151–164.
- Hendrickson, J. B. (1971). Systematic characterization of structures and reactions for use in organic synthesis. *Journal of the American Chemical Society*, 93(25), 6847–6854.
- Henze, H. R., & Blair, C. M. (1931). The number of isomeric hydrocarbons of the methane series. *Journal of the American Chemical Society*, 53, 3077–3085.
- Hopfinger, A. J., Wang, S., Tokarski, J. S., Jin, B. Q., Albuquerque, M., Madhav, P. J., & Duraiswami, C. (1997). Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. *Journal of the American Chemical Society*, 119(43), 10509–10524.
- IUPAC(a). (2016). Goldbook. <http://goldbook.iupac.org>
- IUPAC(b). (2016). Strategy round table. Representations of molecular structure: Nomenclature and its alternatives. http://www.iupac.org/fileadmin/website/news/2000/NRT_Report.pdf
- Iyer, M., Steng, Y. J., Senese, C. L., Liu, J., & Hopfinger, A. J. (2007). Prediction and mechanistic interpretation of human oral drug absorption using MI-QSAR analysis. *Molecular Pharmaceutics*, 4(2), 218–231.
- Karthikeyan, M., & Vyas, R. (2014). *Practical chemoinformatics*. New Delhi: Springer.

- Kubiny, H. (1993). *QSAR: Hansch analysis and related approaches*. Weinheim: VCH.
- Lavine, B. K. (Ed.). (2005). *Chemometrics and chemoinformatics*. Washington, DC: ACS.
- Leach, A. R., & Gillet, V. J. (2003). *An introduction to chemoinformatics*. Dordrecht: Springer.
- Lesk, A. M. (1977). Macromolecular marionettes. *Computers in Biology and Medicine*, 7(2), 113–129.
- Lindsay, R. K., Buchanan, B. G., Feigenbaum, E. A., & Lederberg, J. (1980). *Applications of artificial intelligence for organic chemistry; the DENDRAL project*. New York: McGraw-Hill.
- Lipinski, C., & Hopkins, A. (2004). Navigating chemical space for biology and medicine. *Nature*, 432(7019), 855–861.
- Littman, B. H., & Krishna, R. (Eds.). (2011). *Translational medicine and drug discovery*. New York: Cambridge University Press.
- Martel, S., Gillerat, F., Carosati, E., Maiarelli, D., Tetko, I. V., Mannhold, R., & Carrupt, P.-A. (2013). Large, chemically diverse dataset of logP measurements for benchmarking studies. *European Journal of Pharmaceutical Sciences*, 48(1-2), 21–29.
- Medina-Franco, J. L., Martinez-Mayorga, K., Giulianotti, M. A., Houghten, R. A., & Pinilla, C. (2008). Visualization of the chemical space in drug discovery. *Current Computer-Aided Drug Design*, 4(4), 322–333.
- Mezey, P. G. (1993). *Shape in chemistry an introduction to molecular shape and topology*. New York: VCH.
- Milano Chemometrics and QSAR Research Group (a). (2016). MOLE db – Molecular descriptors data base. http://micchem.disat.unimib.it/mole_db
- Milano Chemometrics and QSAR Research Group (b). (2016). Software and code. <http://micchem.disat.unimib.it/chm/download/software.htm>
- Molinspiration. (2016). Cheminformatics on the Web. www.molinspiration.com
- Morgan, H. L. (1965). Generation of a unique machine description for chemical structures—a technique developed at chemical abstracts service. *Journal of Chemical Documentation*, 5(2), 107–113.
- Musiol, R., Mrozek-Wilczkiewicz, A., & Polanski, J. (2014). Synergy against fungal pathogens: Working together is better than working alone. *Current Medicinal Chemistry*, 21(7), 870–893.
- Netzeva, T. I., Worth, A. P., Waldenberg, T., Benigni, R., Cronin, M. T. D., Gramatica, P., Jaworska, J. S., Kahn, S., Klopman, G., Marchant, C. A., Myatt, G., Nikolova-Jeliazkova, N., Patlewicz, G. Y., Perkins, R., Roberts, D. W., Schultz, T. W., Stanton, D. T., van de Sandt, J. J. M., Tong, W. D., Veith, G., & Yang, C. H. (2005). Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships – The report and recommendations of ECVAM Workshop 52. *Atla-Alternatives to Laboratory Animals*, 33(2), 155–173.
- Paris, G. (1999). Meeting of the American Chemical Society, Aug 1999.
- PDB, Protein Data Bank. (2016). <http://www.rcsb.org>
- Pliska, V., Testa, B., & Waterbeemd, H. (1996). *Lipophilicity in drug action and toxicology*. Weinheim: VCH.
- Polanski, J. (1997). The receptor-like neural network for modeling corticosteroid and testosterone binding globulins. *Journal of Chemical Information and Computer Sciences*, 37(3), 553–561.
- Polanski, J. (2003). Molecular shape analysis. In J. Gasteiger (Ed.), *Handbook of chemoinformatics* (pp. 302–319). Weinheim: Wiley-VCH Verlag.
- Polanski, J. (2009a). Chemoinformatics. In S. D. Brown, R. Tauler, & B. Walczak (Eds.), *Comprehensive chemometrics*. Amsterdam: Elsevier.
- Polanski, J. (2009b). Receptor dependent multidimensional QSAR for modeling drug – Receptor interactions. *Current Medicinal Chemistry*, 16(25), 3243–3257.
- Polanski, J., & Bak, A. (2003). Modeling steric and electronic effects in 3D-and 4D-QSAR schemes: Predicting benzoic pK(a) values and steroid CBG binding affinities. *Journal of Chemical Information and Computer Sciences*, 43(6), 2081–2092.
- Polanski, J., & Bak, A. (2010). *Introduction to chemoinformatics in drug design (in Polish)*. Katowice: WW.
- REACH, Registration, Evaluation, Authorisation and Restriction of Chemicals. (2016). <http://ec.europa.eu/growth/sectors/chemicals/>

- Reymond, J.-L., van Deursen, R., Blum, L. C., & Ruddigkeit, L. (2010). Chemical space as a source for new drugs. *Medicinal Chemistry Communications*, 1(1), 30–38.
- Sasaki, S. I., Abe, H., Outki, T., Sakamoto, M., & Ochiai, S. (1968). Automated structure elucidation of several kinds of aliphatic and alicyclic compounds. *Analytical Chemistry*, 40(14), 2220–2223.
- Schiff, H. (1875). Zur Statistik Chemischer Verbindungen. *Chemische Berichte*, 8, 1542–1547.
- Schreck, J. O. (1971). Nonlinear Hammett relationships. *Journal of Chemical Education*, 48(2), 103–107.
- Shelley, C. A., Hays, T. R., Munk, M. E., & Roman, R. V. (1978). Approach to automated partial structure expansion. *Analytica Chimica Acta-Computer Techniques and Optimization*, 2(2), 121–132.
- StatPages.org, Web Pages that Perform Statistical Calculations. <http://statpages.org/index.html>.
- Todeschini, R., & Consonni, V. (2000). *Handbook of molecular descriptors*. Weinheim: Wiley-VCH.
- VirtualToxLab. (2016). Application for a free OpenVirtualToxLab™ license. <http://www.biograf.ch/data/projects/OpenVirtualToxLab.php>
- Walters, W. P., Green, J., Weiss, J. R., & Murcko, M. A. (2011). What do medicinal chemists actually make? A 50-year retrospective. *Journal of Medicinal Chemistry*, 54(19), 6405–6416.
- Weinhold, F. (2001). Chemistry – A new twist on molecular shape. *Nature*, 411(6837), 539–541.
- Woody, A. I., Hendry, R. F., & Needham, P. (Eds.). (2012). *Philosophy of chemistry*. Amsterdam: Elsevier.
- Worachartcheewan, A., Mandi, P., Prachayasittikul, V., Toropova, A. P., Toropov, A. A., & Nantasenamat, C. (2014). Large-scale QSAR study of aromatase inhibitors using SMILES-based descriptors. *Chemometrics and Intelligent Laboratory Systems*, 138, 120–126.
- Yang, C., Tarkhov, A., Marusczyk, J., Bienfait, B., Gasteiger, J., Kleinoeder, T., Magdziarz, T., Sacher, O., Schwab, C. H., Schwoebel, J., Terfloth, L., Arvidson, K., Richard, A., Worth, A., & Rathman, J. (2015). New publicly available query language, CSRML, to support chemotype representations for application to data mining and modeling. *Journal of Chemical Information and Modeling*, 55(3), 510–528.
- Zartler, E. R., & Shapiro, M. J. (2005). Fragonomics: Fragment-based drug discovery. *Current Opinion in Chemical Biology*, 9, 366–370.
- Zupan, J., & Gasteiger, J. (1999). *Neural networks in chemistry and drug design second edition*. Weinheim: Wiley-VCH.

Natalia Sizochenko, Devashis Majumdar, Szczepan Roszak, and Jerzy Leszczynski

Contents

Introduction	2042
Quantum Chemical Parameters for Chemoinformatics	2044
Atomic Charges	2044
Other Charge-Related Properties	2045
Total and Orbital Energies and Related Properties	2049
Molecular Surface-Related Properties	2053
Quantum Chemical Techniques as Source of Parameter Generations	2057
Molecular Mechanics and QSAR	2059
Conclusions	2060
Bibliography	2060

Abstract

Quantum chemical and molecular mechanics-generated structure and reactivity parameters comprise a part of chemoinformatics, where such parameters are stored and properly indexed for search-information of a related molecule or a set of molecular systems. The present review makes a general survey of the various computable quantum chemical parameters for molecules. These could be used

N. Sizochenko • D. Majumdar (✉)

Department of Chemistry, Jackson State University, Jackson, MS, USA
e-mail: sizochenko@icnanotox.org; devashis@icnanotox.org

S. Roszak

Advanced Materials Engineering and Modelling Group, Faculty of Chemistry, Wrocław University of Technology, Wrocław, Poland
e-mail: szczepan.roszak@pwr.edu.pl

J. Leszczynski (✉)

Department of Chemistry and Biochemistry, Interdisciplinary Center for Nanotoxicity, Jackson State University, Jackson, MS, USA
e-mail: jerzy@icnanotox.org

for quantitative structure activity relation (QSAR) modeling. The applicability of various quantum chemical techniques for such property (QSAR parameters) is also discussed and density functional theory (DFT)-related techniques have been advocated to be quite useful for such purposes. Molecular mechanics methods, although mostly useful for less time consuming structure calculations and important in higher level molecular dynamics and Monte-Carlo simulations, are sometimes useful to generate structure-related descriptors for QSAR analysis. A brief discussion in this connection with molecular mechanics-related QSAR modeling is included to show the use of such descriptors.

Introduction

Chemoinformatics (also known as cheminformatics and chemical informatics) is the use of computer informational techniques applied to a range of problems in the field of chemistry. These *in silico* techniques are used mostly in drug discovery but can also be used in chemical and allied industries which include paper, pulp, dyes, and such other industries. The term chemoinformatics was first properly defined by K.F. Brown in 1998 (Brown 1998), and according to this definition *it is described as the mixing of information resources to transform data into information and information into knowledge for the intended purpose of making better and faster decisions in the area of drug lead identification and optimization*. This basic definition was later extended to the other allied chemical fields. It combines the scientific working fields of chemistry, computer science, and information science in the areas of topology, chemical graph theory, information retrieval, and data mining in the chemical space (Leach and Gillet 2007).

The main goal in chemoinformatics involves storage, indexing, and search of information of the related compounds. The related search topics include unstructured data mining (information retrieval and extraction), structured data mining (database, graph, molecule, sequence and tree mining), and digital libraries. The *in silico* representation of chemicals (involving structure and properties) are usually stored in large chemical databases. They also include visual representations in two and three dimensions for studying physical interactions, modeling, and docking. These chemical databases include virtual libraries (data for real and virtual molecules) and facilities for computationally screening *in silico* libraries of compounds (commonly known as virtual screening). The results are used for the well-known quantitative structure-activity relationship (QSAR) studies for the prediction of specific activity of compounds.

Since the inception of QSAR technique (Hansch 1969), electronic parameters from molecular structures were incorporated in the correlation through Hammett substituents constants (Hammett 1937; Hansch et al. 1991). It was possible since the molecules used were aromatic systems with substituents at meta- and para-positions. Hansch (1969) used the parameters generated through such analysis as measure of the electronic characteristics of molecules. An enormous number of QSAR equations have been reported in literature, many having functional form much more complicated than the original Hansch equation. Various parameters

have been used in QSAR equations in this respect, most of which were designed to represent the hydrophobic, electronic, and steric characteristics of molecules. Extensive tables have been published with values of these common parameters for a wide variety of substituents.

Modern state of the art quantum chemical techniques provide meaningful representation of structures and are source of related physical and electronic properties of molecules from their calculated wave functions. The usual electronic descriptors of molecules are atomic charges, dipole moment, electron affinity, ionization potential, and many other electronic properties related to orbital energies and molecular surfaces (e.g., molecular electrostatic potential (MEP)), and specific and thermodynamic parameters (related to the conformational properties and reactivities). These are of potential importance in QSAR analysis and are usually stored in chemoinformatics libraries as basic molecular information. Molecular mechanics (MM) calculations are also a useful tool for the structure calculations in a less time consuming way (with lower accuracy) but are not useful for generation of electronic parameters like quantum chemical methods. The main importance of various MM approaches actually lies in higher level molecular dynamics and Monte-Carlo simulations. Still in some specific cases, MM has sometimes been used to generate structure-related descriptors for QSAR methods.

A relation between the quantum chemically/MM derivable properties (usually stored in chemoinformatics libraries) and the related mathematical modeling are shown in Fig. 1. The mathematical models are usually of statistical origin and

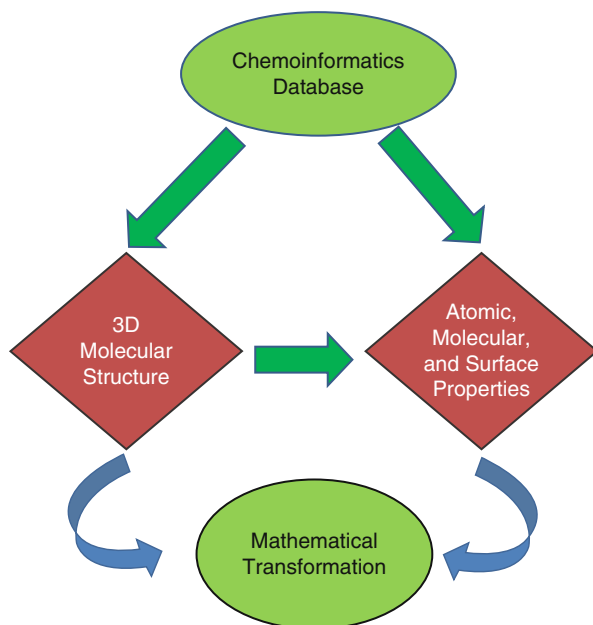


Fig. 1 The use of derived molecular properties (quantum chemical origin) for mathematical drug design is schematically shown in the flow chart. The properties are either from chemoinformatics data base or through direct calculations

correlate the observed biological activities with molecular descriptors to understand their specific properties. They are also useful to predict the properties of designed drug molecules within the statistical limitations of the derived models. In the present review, we will discuss the connection of such *in silico* databases in chemoinformatics used in drug discovery only. The present review would involve brief discussions on the usefulness of various quantum chemical parameters related to the biological properties of molecules (to be used in QSAR) and a brief résumé of the techniques in generating such quantum chemical parameters. In the last section of this review, a brief discussion has been made on the usefulness of MM-derived parameters for QSAR.

Quantum Chemical Parameters for Chemoinformatics

Atomic Charges

The quantum chemically computed atomic charges are one of the most fundamental parameters to be used in QSAR. This is because electrostatic interactions are central in controlling most of the chemical/biological reactions. The advantage of a quantum chemical technique is that it can compute the net charges on each atom in a molecule through gross population analysis. These parameters are also important to investigate surface charge characteristics of a molecule. There are various techniques to generate atomic charges through population analysis. The successful techniques amongst the oldest ones are Mulliken and Löwdin population analysis. In Mulliken's method (Mulliken 1955), the electron population (ρ_A) on an atom A of a molecule is defined in terms of one electron density matrix ($D_{\mu\nu}$) and overlap matrix ($S_{\mu\nu}$) as

$$\rho_A = \sum_{\mu \in A} \sum_v^{AO} D_{\mu\nu} S_{\mu\nu} \quad (1)$$

The diagonal term $D_{\mu\mu}S_{\mu\mu}$ represents the number of electrons in the μ^{th} atomic orbital (AO) and the off-diagonal $D_{\mu\nu}S_{\mu\nu}$ is (half) the number of electrons shared by the AOs $\mu\nu$ (there is an equivalent $D_{\nu\mu}S_{\nu\mu}$ element as the matrix is diagonal). The gross charge on atom A is defined as the sum of nuclear and electronic contributions.

$$Q_A = Z_A - \rho_A \quad (2)$$

The Löwdin technique (Löwdin 1970) uses $S^{1/2}D S^{1/2}$ matrix for population analysis. This is equivalent to a population analysis of the density matrix in the orthogonalized basis set by transforming the original set of functions by $S^{-1/2}$. The electron population associated with an atom is obtained through the equation:

$$Q_A = Z_A - \sum_{\substack{\mu=1 \\ \mu \text{ on } A}}^{AO} \left(S^{1/2} D S^{1/2} \right)_{\mu\mu} \quad (3)$$

The Mulliken and Löwdin population give different atomic charges, but mathematically there is nothing to indicate which partition gives better result. More realistic charge analysis, from a chemist's point of view, is obtained from the natural bond orbital (NBO) analysis by Weinhold and coworkers (Weinhold and Landis 2012). The method uses one-electron density matrix for defining the shape of atomic orbitals in molecular environment and molecular bonds from electron density between atoms. The density matrix D is defined in terms of blocks of basis functions belonging to a particular center. The natural atomic orbitals (NAO) in the molecular orbital are defined as those which diagonalize the original blocks in the D matrix. These NAOs are, in general, not orthogonal and are orthogonalized to achieve well-defined division in electrons. The natural atomic charges are computed as the difference of nuclear charge and the total natural population of NAOs (on the atoms). There are several other atomic charge calculation techniques, viz., electrostatic potential derived charges (Bayly et al. 1993) and populations derived from atoms in molecule (AIM) theory (Popelier 2000). They could be used in QSAR with varying successes, but the most recommended charges are the natural atomic charges through NBO analysis.

The atomic charges are usually used as the static chemical reactivity indices (Franke 1984). Since electrostatic interactions play important role in a chemical reaction, the orientations of the interacting moieties could be decided from the σ and π -electron densities and in such cases they could be used as directional reactivity indices. These orientation natures of the interacting moieties could, of course, be better interpreted from the molecular electrostatic potential (MEP) of the molecules. The atomic charges have been used to explain various other properties also. For example octanol-water partition coefficients of organic compounds (Karelson et al. 1996; Ghose et al. 1988), prediction of anti-HIV-1 activities of HEPT-analog compounds (Alves et al. 2000), hydrogen-bond donor strengths (Schwöbel et al. 2009), and atomic charge selectivity for optimal ligand-charge distribution at protein binding sites (Bhat et al. 2006) are a few to be mentioned. Atomic charges are also used as descriptors of molecular polarity but again charge derived properties like electrostatic moments would be a better choice.

Other Charge-Related Properties

These properties include frontier orbital densities, superdelocalizability, dipole moment, polarizability, and hyperpolarizability. The highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) are usually termed as frontier orbitals (Fukui 1971). These orbitals are the key factors in explaining many organic reactions. Specifically, pericyclic reactions (Woodward

and Hoffmann 1969) were explained from symmetries of the HOMO–LUMO orbitals. The frontier orbital densities were designed later as QSAR descriptor for electrophilic (f_r^E) and nucleophilic (f_r^N) centers in terms of HOMO ($C_{HOMO,n}$) and LUMO ($C_{LUMO,n}$) coefficients of the atomic orbital of the n th atom X and were defined as $\sum(C_{HOMO,n})^2$ and $\sum(C_{LUMO,n})^2$, respectively. The superdelocalizabilities and atom–atom polarizabilities are originated through Hückel theories and are not of much use in recent times in regular quantum chemistries, although they are sometimes used as reactivity descriptors in QSAR analysis (Karelson et al. 1996).

Dipole moment (μ) of a molecule could be used as a polarity descriptor in chromatographic observations of molecular polarity accounts for chromatographic retention on a polar stationary phase (Grunenberg and Herges 1995). Dipole moment of a molecule has several other importance, viz., it is an important parameter to account for the solvation properties of molecules (Milischuk and Matyushov 2002; Jensen 1999). In the case of directional orientation of molecules during interactions, the dipole–dipole interactions become quite important in total electrostatic interactions, if the components have non-zero dipole moments. The higher moments of the molecular systems also play important role in shaping the orientation as well as the nature of interactions in specific cases. The importance of higher moments is seen in cation- π interactions (Ma and Dougherty 1997). It is a very widely studied biological phenomena and the nature of interactions are believed to be controlled through electrostatics. Interactions through charge-dipole (C-D) and charge-quadrupole (C-Q) were believed to be the controlling features in such interactions. Recent studies on model cation- π interactions have shown (Kim et al. 1994; Kadlubanski et al. 2013) that octopole moments also have an important role in such interactions.

The model interactions were investigated for Mg^{2+} , Ca^{2+} , and NH_4^+ ions (acceptors) interacting with benzene, *p*-methylphenol, and 3-methylindole (donors) (Kadlubanski et al. 2013). The calculated strong electron sharing in the complexes showed that they might have considerable influence in shaping the contribution of total electrostatic interactions between the donors and acceptors. Figures 2 and 3 represent the natures of three of such multipolar components of the electrostatic interaction energies in Mg^{2+} , Ca^{2+} , and NH_4^+ complexes as a function of various interaction distances. The total electrostatic interaction energies (T-EI) together with charge-charge (C-C), dipole-charge (D-C), quadrupole-charge (Q-C), and octopole-charge (O-C) components were found to be important and presented in the figures. The influences of the other multipolar components were found to be negligible. The results show that the total interaction energy curves, with few exceptions, are either repulsive or dispersive in nature. The T-EI curves of Mg^{2+} complexes are bound (i.e. having a minimum) while the similar curves for the Ca^{2+} -complexes are dispersive for benzene and 3-methylindole (Fig. 2a, c). In this respect, the interactions of *p*-methylphenol... Ca^{2+} show a weak minimum (Fig. 2c). The nature of the curves representing the Q-C and O-C components of both the complexes as represented in Fig. 2b, d show that they have either repulsive or attractive contributions to the total electrostatic interactions in a significant way around the equilibrium distances and thus revealing the importance of such interactions in strong cation- π complexes.

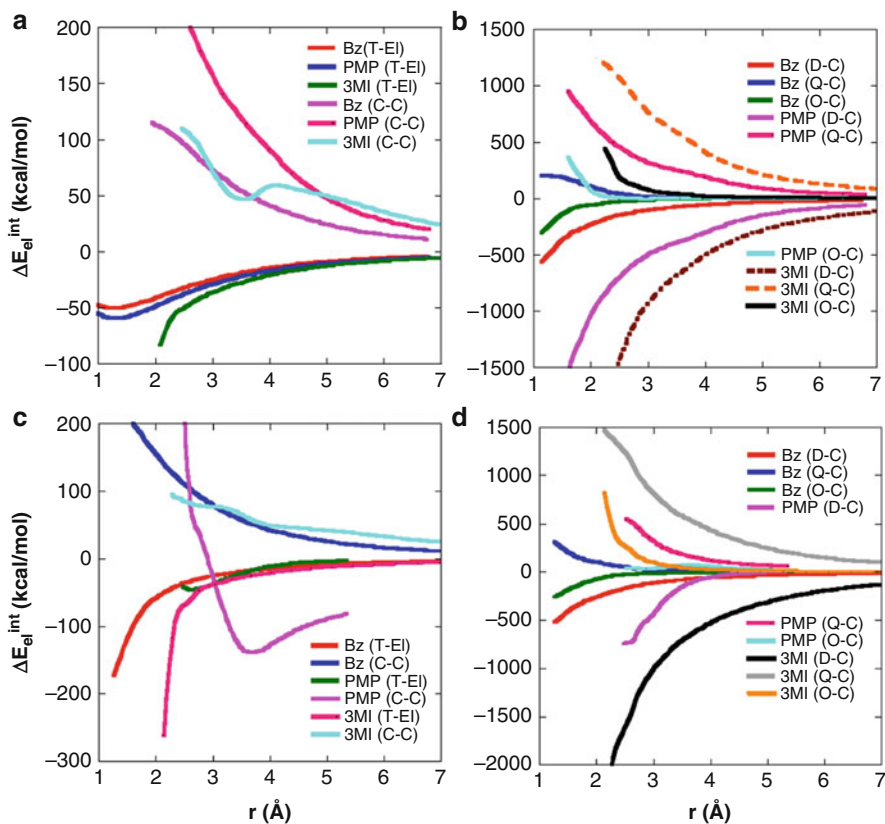


Fig. 2 Plots of total electrostatic interaction energies (*T-EI*) and their multipolar components as a function of *r* (interaction distance) for the complexes of Mg^{2+} and Ca^{2+} ions with benzene (*Bz*), *p*-methylphenol (*PMP*), and 3-methylindole (*3MI*). Panels (a) and (b) represent the curves for the Mg^{2+} complexes, while panels (c) and (d) are for the Ca^{2+} ion complexes. In the figures, C-C, D-C, Q-C, and O-C represent the multipolar electrostatic interaction energy components (C-C charge-charge, D-C dipole-charge, Q-C quadrupole-charge, O-C octopole-charge) (Reproduced from Kadlubanski et al. 2013, with the kind permission of the American Chemical Society 2013)

The D-C and (C-C) components of these metal complexes also behave similarly (Fig. 2a–d) to mostly offset the attractive components to generate overall shape of the T-EI curves.

The curves representing the electrostatic components of the cation- π complexes of NH_4^+ ion show an altogether different nature (Fig. 3). The D-C components are bound around the equilibrium interaction distances (*r*) (Kadlubanski et al. 2013). The rest of the components around the equilibrium *r* behave in such a way that the T-EI curves are dispersive, although such components sometimes show bound character at larger *r* (not of much significance here). The differences of the nature of the curves of NH_4^+ . . . aromatic complexes (Fig. 3) with respect to those of

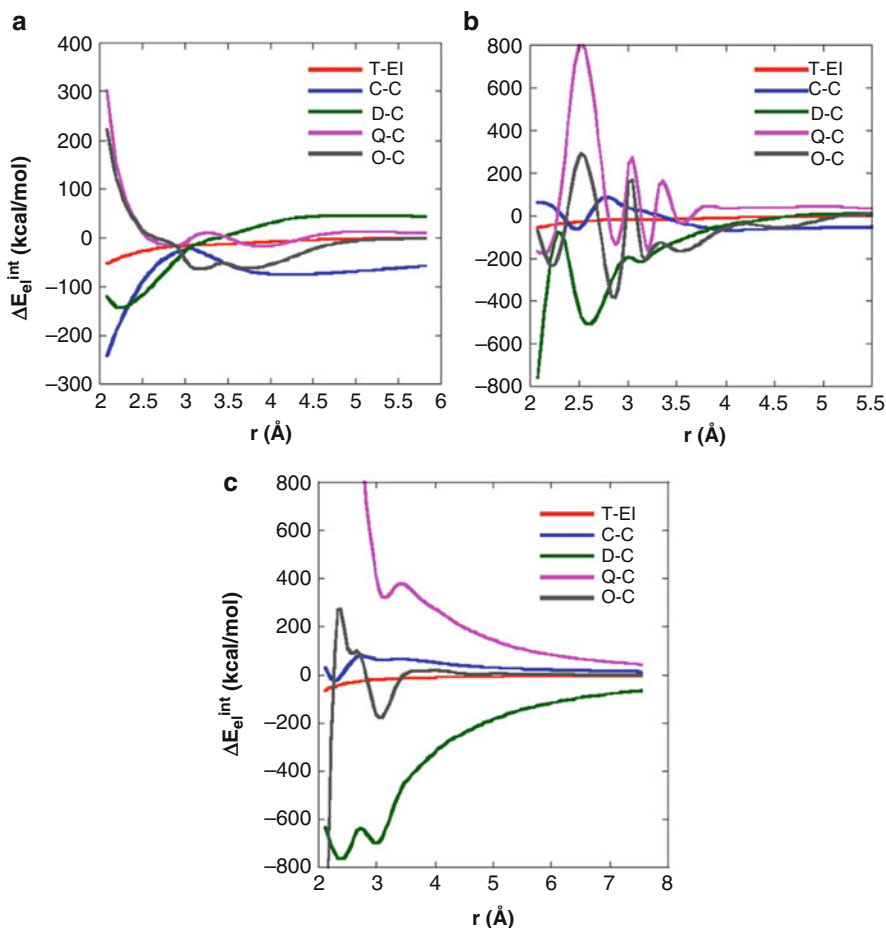


Fig. 3 Plots of total electrostatic interaction energies (*T-EI*) and their multipolar components as a function of interaction distances for the complexes of NH_4^+ ion with benzene (*Bz*) (panel **a**), *p*-methylphenol (*PMP*) (panel **b**), and 3-methylindole (*3MI*) (panel **c**). See Fig. 2 for the definition of the curves representing the cases of C-C, D-C, Q-C and O-C (Reproduced from Kadlubanski et al. 2013, with the kind permission of the American Chemical Society 2013)

the metal ion complexes could be due to the anisotropic potential expansion of the NH_4^+ ion. Similar studies were carried out using the dihydrated species of the metal cations (Kadlubanski et al. 2013). Although dihydrated metal ions are expected to have some anisotropic character in this respect, the curves showed they are not strong enough to make too much difference with respect to non-hydrated cases (where the metal ions are strictly isotropic). Thus it is evident from these curves that the multipolar components of electrostatic interactions play a major role in the formation as well as controlling the nature of interactions in the cation- π complexes. The rationale of this discussion is that although dipole moment is

common descriptor in QSAR studies, the higher moments should also be treated as effective descriptors in many important biological phenomena.

The quadrupole moment gives a rough approximation of the molecular volume, but this volume parameter can be observed through the effect of electric field on dipole moment (induced dipole moment) of a molecule. Generally the induced dipole moment μ could be written as:

$$\mu = \mu^0 + \alpha_{ij} V_j + \frac{1}{2} \beta_{ijk} V_j V_k + \frac{1}{6} \gamma_{ijkl} V_j V_k V_l \dots \quad (4)$$

The terms μ^0 and α denote dipole moment and polarizability, β and γ are the higher order polarizabilities. The $V_j, V_k \dots$ are the components of the applied field V . The term β and γ are not known to make contributions in QSAR studies, but the term α is an important descriptor. There are two such descriptors, isotropic and anisotropic α and they could be written as (Martin et al. 1979):

$$\alpha_{\text{isotropic}} = \frac{1}{3} (\alpha_{xx} + \alpha_{yy} + \alpha_{zz}) \quad (5)$$

and

$$\alpha_{\text{anisotropic}} = \frac{1}{\sqrt{2}} \left[(\alpha_{xx} - \alpha_{yy})^2 + (\alpha_{xx} - \alpha_{zz})^2 + (\alpha_{yy} - \alpha_{zz})^2 + 6(\alpha_{xy}^2 + \alpha_{xz}^2 + \alpha_{yz}^2) \right]^{1/2} \quad (6)$$

The $\alpha_{\text{isotropic}}$ is in a sense related to the molar volume (as it is expressed in terms of \AA^3) and has been found to be related to hydrophobicity and thus to other related biological properties (Hansch and Coats 1970). Furthermore, the electronic polarizability of molecules shares common features with the electrophilic superdelocalizability (Lewis 1987). This parameter also contains information regarding the inductive effects in a molecule (Gaudio et al. 1994). The anisotropic term is also expressed in volume units (\AA^3) and characterizes the properties of a molecule as an electron acceptor (Cartier and Rivail 1987).

Total and Orbital Energies and Related Properties

The total energy (E) of a molecule is generally used to account for several important molecular properties which include ionization potential (IP), electron affinity (EA), protonation energy, etc. The thermochemistry of a system is also important and the related parameters are computed in quantum chemical calculations as part of vibrational frequency calculations, as the partition functions needed for Gibbs free energy (G), enthalpy (H), and entropy (S) calculations are related to the molecular vibrations. The changes of these thermodynamic parameters, viz., ΔG , ΔH , ΔS , are important in accounting for experimental interaction energies between two

molecular systems, for heat of reactions, and to calculate reaction rates. The last parameter is not probably that important in QSAR, but the two others are quite important. The IP and EA of a molecule are computed as the difference of energy between the neutral and the corresponding cationic and anionic species, respectively. The HOMO (E_{HOMO}) and LUMO (E_{LUMO}) energies also account for the IP and EA of a molecule (Koopman's theorem), but the accuracies of such results are very low and are not used any more.

The important applications of E_{HOMO} and E_{LUMO} are to account for reactivity of molecules has already been discussed in connection with frontier electron densities. But the conception of hardness and softness are considered as more reliable descriptors for reactivities and they are also derived through E_{HOMO} and E_{LUMO} . Parr and Pearson provided the analytical definitions of hardness (η) and softness (S) as (Parr and Pearson 1983):

$$\eta = \frac{1}{2} \left(\frac{\partial^2 E}{\partial N^2} \right)_{v(\bar{r})} = \frac{1}{2} \left(\frac{\partial \mu}{\partial N} \right)_{v(\mathbf{r})} \quad (7)$$

$$S = \frac{1}{2} = \left(\frac{\partial N}{\partial \mu} \right)_{v(\mathbf{r})} \quad (8)$$

where E is the total energy, N is the number of electrons of the chemical species, and μ represents the chemical potential, which is identified as negative of the electronegativity (Parr and Pearson 1983), and $v(\mathbf{r})$ is the external electrostatic potential of an electron at \mathbf{r} due to the nuclei. The operational form of η and S could be written in terms of IP and EA as (Parr and Pearson 1983; Roy et al. 1998):

$$\eta = \frac{IP - EA}{2}; S = \frac{1}{IP - EA} \quad (9)$$

These definitions of hardness and softness lead to the hard and soft acid–base theory, which is quite useful to explain chemical reactions. According to this principle, all factors being equal, hard bases prefer to react with hard acids and soft bases prefers to react with soft acids, and atomic and molecular systems were classified as hard and soft systems based on equation (9). The concept was further extended to explain site-specific nucleophilic, electrophilic, and radical reactions through development of the concepts of local softness and hardness ($\eta_D(\mathbf{r})$ and Fukui functions ($f(\mathbf{r})$). The local softness ($s(\mathbf{r})$) is defined as (Roy et al. 1998):

$$s(\mathbf{r}) = f(\mathbf{r})S = \left(\frac{\partial \mu}{\partial v(\mathbf{r})} \right)_N S \quad (10)$$

Three types of $f(\mathbf{r})$ can be defined, which when multiplied by S generates three types of local softness, viz., s_k^+ , s_k^- , and s_k^0 for the nucleophilic, electrophilic, and radical attack, respectively, on the atom center k of a molecule.

$$\left. \begin{aligned} s_k^+ &= [\rho_k(N_0 + 1) - \rho_k(N_0)] S \\ s_k^- &= [\rho_k(N_0) - \rho_k(N_0 - 1)] S \\ s_k^0 &= \frac{1}{2} [\rho_k(N_0 + 1) - \rho_k(N_0 - 1)] S \end{aligned} \right\} \quad (11)$$

$\rho_k(N_0)$ represents electronic population on atom k for N_0 electron system. The local hardness ($\tilde{\eta}(\mathbf{r})$) is defined using electron density $\rho(\mathbf{r})$ as:

$$\tilde{\eta}(\mathbf{r}) = \left[\frac{\partial \mu}{\partial \rho(\mathbf{r})} \right]_{v(\mathbf{r})} \quad (12)$$

The local hardness is usually derived in working form ($\eta_D(\mathbf{r})$) from this formula using Thomas-Fermi-Dirac (TFD) approach to DFT using the electron density ρ as:

$$\eta_D(\mathbf{r}) = \frac{V_{el}(\mathbf{r})}{2N} \quad (13)$$

where $V_{el}(\mathbf{r})$ is the electronic part of the electrostatic potential. The electron density ρ is embedded in the explicit expression of $\eta_D(\mathbf{r})$ (Roy et al. 1998) and is not shown here.

The applicability of hard and soft acid–base (HSAB) principle together with their local counterparts (hardness and softness) has been discussed in details by Roos et al. (2009). The authors have discussed the use of these parameters in enzyme catalysis, which include phosphoenolpyruvate (the cofactor of several syntheses), the disulfide cascade mechanism in arsenate reductase, the preferential attack of cys10 on dianionic arsenate, mixed disulfide dissociation mechanism of thioredoxin, and hydride transfer in flavoenzymes. The local softness and electrophilicity parameters were found to be quite effective in explaining site-specific reactivity involved in these complex reactions. QSAR combined with DFT concepts have been applied in histone deacetylase (HDAC) inhibitors as they are very promising targets as anticancer drugs. The active site HDAC consists of Zn^{2+}

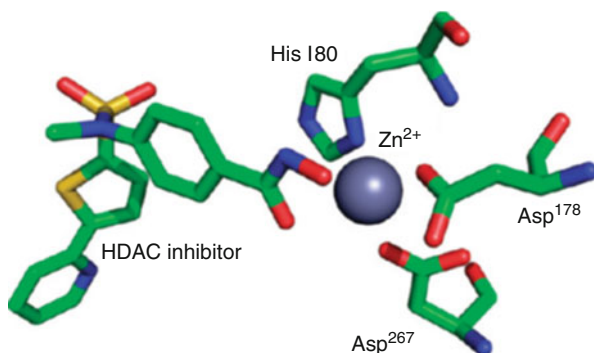


Fig. 4 Ribbon diagram of the active site of human histone deacetylase (HDAC8) (PDB: 1 W22) with a bound hydroxamic acid inhibitor (Reproduced from Roos et al. 2009, with the kind permission of the American Chemical Society 2009)

ion (Fig. 4). The general QSAR analysis in this respect was found to be qualitative in nature (Chattaraj et al. 2006). The introduction of electrophilicity concept in QSAR (Chattaraj et al. 2006) increased the interpretative character and chemical hardness was found to be useful descriptor in this respect, as the hardness of the zinc binding group of HDAC correlated well with the calculated relative interaction energies of these inhibitors (Roos et al. 2009). In a recent review (Schwöbel et al. 2011), the importance of chemical hardness descriptors has been discussed for the electrophilic reactivity in toxicity prediction. An electrophilic index ω_{el} ($= \mu^2/2\eta$) has been proposed to be important descriptor for the QSAR analysis of such reactions.

Hardness and electronegativity parameters have been further used to develop descriptors for hydrogen bond basicity and acidity for applications in QSAR. Two theoretical scales, quantifying the hydrogen bond ability of compounds, are proposed: hydrogen bond basicity, B^* , and hydrogen bond acidity, A^* (Oliferenko et al. 2004). Both of these scales are based on clear physical considerations and use molecular topology and orbital valence-state energetic characteristics as parameters. These parameters are quantitatively defined including the geometric features originating through molecular graphs. The total hydrogen bond basicity B^* is defined as:

$$B^* = \frac{\eta^{lp}}{\chi^{lp} n^*} \sum_i^D \frac{\chi_i^\sigma - \chi_1^{ref}}{\eta_i^\sigma d_i^2 n^*} \quad (14)$$

where χ^{lp} and η^{lp} are the lone pair electronegativity and corresponding hardness values, n^* is the effective principal quantum number, χ_1^{ref} is the reference electronegativity of the basic center, χ_i^σ , η_i^σ represent electronegativity and hardness of the σ -orbitals, and d_i is the topological distance for the i th substituent atom. The summation runs over all atoms positioned at particular topological distances from the basic center D , where D is the diameter (the shortest maximum length) of the corresponding molecular graph. The total hydrogen bond acidity A^* is defined using a similar set of parameters in the following way:

$$A^* = \frac{\chi_i^\sigma - \chi_1^{ref}}{\eta_i^\sigma D_{1H}^2} + \sum_{i=2}^N \frac{\chi_i^\sigma - \chi_i^{ref}}{\eta_i^\sigma d_{iH}^2} \quad (15)$$

where χ_i^σ and η_i^σ are the electronegativity and hardness of the sigma orbital of the acidic center and the i th atom, respectively, χ_i^{ref} is reference electronegativity of the i th atom, D_{1H} is the geometric distance between the hydrogen and the nearest atom, and d_{iH} is the topological distance between i th and the hydrogen atom. These new HB descriptors were found to significantly improve the performance of QSAR software and be useful in molecular modeling and drug design. The latter becomes possible if one considers A_i^* and B_i^* values of each atom in the molecule rather than the overall molecular features. Two sets of such atomic values, $\{A_i^*\}$ and

$\{B_i^*\}$, define two scalar potential fields of hydrogen bonding which can be used as molecular fields in comparative molecular field analysis (CoMFA) and/or scoring functions in ligand docking.

Molecular Surface-Related Properties

Molecular Electrostatic Potential

Molecular complementarity plays a central role in the molecular recognition process, and molecular electrostatic potential (MEP) is one of the oldest known properties of molecular systems to interpret such interactions. Since the introduction of the MEP by Scrocco et al. (Scrocco and Tomasi 1973), different investigators have applied this property for predicting biological recognition processes and it has been shown that an MEP profile leads to direct inferences about the nature of the corresponding binding site (Scrocco et al. 1973) and about interaction between these sites and an approaching molecule. The MEP of molecule $V(\vec{r})$ at a point \vec{r} is defined as:

$$V(\vec{r}) = \sum_A \frac{Z_A}{|\vec{R}_A - \vec{r}|} - \int \frac{\rho(\vec{r}_1)}{|\vec{r}_1 - \vec{r}|} d\vec{r} \quad (16)$$

The MEP has been used for molecular recognition process of drug molecules and to investigate the nature of long-range interactions with the corresponding drug-binding/receptor sites (Scrocco and Tomasi 1978). This approach was used to account for the binding properties of the neurotransmitter γ -aminobutyric acid (GABA) (Guha et al. 1992), cardiotonic drugs (Bhattacharjee et al. 1992), various nucleic acid bases (Weiner et al. 1982), and various other molecular systems (Scrocco and Tomasi 1978; Weiner et al. 1982). In recent times, MEP analysis have been found to be quite useful to interpret the difference in the nature of binding of acetylcholine (Ach) and various toxic nerve agents to the acetylcholinesterase (ACHE) binding sites (Majumdar et al. 2006). A typical example in this respect could be found in the interaction natures of Ach and R- and S-sarin enantiomers to the ACHE binding site (Fig. 5).

The ACHE(I) and ACHE(II) represents two forms of ACHE active cavity due to the two different orientations of the tautomeric His-440 moiety. The interactions in the ACHE(I) . . . Ach and ACHE(II) . . . Ach complexes (Fig. 5a, b) are dominated by the two-body electrostatic interactions between Tyr-121 residue and Ach. The orientation of the $-N^+(\text{CH}_3)_3$ group over the aromatic ring is ideal for cation- π interaction. However, it cannot be classified as purely cation- π , as MEP maps of Fig. 5a, b show that there is a possibility of electrostatic interactions of the $-N^+(\text{CH}_3)_3$ electron density with the side chain and the $-\text{OH}$ group of the Tyr-121 residue. This electrostatic effect together with the dispersion interactions makes this interaction more attractive than the normal cation- π interaction. The electrostatic interactions with the rest of the fragments are repulsive.

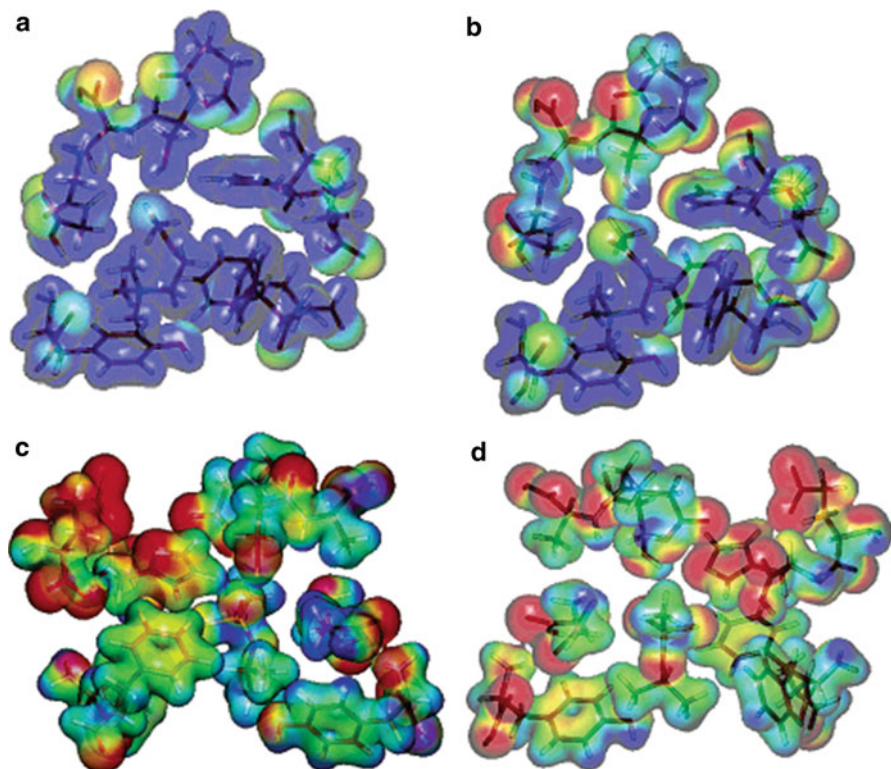


Fig. 5 Calculated MEP on the isodensity surfaces of the (a) ACHE(I)...Ach, (b) ACHE(II)...Ach, (c) ACHE(I)...(*S*)-sarin, and (d) ACHE(II)...(*S*)-sarin complexes. The various colored regions on the surfaces are *deep blue* (highly positive, >0.1 au), *light blue* (<0.1 au and >0.05 au), *green* (0.0 au), *yellow* (<-0.1 au and >-0.05 au), and *red* (>0.1 au). The minimum negative MEP values of the substrate are within the range -20 to -10 kcal/mol (*greenish yellow region*), while it is around -50 kcal/mol (*reddish yellow region*) for the inhibitor (Reproduced from Majumdar et al. 2006, with the kind permission of the American Chemical Society 2006)

The MEP map of the ACHE(I)...(*S*)-sarin complex (Fig. 5c) indicates strong hydrogen bonding interactions with the Ala-201, Ser-200, Glu-199 (X1), His-440, and Glu-327 (X3) fragments of ACHE (Majumdar et al. 2006). However, these are not pure hydrogen bonding interactions. The MEP isopotential surface of the ACHE(I)...(*S*)-sarin complex shows that the attractive interactions are also possible from the other parts of the X1-Ach and X3-Ach fragments. These effects make the two-body interactions more attractive than an ordinary hydrogen bond. The nature of interactions is totally changed in the ACHE-(II)...(*S*)-sarin complex (Fig. 5d). Since the interaction of P = O [of (*S*)-sarin] with the -N-H of the His-400 residue is lost in this case, the complex becomes a weakly bound system. The MEP map further shows that the electrostatic terms are quite important.

The MEP properties could be used as QSAR descriptor and important development in this direction was made by Carbo and coworkers (1980) and later by Good

et al. (Good and Richards 1996). These are concerned with molecular similarity and are used in 3D-QSAR modeling. The Carbo index for molecular similarity (C_{AB}) is defined as:

$$C_{AB} = \frac{\int P_A P_B dv}{\left(\int P_A^2 dv\right)^{1/2} \left(\int P_B^2 dv\right)^{1/2}} \quad (17)$$

where P_A and P_B are electron density dependent structural properties. The numerator measures property overlap while the denominator normalizes the similarity result. Initially, electron density was used as the structural property P , and later MEP was used to evaluate P (Eq. 18).

$$P_r = \sum_{i=1}^n \frac{q_i}{(r - R_i)} \quad (18)$$

where P_r is the electrostatic potential at point r associated with the charge q_i of the i th atom and the summation runs over the total n atoms. R_i is the nuclear coordinate of the atom i . In order to avoid singularities at the atomic nuclei, where $1/(r - R_i)$ tends to infinity, the MEP is normally determined outside the van der Waals (VDW) volume of the molecules in the calculation. The Carbo index was sensitive to the shape of property distribution rather than to its magnitude. Hodgkin and Richards (Good and Richards 1996) introduced the following modified index (H_{AB}) to increase the sensitivity of the formula to property magnitude:

$$H_{AB} = \frac{2 \int P_A P_B dv}{\int P_A^2 dv + \int P_B^2 dv} \quad (19)$$

When considering receptor drug binding, the electrostatic interaction energy of the system is proportional to the electrostatic potential exerted by the drug. The drug molecule similarity is related to the difference in binding energies (ignoring entropy and solvent effects), and this disparity is proportional to the difference in drug electrostatic potentials. Richards et al. (Good and Richards 1996) further developed two similarity indices based on this premise. They are known as exponential (E_{AB} , Eq. 20) and linear (L_{AB} , Eq. 21) indices.

$$E_{AB} = \frac{\sum_{i=1}^n \exp\left(-\frac{|P_A - P_B|}{\max(|P_A|, |P_B|)}\right)}{n} \quad (20)$$

$$L_{AB} = \frac{\sum_{i=1}^n \left(1 - \frac{|P_A - P_B|}{\max(|P_A|, |P_B|)} \right)}{n} \quad (21)$$

The term $\max(|P_A|, |P_B|)$ ($= P_{\max}$) equals larger MEP magnitude between P_A and P_B at grid point where similarity is being calculated. Although C_{AB} indices were used in QSAR studies, this was later replaced by the indices developed by Richard and coworkers. They were used to study dopamine agonistic properties as well as studies on hypoglycemic agents and np-apomorphine molecules and were employed in QSAR studies of steroids (Good and Richards 1996). These 3D-QSAR indices were also extended in graphical structure activity analysis using neural networks (Good and Richards 1996) and QSAR models from partial least squares [PLS] studies (Good and Richards 1996). Klebe et al. (1994) further developed MEP based comparative molecular similarity indices analysis technique for 3D-QSAR studies. The method was used to examine the correlations between calculated physicochemical properties and in vitro activities of a series of human immunodeficiency virus type 1 (HIV-1) integrase inhibitors (Makhija and Kulkarni 2001).

The more advanced studies of the MEP properties of molecules have been directed to macromolecular systems. In recent times, average electrostatic potential (ξ_k) have been computed using Markov model to study indirect interaction between amino acids placed at a topological distance (k) within a given protein backbone. A specific example in this context is the case of Arc repressor, which is a model protein of relevance for biochemical studies on bioorganics and medicinal chemistry. It was used to model the effect of alanine scanning on thermal stability (González-Díaz and Uriarte 2005). The potential ξ_k is defined as:

$$\xi_k = \sum_{j=1}^n p_k^A(j) \cdot V(j) \quad (22)$$

The potential ξ_k , in this definition, is dependent on the absolute probability $p_k^A(j)$ with which the amino acids interact with other amino acids placed at a distance k . The potential also depends on the initial unperturbed MEP $V(j)$ of the amino acid. The detailed mathematical analysis is available in González et al. (2005), and we do not need to elaborate it here further. Recently Murray, Politzer, and coworkers (Murray et al. 2014, 2015) have designed reactivity index of molecules based on their MEP surface properties and have applied to account for several interactions, but the applications of such indices as descriptor in QSAR are not yet available.

Atoms in Molecule (AIM) Theory Based Descriptors

The 3D molecular structure representation based on Bader's quantum topological Atoms in Molecules (AIM) theory for use in QSAR have been developed by Alsberg et al. (2000). Critical points located on the electron density distribution of the molecules are central to this structure representation using quantum topology

(StruQT). Other gradient fields such as the Laplacian of the electron density distribution can also be used. The type of critical point of particular interest is the bond critical point (BCP), which is characterized by using the three parameters, viz., electron density (ρ), the Laplacian ($\nabla^2\rho$), and the ellipticity (ϵ). This representation has the advantage that there is no need to probe a large number of lattice points in 3D space to capture the important parts of the 3D electronic structure as is necessary in, e.g., comparative field analysis (CoMFA). The details of AIM parameters are well documented (Popelier 2000) and need not be discussed here within the short span of this review. Interested readers are advised to go through the cited references for details.

Alsberg et al. (2000) used the structure representation through AIM to compute the wavelength of the lowest UV transition for a system of 18 anthocyanidins. Different QSAR models were also constructed using several chemo-metric/machine learning methods such as standard partial least squares regression (PLS), truncated PLS variable selection, genetic algorithm-based variable selection, and genetic programming (GP). These models identified bonds that either took part in decreasing or increasing the dominant excitation wavelength (Alsberg et al. 2000). The models also correctly emphasized on the involvement of the conjugated π system for predicting the wavelength through flagging the BCP ellipticity parameters as important for this particular data set. One advantage of AIM formalism is that it can be used to compute various electrostatic moment terms (up to octopole moment) and the MEP generated by molecular electron distribution can be expanded using multipole moments to represent the potential accurately. The AIM has further advantage of generating transferable atomic contribution of electrostatic potential in the total MEP of a molecule. This enables construction of potential maps of macromolecules (e.g., proteins) using the results of smaller fragments (Popelier 2000). The applications of AIM formalism, although promising, are still in developing stage and not used yet on regular basis.

Quantum Chemical Techniques as Source of Parameter Generations

The quantum chemical techniques to compute the parameters described above are quite large in number and still many new techniques are emerging to meet the challenges of breaking the molecular size barrier together with accuracy of the parameters. The guide to the accuracy of the computed parameters lies in the reproducibility of the experimental observables (e.g., dipole moment, vibrational spectra, binding energies, etc.) through the techniques used. In earlier days of QSAR, mostly Hückel's method was used and the molecular systems were restricted to planar π -systems (Hansch 1969; Hansch and Coats 1970; Hansch et al. 1991). The development of various semiempirical techniques broke this barrier from the mid-1970s through the development of complete neglect of differential overlap (CNDO) (Jensen 1999), intermediate neglect of differential overlap (INDO) (Jensen 1999), and modified intermediate neglect of differential overlap (MINDO group of

techniques) (Jensen 1999) of atomic orbitals in a molecule. These methods were developed through approximations (attached to the name of the techniques) from Hartree-Fock (HF) equations. The other variants CNDO/S and ZINDO (Jensen 1999) were not used unless spectral parameters were needed in QSAR. The later developments of these semiempirical techniques, viz., modified neglect of diatomic differential overlap (MNDO), its AM1 version, and the PM3-PM6 (Jensen 1999) group of techniques are still in use, especially PM6, for some specific cases. These methods were mostly generated from their earlier version of neglect of diatomic differential overlap (NDDO) technique (Jensen 1999). Most of the semiempirical techniques are not considered as the state of the art after the early 1990s due to their inability to reproduce experimental observables within accuracy limit. The systematic error embedded in the approximations used in such techniques is the responsible factor and development of various HF theory-based ab initio methods from the mid-1960s started to replace the semiempirical computations.

The simple HF-theory based ab initio techniques were later improved further to include high electron-correlations. These include Möller-Plesset second-order perturbation (MP2) and its higher versions (MP3, MP4), configuration interactions (CI), and multiconfiguration self-consistent field (MCSCF) techniques; their variants at the multireference levels; and various coupled cluster theory based methods (viz., CCSD, CCSD(T), and their multireference models) (Jensen 1999). These are highly accurate methods for computing molecular properties for both the ground and excited states but are enormously time consuming and need large computer memories even for small molecules. Thus their usage is not very much welcome in QSAR as such analysis needs substantial number of molecular data set for a specific property analysis of mostly small and mid-sized molecules. Sometimes the molecular systems are quite large and the properties are not easily extractable through such higher-order techniques.

The compromise between accuracy and molecular size could be resolved through the use of density functional theory (DFT) (Jensen 1999). The Kohn-Sham variational principle for DFT is quite old (Kohn and Sham 1965), but real computational approach of DFT in molecular quantum chemistry started in the mid-1980s. At present this is one of the cheapest and reliable techniques and quite a large number of functionals have been designed (and still under development) in this context to resolve diverse molecular problems. DFT could be used for fairly large molecular systems with affordable computational time and the technique is ideal for generating reliable parameters for QSAR. Several quantum chemical parameters for QSAR, discussed earlier, have either no semiempirical version or are not so accurate after computation. DFT technique is also very much desirable in this context. The semiempirical techniques are still directly used for very large molecular systems or systems where use of QM/MM (Jensen 1999) type of formalism is essential. The QM/MM technique, which divides a molecule/molecular system into two or three regions (called layers), e.g., methods like ONIOM (Svensson et al. 1996), mostly uses semiempirical formalism as second layer in the total computational strategy. This layer distribution is, of course, flexible and depends on the level of ab initio technique chosen for the first layer and the accuracy of computations needed for the

second and third layers. The final choice of the use of quantum chemical techniques in QSAR depends on the requirements for the model and, of course, user's intuition.

Molecular Mechanics and QSAR

The molecular mechanics (MM) as such is a tool for generating structure of molecules through much cheaper way than quantum chemical techniques using classical potentials. There are various forms of such potentials and they are formulated through transferable parameters called force-fields. The details of such potentials cannot be discussed in this short review, but the general views on it are well documented. The interested readers can refer to more recent books in this area (Schlick 2002; Warshel 1991). The most important aspect is that these potentials and related force-field parameters are the important essential step in designing strategies for large molecular structure simulations. The large structures are usually proteins, enzymes, DNA, and RNA like molecules and the simulation strategies involve classical molecular dynamics (MD) and Monte-Carlo (MC) techniques. The literature involving such calculation strategies as well as their use in large structure simulations, simulations involving their free energy surfaces (involving phase transition from one form to the other, free energy of interactions with large and small molecular units, etc.) have been accumulating for the last 50 years with various levels of sophistications. It is out of scope to discuss such topics here, but they are well documented (Schlick 2002) for interested readers. The MM calculations could also be used as an initial step to calculate the structures of larger species, like metal nanoclusters in desired symmetry (Majumdar et al. 2012).

The other important part of MM calculations is that the methods are sometimes used as a cheap way to generate structure dependent descriptor in QSAR analysis. Such descriptors could be generated from other sources also (Stanon et al. 2002), and we will wrap up this discussion by citing one important example of MM applications in this direction. The example is related to the generation of surface area of molecule from the van der Waals radii of atoms from the force-field data (MMFF94 force-field) after optimizing the geometry through MM calculations. For a given numerical property P_i for each atom i in a molecule, the descriptor $P_VSA(u, v)$ (VSA: van der Waals surface area) for a specific range $[u, v]$ was defined using sum of atomic VSA contribution of each atom i as (Labute 2000):

$$P_VSA(u, v) = \sum_i V_i \delta(P_i \in [u, v]) \quad (23)$$

where V_i is the atomic contribution of atom i to the VSA of the molecule. For a set of n descriptors associated with a property P ,

$$P_VSA_k = \sum_i V_i \delta(P_i \in [a_{k-a}, a_k]) \quad k = 1, 2, \dots, n \quad (24)$$

where $a_0 < a_k < a_n$ are interval boundaries such that $[a_0, a_n]$ bound all values of P_i in any molecule. The mathematical expression shows that VSA-type of descriptors correspond to a subdivision of the molecular surface area. The concept was used to construct QSAR/QSPR models for boiling point, vapor pressure, free energy of solvation in water, solubility in water, receptor class, and activity against thrombin, trypsin, and factor Xa (Labute 2000).

Conclusions

Quantum chemical parameters are getting more and more useful in QSAR as they are quite reliable because of the proper reproducibility of many key molecular observables with respect to experiments. In the present review, we have discussed various quantum chemical parameters involving atomic charges in molecules; electrostatic moments and their implications; orbital-related properties like hardness, softness, and their implications in reactivities; surface-related properties like MEP; and the related important QSAR indices. AIM properties are also discussed and examples have been cited regarding the usage of all these parameters in QSAR to model drug-related molecular properties. The sources of generating such quantum-chemical parameters are also discussed in terms of semiempirical and ab initio-related methods. DFT techniques have been suggested to be the most suitable ones after considering the pros and cons of these methods. The discussion on molecular mechanics (MM) also constitutes a part of the review and although it is mostly used to construct molecular structures, the structural parameters generated through MM are also sometimes used in QSAR/QSPR-related modeling. Such possibilities are also discussed in terms of van der Waals surface-related parameters. Various parameters for QSAR are still emerging through various quantum chemical/MM-related techniques, but ultimately it depends on the intuition of the user to generate meaningful structure-activity correlations for practical use in chemoinformatics.

Acknowledgments The authors acknowledge the support of NSF CREST (No.: HRD-0833178) grant. One of the authors (S.R.) acknowledges the financial support by a statutory activity subsidy from Polish Ministry of Science and Technology of Higher Education for the Faculty of Chemistry of Wrocław University of Science and Technology and NCN grant no UMO-2013/09/B/ST4/00097.

Bibliography

- Alsberg, B. K., Marchand-Geneste, N., & King, R. D. (2000). A new 3D molecular structure representation using quantum topology with application to structure–property relationships. *Chemometrics and Intelligent Laboratory Systems*, 54, 75.
- Alves, C. N., Pinheiro, J. C., Camargo, A. J., Ferreira, M. M. C., & da Silva, A. B. F. (2000). A structure–activity relationship study of HEPT-analog compounds with anti-HIV activity. *Journal of Molecular Structure (THEOCHEM)*, 530, 39.
- Bayly, C. I., Cieplak, P., Cornell, W., & Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *The Journal of Physical Chemistry*, 97, 10269.

- Bhat, S., Sulea, T., & Purisima, E. O. (2006). Coupled atomic charge selectivity for optimal ligand-charge distributions at protein binding sites. *Journal of Computational Chemistry*, 27, 1899.
- Bhattacharjee, A. K., Majumdar, D., & Guha, S. (1992). Theoretical studies on the conformational properties and pharmacophoric pattern of several bipyridine cardiotonics. *Journal of the Chemical Society, Perkin Transactions*, 2, 805.
- Brown, F. K. (1998). Chemoinformatics: What is it and how does it impact drug discovery. *Annual Reports in Medicinal Chemistry*, 33, 375.
- Carbó, R., Leyda, L., & Arnau, M. (1980). How similar is a molecule to another? An electron density measure of similarity between two molecular structures. *International Journal of Quantum Chemistry*, 17, 1185.
- Cartier, A., & Rivail, J. L. (1987). Electronic descriptors in quantitative structure – Activity relationships. *Chemometrics and Intelligent Laboratory Systems*, 1, 335.
- Chattaraj, P. K., Sarkar, U., & Roy, D. R. (2006). Electrophilicity index. *Chemical Reviews*, 106, 2065.
- Franke, R. (1984). *Theoretical drug design methods*. Amsterdam: Elsevier.
- Fukui, K. (1971). Recognition of stereochemical paths by orbital interaction. *Accounts of Chemical Research*, 4, 57.
- Gaudio, A. C., Korolkovas, A., & Takahata, Y. (1994). Quantitative structure-activity relationships for 1,4-dihydropyridine calcium channel antagonists (nifedipine analogues): A quantum chemical/classical approach. *Journal of Pharmaceutical Sciences*, 83, 1110.
- Ghose, A. K., Pritchett, A., & Crippen, G. M. (1988). Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships III: Modeling hydrophobic interactions. *Journal of Computational Chemistry*, 9, 80.
- González-Díaz, H., & Uriarte, E. (2005). Proteins QSAR with Markov average electrostatic potentials. *Bioorganic & Medicinal Chemistry Letters*, 15, 5088.
- Good, A. C., & Richards, W. G. (1996). The extension and application of molecular similarity calculations to drug design. *Drug Information Journal*, 30, 371.
- Grunenberg, J., & Herges, R. (1995). Prediction of chromatographic retention values (rm) and partition coefficients (log Poct) using a combination of semiempirical self-consistent reaction field calculations and neural networks. *Journal of Chemical Information and Computer Sciences*, 35, 905.
- Guha, S., Majumdar, D., & Bhattacharjee, A. K. (1992). Molecular electrostatic potential: A tool for the prediction of the pharmacophoric pattern of drug molecules. *Journal of Molecular Structure (THEOCHEM)*, 256, 61.
- Hammett, L. P. (1937). The effect of structure upon the reactions of organic compounds. Benzene derivatives. *Journal of the American Chemical Society*, 59, 96.
- Hansch, C. (1969). Quantitative approach to biochemical structure-activity relationships. *Accounts of Chemical Research*, 2, 232.
- Hansch, C., & Coats, E. (1970). α -chymotrypsin: A case study of substituent constants and regression analysis in enzymic structure – activity relationships. *Journal of Pharmaceutical Sciences*, 59, 731.
- Hansch, C., Leo, A., & Taft, R. W. (1991). A survey of Hammett substituent constants and resonance and field parameters. *Chemical Reviews*, 91, 165.
- Jensen, F. (1999). *Introduction to computational chemistry*. New York: Wiley.
- Kadlubanski, P., Calderón-Mojica, K., Rodríguez, W. A., Majumdar, D., Roszak, S., & Leszczynski, J. (2013). Role of the multipolar electrostatic interaction energy components in strong and weak cation – π interactions. *The Journal of Physical Chemistry. A*, 117, 7989.
- Karelson, M., Lobanov, V. S., & Katritzky, A. R. (1996). Quantum-chemical descriptors in QSAR/QSPR studies. *Chemical Reviews*, 96, 1027.
- Kim, K. S., Lee, J. Y., Lee, S. J., Ha, T.-K., & Kim, D. H. (1994). On binding forces between aromatic ring and quaternary ammonium compound. *Journal of the American Chemical Society*, 116, 7399.
- Klebe, G., Abraham, U., & Mietzner, T. (1994). Molecular similarity indices in a comparative analysis (comsia) of drug molecules to correlate and predict their biological activity. *Journal of Medicinal Chemistry*, 37, 4130.

- Kohn, W., & Sham, L. J. (1965). Self-consistent equations including exchange and correlation effects. *Physical Review*, *140*, A1133.
- Labute, P. (2000). A widely applicable set of descriptors. *Journal of Molecular Graphics and Modelling*, *18*, 464.
- Leach, A. R., & Gillet, V. J. (2007). *An introduction to chemoinformatics*. Dordrecht: Springer.
- Lewis, D. F. V. (1987). Molecular orbital calculations on solvents and other small molecules: Correlation between electronic and molecular properties v , α MOL, π^* , and β . *Journal of Computational Chemistry*, *8*, 1084.
- Löwdin, P.-O. (1970). On the nonorthogonality problem. *Advances in Quantum Chemistry*, *5*, 185.
- Ma, J. C., & Dougherty, D. A. (1997). The cation $-\pi$ interaction. *Chemical Reviews*, *97*, 1303.
- Majumdar, D., Roszak, S., & Leszczynski, J. (2006). Probing the acetylcholinesterase inhibition of sarin: A comparative interaction study of the inhibitor and acetylcholine with a model enzyme cavity. *The Journal of Physical Chemistry. B*, *110*, 13597.
- Majumdar, D., Roszak, S., & Leszczynski, J. (2012). Theoretical studies on the structure and electronic properties of cubic gold nanoclusters. *The Canadian Journal of Chemical Engineering*, *90*, 852.
- Makhija, M., & Kulkarni, V. (2001). Molecular electrostatic potentials as input for the alignment of HIV-1 integrase inhibitors in 3D QSAR. *Journal of Computer-Aided Molecular Design*, *15*, 961.
- Martin, R. L., Davidson, E. R., & Eggers, D. F. (1979). Ab initio theory of the polarizability and polarizability derivatives in H_2S . *Chemical Physics*, *38*, 341.
- Milischuk, A., & Matyushov, D. V. (2002). Dipole solvation: Nonlinear effects, density reorganization, and the breakdown of the onsager saturation limit. *The Journal of Physical Chemistry. A*, *106*, 2146.
- Mulliken, R. S. (1955). Electronic population analysis on LCAO–MO molecular wave functions I. *The Journal of Chemical Physics*, *23*, 1833.
- Murray, J. S., Macaveiu, L., & Politzer, P. (2014). Factors affecting the strengths of σ -hole electrostatic potentials. *Journal of Computational Science*, *5*, 590.
- Murray, J. S., Shields, Z. P.-I., Seybold, P. G., & Politzer, P. (2015). Intuitive and counterintuitive noncovalent interactions of aromatic π regions with the hydrogen and the nitrogen of HCN. *Journal of Computational Science*, *10*, 209.
- Oliferenko, A. A., Oliferenko, P. V., Huddleston, J. G., Rogers, R. D., Palyulin, V. A., Zefirov, N. S., & Katritzky, A. R. (2004). Theoretical scales of hydrogen bond acidity and basicity for application in qsar/qspr studies and drug design. Partitioning of aliphatic compounds. *Journal of Chemical Information and Computer Sciences*, *44*, 1042.
- Parr, R. G., & Pearson, R. G. (1983). Absolute hardness: Companion parameter to absolute electronegativity. *Journal of the American Chemical Society*, *105*, 7512.
- Popelier, P. L. A. (2000). *Atoms in molecules: An introduction*. New York: Prentice Hall.
- Roos, G., Geerlings, P., & Messens, J. (2009). Enzymatic catalysis: The emerging role of conceptual density functional theory. *The Journal of Physical Chemistry. B*, *113*, 13465.
- Roy, R. K., Krishnamurti, S., Geerlings, P., & Pal, S. (1998). Local softness and hardness based reactivity descriptors for predicting intra- and intermolecular reactivity sequences: Carbonyl compounds. *The Journal of Physical Chemistry. A*, *102*, 3746.
- Schlick, T. (2002). *Molecular modeling and simulation: An interdisciplinary guide*. New York: Springer.
- Schwöbel, J., Ebert, R.-U., Kühne, R., & Schüürmann, G. (2009). Modeling the H bond donor strength of -OH, -NH, and -CH sites by local molecular parameters. *Journal of Computational Chemistry*, *30*, 1454.
- Schwöbel, J. A. H., Koleva, Y. K., Enoch, S. J., Bajot, F., Hewitt, M., Madden, J. C., Roberts, D. W., Schultz, T. W., & Cronin, M. T. D. (2011). Measurement and estimation of electrophilic reactivity for predictive toxicology. *Chemical Reviews*, *111*, 2562.
- Scrocco, E., & Tomasi, J. (1973). The electrostatic molecular potential as a tool for the interpretation of molecular properties. *Topics in Current Chemistry*, *42*, 95.

- Scrocco, E., & Tomasi, J. (1978). Electronic molecular structure, reactivity and intermolecular forces: An euristic interpretation by means of electrostatic molecular potentials. *Advances in Quantum Chemistry*, *11*, 115.
- Stanon, D. L., Dimitrov, S., Gruncharov, V., & Mekenyan, O. G. (2002). Charged partial surface area (CPSA) descriptors QSAR applications. *SAR and QSAR in Environmental Research*, *13*, 341.
- Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S., & Morokuma, K. (1996). ONIOM: A multilayered integrated mo + mm method for geometry optimizations and single point energy predictions. a test for diels – alder reactions and $\text{Pt}(\text{P}(\text{t-Bu})_3)_2 + \text{H}_2$ oxidative addition. *The Journal of Physical Chemistry*, *100*, 19357.
- Warshel, A. (1991). *Computer modeling of chemical reactions in enzymes and solutions*. New York: Wiley.
- Weiner, P. K., Langridge, R., Blaney, J. M., Schaefer, R., & Kollman, P. A. (1982). Electrostatic potential molecular surfaces. *Proceedings of the National Academy of Sciences of the United States of America*, *79*, 3754.
- Weinhold, F., & Landis, C. R. (2012). *Discovering chemistry with natural bond orbitals*. Hoboken: Wiley.
- Woodward, R. B., & Hoffmann, R. (1969). The conservation of orbital symmetry. *Angewandte Chemie International Edition in English*, *8*, 781.

Andrea Mauri, Viviana Consonni, and Roberto Todeschini

Contents

Introduction	2066
How to Represent Molecules	2068
How to Codify Chemical Structures	2069
Molecular Structure Curation	2069
Constitutional Descriptors	2070
Topological Indices (TIs)	2071
Molecular Graph	2071
Matrix-Based Representation	2072
Generalized Topological Indices	2076
Autocorrelation Descriptors	2079
Geometrical Descriptors	2082
Structural Keys and Molecular Fingerprints	2086
Future Perspectives	2087
Molecular Descriptors for Disconnected Structures	2088
Molecular Descriptors for Nanostructures	2089
Conclusions	2089
Bibliography	2090

Abstract

Despite the number of available chemicals growing exponentially, testing of their toxicological and environmental behavior is often a critical issue and alternative strategies are required. Additionally, there is the need to predict properties of not yet synthesized compounds to reduce the costs of synthesis, selecting only those that have the maximal potential to be active and nontoxic compounds. In order to evaluate chemical properties avoiding chemical synthesis and reducing expensive and time-demanding laboratory testing, it is necessary to build *in silico* models

A. Mauri (✉) • V. Consonni • R. Todeschini

Department of Earth and Environmental Sciences, University of Milano-Bicocca, Milan, Italy

e-mail: andrea.mauri@unimib.it; viviana.consonni@unimib.it; roberto.todeschini@unimib.it

establishing a mathematical relationship between the structures of molecules and the considered properties (quantitative structure–activity relationships, QSARs). Molecular descriptors play a fundamental role in QSAR and other *in silico* models since they formally are the numerical representation of a molecular structure. Molecular descriptors can be classified using different criteria. Among them, there are two main categories, experimental and theoretical descriptors. The basis to understand and perform molecular descriptor calculation, the different theoretical descriptor categories together with their perspectives are described in this chapter.

Introduction

The overall number of chemicals synthesized, isolated, or marketed is increasing exponentially. The CAS REGISTRYSM database (www.cas.org) currently contains more than 93 million unique organic and inorganic chemical substances; it is updated daily and approximately includes 15,000 new substances each day.

Combinatorial chemistry and dynamic combinatorial chemistry (DCC) (Lehn 1999; Corbett et al. 2006) allow the synthesis of a large number of compounds in a single process, fastening the synthesis of different molecules with predefined characteristics.

At the same time, techniques like molecule structure generation and computer-assisted structure elucidation (CASE) of chemical compounds (Kind and Fiehn 2010; Elyashberg et al. 2011) give the possibility to design virtual molecules using *in silico* methods and potentially allow researchers to design millions of molecules in seconds.

Unfortunately, testing of physicochemical, toxicological, environmental, and other properties of chemicals is growing more slowly, hence, there is the need to have tools able to estimate chemical properties without performing experimental tests.

Quantitative structure–activity relationships (QSARs) and the similar quantitative structure–property relationships (QSPRs) (Hansch et al. 1996) are chemoinformatic techniques that allow establishing a relationship between the molecular structure and a defined endpoint. In order to apply these techniques, it is necessary to define, from a chemoinformatic point of view, what a molecular structure is. The molecular structure is a set of parameters able to describe a molecule in quantitative terms. Unfortunately, this set of parameters is not unique, that is, different mathematical representations of a molecular structure are available. Depending on the selected molecular representation, we can describe a molecule by means of experimental descriptors or theoretical descriptors. Experimental descriptors are all the experimental measurements such as octanol–water partition coefficient, molar refractivity, polarizability, and, in general, any physicochemical property obtained applying a well-specified experimental procedure. On the contrary, theoretical molecular descriptors are those descriptors obtained by a well-specified chemoinformatic algorithm applied to an unequivocal molecular representation.

Research on molecular descriptor is growing fast during the last years; the major drives that foster this development are essentially two. The first one is the rapid growth of number of available chemicals together with freely accessible databases collecting experimental data (Irwin and Shoichet 2005; Bolton et al. 2008; Pence and Williams 2010; Williams and Tkachenko 2014). Data availability changes the approaches to QSAR models (Cherkasov et al. 2014). While the first QSARs were based on few congeneric compounds possibly having the same mode of action (MoA) (Bobra et al. 1985; Vighi and Calamari 1985), the increased data availability enables researchers to develop more complex and global models based on hundreds of heterogeneous chemicals (Mannhold et al. 2009; Kühne et al. 2013; Cassotti et al. 2014a; Gissi et al. 2015).

The second stimulus comes from institutions. REACH regulation on the registration, evaluation, authorization, and restriction of chemicals in Europe (Worth 2010) promotes the use of QSAR methodologies to predict not available experimental data. There are analogous perspectives in the US EPA (Zeeman et al. 1995) and governmental agencies including Health Canada (Hughes et al. 2009).

In this scenario, molecular descriptors play a crucial role. Although until the 1960s–1970s, the study of relationships between experimentally measured quantities and molecular structure was based on experimental descriptors; nowadays, this analysis is more focused on the recognition of the relationships among theoretical descriptors and the considered experimental property. Thousands of molecular descriptors have been proposed in literature (Todeschini and Consonni 2009); this impressive amount of work is just one of the evidences of the scientific community interest in this field.

The huge amount of available molecular descriptors requires the delineation of few basic rules that molecular descriptors should comply with (Randić 1996; Guha and Willighagen 2012).

A molecular descriptor must:

1. Be invariant to atom labeling and numbering
2. Be invariant to the molecule roto-translation
3. Be defined by an unambiguous algorithm
4. Have a well-defined applicability on molecular structures

Additionally, in order to be potentially useful, a molecular descriptor should meet the following requirements:

5. Should have structural interpretation
6. Should have a good correlation with at least one experimental property
7. Should not have trivial relation with other molecular descriptors
8. Should not be based on experimental properties
9. Should preferably be continuous
10. Should preferably show minimal degeneracy
11. Should preferably be simple
12. Should preferably be applicable to a broad class of molecules

13. Should preferably be able to discriminate among isomers
14. Should preferably have calculated values in a suitable numerical range for the set of molecules where it is applicable to

Basically, the first rules can be used to understand if a molecular descriptor is well defined, but they do not permit to realize if a descriptor is useful or not to predict a defined property. Molecular descriptors should be invariant to all the operations related to molecule manipulation, which do not affect the molecular structure. Additionally, a molecular descriptor should have an applicability domain; it means that it should be clear to which class of molecules can be applied. Most of the first proposed molecular descriptors have been thought of for small organic molecules. Recently, one of the major challenges that molecular descriptors should face with is the possibility to be applied on different classes of compounds like salts, ionic liquids, peptides, polymers, and nanostructures.

The second set of rules are more related to the usefulness of a descriptor that should be interpretable, possibly related to an experimental property, but not excessively related to already existing molecular descriptors. A molecular descriptor should be a continuous variable and show a low degeneracy; it means that a molecular descriptor should be able to account for variations in molecular structure, even if variation is small.

Regardless of the molecular descriptor definition, the information gatherable by a molecular structure depends on the selected molecular representation.

How to Represent Molecules

Chemical compounds can be represented in different ways using different rules and criteria. Depending on the molecule representation, different information, codified in molecular descriptors, can be collected (Testa and Kier 1991; Jurs et al. 1995).

The simplest molecule representation is the chemical formula that codifies the atom types and their occurrences within a molecule. Chemical formula does not carry any information about atom connections, and molecular descriptors obtained by this kind of representation are referred to as 0D molecular descriptors.

The substructure list consists of a partial list of structural fragments included in a molecule. Molecular descriptors derived from this representation are usually cited as 1D molecular descriptors. They cannot represent the whole topology of the chemical structure, can be easily calculated and interpreted, and are typically used for similarity/diversity analysis and virtual screening of large chemical databases.

The topological representation is the most popular one. This is a two-dimensional representation that, in addition to the atomic composition, includes the information about connectivity of atoms in the molecule in terms of a molecular graph. Molecular descriptors derived from two-dimensional representation of molecule are referred to as 2D molecular descriptors.

3D molecular descriptors derive the molecular representation where the spatial information about atom positions is added to atom connectivities; in other words,

every atom of the molecule is associated with a set of three-dimensional coordinates determining its own position in the three-dimensional space. This representation is referred to as geometrical representation of a molecule where the molecule is considered as a rigid geometrical structure.

Additional representations of molecular structures have been defined. For example, 4D representation takes also into account molecular properties deriving from electron distribution and the interactions of the molecules with probes.

How to Codify Chemical Structures

Molecular structure information is usually stored in the so-called chemical files. Despite the huge number of chemical file formats (Boyle et al. 2011), all of them represent molecular structure focusing on specific atomic or molecular information and cannot store all possible chemical data.

With regard to the topic of this chapter, two main types of chemical files can be distinguished, one concerning the topological representation and the other the geometrical representation of molecules.

Typically, the former includes all types of linear notations; among them the most used is the SMILES file format (Weininger 1988). The latter includes all chemical file formats that can store the three-dimensional coordinates of the atoms, such as MDL and SYBYL file formats. An attempt to provide a standard format to store chemical data based on an XML dialect is the Chemical Markup language (CML) (Murray-Rust 1999; Murray-Rust and Rzepa 2001).

Molecular Structure Curation

The fundamental assumption at the basis of molecular descriptor calculation is the correctness of the molecular structures on which calculation algorithms are applied. Unfortunately, chemical structure annotation, both in chemistry publications and in public and commercial databases, is not exempt from errors.

Molecular structure curation has been severely emphasized in the last years in many publications highlighting problems and consequences related to the use of incorrect data in chemoinformatic and QSAR analysis (Young et al. 2008; Tropsha 2010). Data curation is usually composed of different steps to be applied to a whole dataset. Chemical structure curation has been recently detailed and has been split in five main steps (Fouches et al. 2010):

1. Removal of chemical structures that cannot be treated by the considered application or software, typically inorganics, mixtures, and eventually organometallics
2. Structural conversion and cleaning
3. Standardization and normalization of specific chemotypes
4. Removal of duplicates
5. Final manual checking

In this chapter, we will focus on second, third, and fifth steps. The second step is related to the preparation of the dataset. To calculate molecular descriptors, chemical structures must be encoded into a specific computational format, such as SMILES, SYBYL, or MDL. This conversion can be performed drawing the molecular structure by a molecule editor or directly querying a chemical database. Once this operation has been done, a consistency check must be performed on the generated molecular files finding out not realistic or incorrect molecular structures.

The third step of the curation procedure takes into account the fact that a molecule, or even a single fragment, can be represented in different ways. For example, aromatic rings can be represented either in Kekulé (i.e., bond configuration double–single–double) or aromatic form (i.e., conjugated bonds); since the representation is different, in principle it can affect the molecular descriptor values even if the chemical information is identical. Also some functional groups can have alternative representations; this is the case of nitro groups, which have multiple mesomers and can be represented using different bond notations (e.g., neutral and charged). The choice of one form with respect to another can affect significantly the performance of the models built with such data. Much more important is that the same normalization procedure is applied to all the molecular structures regardless of their specific representation.

The last step concerns the final check of all the codified molecular structures, which has to be done manually and, consequently, could be tedious and time demanding especially for big datasets. Nevertheless, it is crucial in order to identify issues not properly handled by automatic curation and to fix, if possible, the problems raised during the aforementioned steps.

Constitutional Descriptors

Constitutional descriptors are the simplest molecular descriptors that can be calculated starting from a molecular structure. Constitutional descriptors comprise all those descriptors representing a molecular structure, which take into account only chemical composition and do not encode information about the overall topology and geometry.

Typical constitutional descriptors are the molecular weight, the number of atoms and bonds, absolute and relative occurrence frequency of specific atom and bond types, and ring descriptors obtained considering only molecular structure composition.

The cyclomatic number (c), also known as circuit rank, is the most known ring descriptor and it is defined as:

$$c = |E| - |V| + D$$

where $|E|$ is the number of bonds, $|V|$ is the number of atoms, and D is the number of disconnected fragments in the molecular structure. The cyclomatic number is the minimum number of edges to be removed from a molecular graph in order to

remove all its cycles, making it into a forest, that is, an acyclic graph. It is widely used in the description of ring systems and gives the cardinality of the Smallest Set of Smallest Rings (SSSR).

Topological Indices (TIs)

Topological indices are two-dimensional descriptors derived from the topological representation of molecular structures (i.e., molecular graph); no information about spatial distribution of the atoms is considered in the calculation of these descriptors.

There exist different categories of topological indices; some of them will be described in the following paragraphs.

Molecular Graph

The typical 2D representation of a molecule is the molecular graph. A graph, $G = (V, E)$, is a representation of a set V of objects, called vertices, and a set E of links between vertices, called edges (Cormen et al. 2001). The molecular graph is a graph where the atoms are the vertices and the chemical bonds between atom pairs are the edges.

The molecular graph is typically an undirected, weighted, sparse graph. A graph is defined as undirected if the edges have no orientation and is defined as weighted if a number (i.e., the weight) is assigned to each edge. A sparse graph is a graph for which the number of edges $|E|$ is much less than the square of the number of vertices $|V|^2$. The weights associated to edges are typically the bond orders.

Application of graph theory on molecular structure (Balaban 1985) permits to handle chemical graphs adopting well-defined algorithms useful to explore structural properties of the molecules. Typically, graph theory can be used to extract information from the molecular graph, such as the calculation of matrices depending on graph connections and the definition of graph invariants that are those properties that depend only on the abstract structure of the graphs and do not depend from graph representations. Graph invariants are numerical representations of a graph; they can be a single number, a sequence of numbers, or a characteristic polynomial derived from the graph by a well-defined numerical manipulation of the graph structure. Graph invariants can be used for fast recognition of graph isomorphism since, for any invariant, two graphs with different values cannot be isomorphic. Anyway, two graphs with the same invariant value may or may not be isomorphic; a graph invariant is complete if it has equal values only for isomorphic graphs.

Graph invariant theory is the basis for the definition of topological indices.

A graph is commonly represented in two standard ways, as a collection of adjacency lists or as an adjacency matrix.

The adjacency lists are a collection of $|V|$ lists, one for each atom i , where every list $Adj[i]$ includes the atoms connected to the i -th atom together with the corresponding bond orders.

The adjacency matrix \mathbf{A} is essentially a $|\mathbf{V}| \times |\mathbf{V}|$ symmetric square matrix where the element a_{ij} is equal to one if atoms i and j are adjacent, zero otherwise:

$$[\mathbf{A}]_{ij} \equiv a_{ij} = \begin{cases} 1 & \text{if } (i, j) \in \mathbf{E} \\ 0 & \text{otherwise} \end{cases}$$

Atoms are adjacent if there exists a chemical bond connecting them.

Matrix-Based Representation

Adjacency matrix is just one of the possible matrices that can encode information about a molecular graph. Many matrices can be calculated from a graph extracting different information. These matrices, known as the graph-theoretical matrices, are the starting point to derive lots of molecular descriptors depending on the information included in the considered matrix.

Graph-theoretical matrices can be classified as vertex matrices, edge matrices, or incidence matrices. Vertex matrices are those matrices whose rows and columns refer to graph vertices, the atoms, and each element of the matrix encodes a property associated to a pair of vertices. Vertex matrices are square matrices whose dimension is $|\mathbf{V}| \times |\mathbf{V}|$. Edge matrices are $|\mathbf{E}| \times |\mathbf{E}|$ square matrices whose elements encode information on pair of edges, the bonds. Incidence matrices contain information about the relationships between two different sets of objects, such as atoms, bonds, cycles, substructures, or molecular paths. Usually incidence matrices are not square matrices and their dimension depends on the considered objects.

Most of the graph-theoretical matrices are calculated on a graph that does not include the hydrogen atoms, the so-called H-depleted molecular graph; however, they can be also derived from the H-filled molecular graph, if hydrogens are required to better represent the molecular structure.

Despite a large number of graph-theoretical matrices have been proposed in the literature (Ivanciuc and Ivanciuc 2000; Todeschini and Consonni 2009), only some of them along with the corresponding molecular descriptors will be here described with some details.

Vertex matrices are more used to derive molecular descriptors than edge and incidence matrices. Among them, the most important are the adjacency matrix and the topological distance matrix.

One of the first molecular descriptors derived from the adjacency matrix \mathbf{A} is the Lovasz–Pelikan index, which is defined as the leading eigenvalue (i.e., the largest eigenvalue) of the adjacency matrix (Lovasz and Pelikan 1973):

$$\lambda_1^{LP} \equiv \max_i \{\lambda_i\}$$

The topological distance matrix \mathbf{D} is a symmetric square $|\mathbf{V}| \times |\mathbf{V}|$ matrix whose element d_{ij} is equal to the topological distance between vertices i and j :

$$[\mathbf{D}]_{ij} \equiv d_{ij} = \begin{cases} |\min p_{ij}| & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}$$

where $|\min p_{ij}|$ is the number of edges along the shortest path between i and j .

One of the most known topological indices is the Wiener index (W), which is defined as the half-sum of all the elements d_{ij} of the distance matrix (Wiener 1947):

$$W = \frac{1}{2} \cdot \sum_{i=1}^{|\mathbf{V}|} \sum_{j=1}^{|\mathbf{V}|} d_{ij}$$

The Laplacian matrix \mathbf{L} , also known as Kirchhoff matrix, is a symmetric square $|\mathbf{V}| \times |\mathbf{V}|$ matrix (Merris 1994) defined as an augmented matrix, it being obtained as the result of a square matrix and a diagonal matrix whose diagonal elements are the atomic properties. More specifically the Laplacian matrix is defined as the difference between the vertex degree matrix and the adjacency matrix. The vertex degree matrix is a diagonal matrix whose diagonal elements are the vertex degrees, the vertex degree δ_i being the number of non-H atoms adjacent to atom i :

$$[\mathbf{L}]_{ij} = \begin{cases} \delta_i & \text{if } i = j \\ -1 & \text{if } (i, j) \in \mathbf{E} \\ 0 & \text{if } (i, j) \notin \mathbf{E} \end{cases}$$

Vertex degrees can be calculated as the row sums of the adjacency matrix \mathbf{A} :

$$\delta_i = \sum_{j=1}^{|\mathbf{V}|} a_{ij}$$

The Laplacian spectrum is the set of eigenvalues derived from the diagonalization of the Laplacian matrix. These eigenvalues are real nonnegative numbers; the last eigenvalue is always equal to zero, and, if the graph is connected, the eigenvalues, except for the last one, are greater than zero.

Quasi-Wiener index (W^*) is calculated as the product of the number $|\mathbf{V}|$ of vertices of the graph and the sum of the reciprocal non-zero eigenvalues of the Laplacian matrix (Mohar et al. 1993):

$$W^* = |\mathbf{V}| \cdot \sum_{i=1}^{|\mathbf{V}|-1} \frac{1}{\lambda_i}$$

where $|\mathbf{V}|-1$ is the number of positive eigenvalues and λ_i the i -th eigenvalue of the Laplacian matrix. The quasi-Wiener index coincides with the Wiener index for acyclic graphs.

The detour matrix Δ , also known as maximum path matrix, is a symmetric vertex matrix defined as (Harary 1969; Trinajstić et al. 1997):

$$[\Delta]_{ij} = \begin{cases} |\max p_{ij}| & \text{if } i \langle \rangle j \\ 0 & \text{if } i = j \end{cases}$$

where $|\max p_{ij}|$ is the number of edges along the longest path between vertices i and j , that is, the detour distance.

The maximum path sum (MPS) index (Ivanciuc and Balaban 1994), also known as the detour index and usually denoted as w , is a Wiener-type index defined as the half-sum of the detour distances between every pair of vertices included in the molecular graph:

$$w = \frac{1}{2} \cdot \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} [\Delta]_{ij}$$

From the standard graph-theoretical matrices, additional matrices can be calculated applying mathematical operators to a single or a set of matrices, hence obtaining, for instance, augmented matrices, reciprocal matrices, and weighted matrices.

Reciprocal matrices are obtained by raising to different powers the elements of the considered matrix. A typical example of these matrices is the reciprocal distance matrix \mathbf{D}^{-1} , also known as Harary matrix \mathbf{H} :

$$[\mathbf{D}^{-1}]_{ij} = \begin{cases} 1/d_{ij} & \text{if } i \langle \rangle j \\ 0 & \text{if } i = j \end{cases}$$

The Harary index H is calculated from the reciprocal distance matrix as a Wiener-type index:

$$H = \frac{1}{2} \cdot \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} d_{ij}^{-1}$$

Weighted matrices are usually vertex or edge matrices describing a molecular graph containing heteroatoms or multiple bonds. Weighted matrices are obtained associating to each vertex or edge a defined property, the weight, which is used to differentiate the diverse atom or bond types. The set of weights is the weighting scheme associated to the molecular graph. Typically vertex weights are atomic properties such as atomic numbers, mass, or any physicochemical property that can be associated to the atoms. The edge weights are bond properties like the conventional bond order, the bond length, some combination of the properties of the connected atoms, or any other property that can be associated to the considered bond.

The atom connectivity matrix **C** is the most known example of weighted matrices. This is a weighted adjacency matrix whose elements c_{ij} are defined as follows:

$$[\mathbf{C}]_{ij} \equiv c_{ij} = \begin{cases} \pi_{ij}^* & \text{if } (i, j) \in E \\ 0 & \text{otherwise} \end{cases}$$

where π^* is the conventional bond order.

Another example of weighted matrices is the χ matrix (Randić 1992) that is defined as follows:

$$[\chi]_{ij} = \begin{cases} \frac{1}{\sqrt{\delta_i \cdot \delta_j}} & \text{if } (i, j) \in E \\ 0 & \text{otherwise} \end{cases}$$

where δ is the vertex degree of the two connected atoms.

Randić connectivity index (χ_R) is among the first proposed topological indices (Randić 1975, 2001); it can be calculated from the χ matrix as a Wiener-type index:

$$\chi_R \equiv {}^1\chi = \frac{1}{2} \cdot \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} [\chi]_{ij} = \frac{1}{2} \cdot \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} a_{ij} (\delta_i \cdot \delta_j)^{-1/2}$$

where a_{ij} are the adjacency matrix elements, which equal 1 for pairs of connected atoms, 0 otherwise.

Burden matrix **B** is another weighted matrix, which provides the well-known Burden eigenvalue descriptors (Burden 1989). Burden matrix is defined as follows:

$$[\mathbf{B}]_{ij} = \begin{cases} \pi_{ij}^* \cdot 10^{-1} & \text{if } (i, j) \in E \\ Z_i & \text{if } i = j \\ 0.001 & \text{if } (i, j) \notin E \end{cases}$$

A generalization of Burden matrix has been proposed in Dragon software (Mauri et al. 2014).

Among the edge matrices, the most common is the edge adjacency matrix. This is a symmetric matrix whose dimension is $|E| \times |E|$, which collects information about the connectivities of molecular bonds. Formally, it is defined as:

$$[\mathbf{E}]_{ij} \equiv e_{ij} = \begin{cases} 1 & \text{if } (i, j) \text{ are adjacent bonds} \\ 0 & \text{otherwise} \end{cases}$$

Analogously to the calculation of the vertex degree δ , the edge degree ε can be calculated from the edge adjacency matrix as:

$$\varepsilon_i = \sum_{j=1}^{|\mathbf{E}|} [\mathbf{E}]_{ij}$$

Edge adjacency matrix has been used to derive different graph-theoretical invariants, such as the edge connectivity index (Estrada 1995) defined as:

$$\varepsilon = \frac{1}{2} \cdot \sum_{i=1}^{|\mathbf{E}|} \sum_{j=1}^{|\mathbf{E}|} e_{ij} (\varepsilon_i \cdot \varepsilon_j)^{-1/2}$$

where e_{ij} are the elements of the edge adjacency matrix.

The advantage of the matrix representation of a molecular graph is that a lot of matrix operators and manipulations can be applied to graph-theoretical matrices allowing the calculation of different sets of descriptors highlighting diverse information depending on the considered representation.

Generalized Topological Indices

Topological indices are single graph invariants derived from a graph-based representation of a molecular structure. Topological indices are usually calculated on an H-depleted molecular graph, and depending on their definition, they are responsive to one or more structural features of the molecule, like size, shape, symmetry, and ring system characteristics. Topological indices taking only into account the structural information of the molecule are referred to as topostructural indices, whereas topological indices taking also into account the chemical information of the molecule such as the atom types or the hybridization states are named topochemical indices (Basak et al. 1997).

Topological indices can be derived applying mathematical operators directly to graph-theoretical matrices, such as the Wiener operator that consists in the half-sum of the matrix elements, or applying mathematical functions to local vertex invariants (LOVIs). Local vertex invariants are those numerical quantities of graph vertices (i.e., the atoms) used to characterize specific properties of the molecule's atoms. In analogy with the local vertex invariants, local edge invariants are those invariants calculated on graph edges and codifying bond properties.

Local vertex invariants can be derived taking into account only the topological representation of the molecule or considering also the physicochemical properties of the atoms associating a weight, the considered physicochemical property, to each atom in the molecule. This approach enables the topological indices to discriminate among heteroatoms.

Typical local invariants are the vertex degree δ , the valence vertex degree, and the bond vertex degree ε .

Local invariants can be calculated from any graph-theoretical matrix. One of the most common approaches is to sum the values of a generic symmetric graph-theoretical matrix along a row:

$$\mathcal{L}_i = \sum_{j=1}^A [\mathbf{M}]_{ij}$$

where \mathbf{M} is the considered matrix of dimension $A \times A$.

Applying this operator to the adjacency matrix \mathbf{A} results in the calculation of the most common local vertex invariant, the vertex degree δ_i :

$$\delta_i = \sum_{j=1}^{|\mathbf{V}|} a_{ij}$$

Most of the topological indices proposed in the literature have been studied to be applied on a specific graph-theoretical matrix; for instance, the Wiener index is defined as the sum of the topological distances between all the atom pairs, but it is commonly calculated as the half-sum of the distance matrix entries.

One of the characteristics of the most common topological indices is that they can be easily generalized in order to be applied on almost any graph-theoretical matrix. This generalization allows the evaluation of the behavior of these indices when applied on matrices including different types of information.

Let \mathbf{M} be a generic $A \times A$ symmetric graph-theoretical matrix, then a collection of generalized topological indices can be defined.

The formula $D1$ is a generalization of the Wiener index:

$$D1_{(\mathbf{M})} = \alpha \cdot \sum_{i=1}^A \sum_{j=1}^A [\mathbf{M}]_{ij}^\lambda$$

where two parameters have been added to the original formula, a scale parameter α and a power parameter λ . If the matrix \mathbf{M} coincides with the distance matrix (\mathbf{D}), $\alpha = 1/2$ and $\lambda = 1$, then the formula $D1$ corresponds to the Wiener index. Analogously, if matrix \mathbf{M} is the distance matrix (\mathbf{D}), $\alpha = 1/2$ and $\lambda = -1$, then the Harary index is obtained.

An additional parameter a_{ij} can be introduced in formula $D2$ to restrict the summation to adjacent elements, a_{ij} being the elements of the adjacency matrix:

$$D2_{(\mathbf{M})} = \alpha \cdot \sum_{i=1}^A \sum_{j=1}^A a_{ij} [\mathbf{M}]_{ij}^\lambda$$

Based on generic local vertex invariants \mathcal{L} , a general formula for the calculation of generalized topological indices is defined as the following:

$$D3_{(\mathcal{L};\alpha,\lambda)} = \alpha \cdot \sum_{i=1}^A \sum_{j=1}^A a_{ij} (\mathcal{L}_i \cdot \mathcal{L}_j)_{ij}^\lambda$$

The original Randić-connectivity index is obtained substituting \mathcal{L}_i with the vertex degree of the i -th atom and setting $\alpha = 1/2$ and $\lambda = -1/2$. Analogously, the Balaban index J (Balaban 1982) is obtained setting $\alpha = |E|/2(c + 1)$, where r is the cyclomatic number, and $\lambda = -1/2$ and considering the \mathcal{L}_i as the sum of all entries of the i -th row of the distance matrix \mathbf{D} .

A further generalization of the formula $D3$ to be applied on any type of graph fragments can be written as:

$$D4_{(\mathcal{L};\alpha,\lambda,k)} = \alpha \cdot \sum_{k=1}^K \left(\prod_{i=1}^{n_k} \mathcal{L}_i \right)_k^\lambda$$

where \mathcal{L} is a generic local vertex invariant, K is the total number of selected graph fragments, and n_k is the number of vertices in the k -th fragment. Function $D4$ is the basis for the calculation of Kier–Hall connectivity indices (Kier and Hall 1977). Kier and Hall proposed a general scheme to calculate connectivity indices based on the Randić index formula but considering larger fragments than bonds (i.e., pairs of adjacent atoms):

$${}^m\chi = \sum_{k=1}^K \left(\prod_{i=1}^{n_k} \delta_i \right)_k^{-1/2}$$

where δ_i is the vertex degree of the i -th atom.

Unlike Randić-connectivity index that is calculated on all pairs of adjacent vertex, Kier and Hall connectivity indices are calculated on all the subgraphs, of a predefined order m , constituting the molecule. Zero-order subgraphs ($m = 0$) are all the atoms in the molecule; it means that the summation goes over all vertex degrees, K is equal to the number of atoms, and $n_k = 1$. Analogously, 1-order subgraphs ($m = 1$) are all pairs of adjacent atoms, that is, all paths of length one (i.e., bonds), and for $m = 1$ the Randić-connectivity index is obtained, with K equal to the number of edges $|E|$ and n_k equal to 2. Second-order connectivity indices ($m = 2$) are calculated on all paths of order 2, K is equal to the number of path of length 2 in the graph, and n_k is equal to 3 since any path of length two involves three atoms.

As described, topological indices can be easily derived from a molecular graph. The collected information depends on the considered graph-theoretical matrix, on the chosen invariants, and on the operators applied to the examined matrix. This approach, besides being a theoretical generalization of existing and well-known topological indices, enables the definition of an extremely large number of theoretical topological indices whose tangible usefulness must be verified both analyzing their chemical meaning and applying them on real cases evaluating their ability to properly correlate physicochemical, biological, or other experimental properties.

Autocorrelation Descriptors

Autocorrelation descriptors denote a heterogeneous class of molecular descriptors whose general mathematical formalization is the following:

$$D5_{(\mathcal{L};\alpha,\lambda,k)} = \alpha \cdot \sum_{i=1}^A \sum_{j=1}^A (\mathcal{L}_i \cdot \mathcal{L}_j)_{ij}^\lambda \cdot \delta(d_{ij};k)$$

where \mathcal{L} is a generic local vertex invariant; α and λ are a scaling and a power parameter, respectively; and $\delta(d_{ij};k)$ is a Kronecker delta function equal to one for pairs of substructure centers at topological distance $d_{ij} = k$ and zero otherwise; A is the number of substructure centers that typically are the molecule atoms.

Autocorrelation descriptors are related to the spatial distribution of a generic molecular property, the local vertex invariant \mathcal{L} , on the molecular structure space and measure the strength of a relationship between atoms at a predefined distance equal to k .

The most known autocorrelation descriptors are the Moreau–Broto (1980; Broto et al. 1984), Moran (1950), and Geary (1954) autocorrelation descriptors.

Moreau and Broto applied for the first time the spatial autocorrelation function to a molecular graph to measure the distribution of atomic properties on the molecular graph. They named the proposed descriptors as the autocorrelation of a topological structure (ATS); these can be derived from the general equation $D5$ setting $\alpha = 1/2$ and $\lambda = 1$ and substituting the generic local vertex invariant \mathcal{L}_i with any atomic property w_i as:

$$ATS_k = \frac{1}{2} \cdot \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} w_i \cdot w_j \cdot \delta(d_{ij};k)$$

where V is the set of vertices of the H-filled molecular graph.

Moreau–Broto autocorrelation calculated for path length zero is defined as the sum of the squares of the atomic properties:

$$ATS_0 = \sum_{i=1}^{|V|} w_i^2$$

The weight w can be defined arbitrarily. Common atomic properties are the atomic masses, electronegativities, polarizabilities, ionization potentials, and van der Waals volumes. Atomic properties should be properly centered in order to obtain uncorrelated autocorrelation descriptors (Hollas 2003).

Moran and Geary autocorrelation functions have been originally applied in ecological studies to evaluate the spatial distribution of environmental properties

and later proposed in Dragon software to evaluate the spatial distribution of atomic properties as alternative functions to ATS descriptors. Unlike Moreau–Broto autocorrelation descriptors (ATS), they explicitly take into account the mean and the standard deviation of the properties used as the atom-weighting scheme. Both Moran and Geary coefficients can be derived from the general equation *D5* just replacing the scaling and power parameters and the local vertex invariants.

Moran autocorrelation descriptors (MATS) have been defined as:

$$\text{MATS}_k = \frac{\frac{1}{\Delta_k} \cdot \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} (w_i - \bar{w}) \cdot (w_j - \bar{w}) \cdot \delta(d_{ij}; k)}{\frac{1}{|V|} \sum_{i=1}^{|V|} (w_i - \bar{w})^2}$$

where w_i is any atomic property, \bar{w} is the property mean on the molecule, and Δ_k is the sum of the Kronecker deltas, that is, the number of atom pairs at distance equal to k . Moran coefficient usually takes value in the interval $[-1, +1]$. Positive values of the coefficients reflect positive spatial autocorrelation, whereas negative values of the coefficients reflect negative spatial autocorrelation; values around 0 mean that there is no specific pattern in the atom property distribution in the molecule.

Geary autocorrelation descriptors (GATS) can be calculated using the following equation:

$$\text{GATS}_k = \frac{\frac{1}{2\Delta_k} \cdot \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} (w_i - w_j)^2 \cdot \delta(d_{ij}; k)}{\frac{1}{(|V|-1)} \sum_{i=1}^{|V|} (w_i - \bar{w})^2}$$

where w_i , \bar{w} , and Δ_k have the same definition as for MATS descriptors. Geary coefficient is a distance-type function whose values vary from zero to infinite. Small values of Geary autocorrelation coefficients reflect strong spatial autocorrelation; values between 0 and 1 reveal positive autocorrelation, while values larger than 1 are associated to negative spatial autocorrelation.

A particular case of autocorrelation descriptors is the atom-type autocorrelation (ATAC). Atom-type autocorrelation descriptors are discrete descriptors, while Moreau–Broto, Moran, and Geary descriptors are continuous functions. Atom-type autocorrelation descriptors count the occurrences of atom pairs, where atom pairs are defined as any pair of atoms having particular features at a predefined topological distance k . Also in this case, the formula of atom-type autocorrelation descriptors can be derived from the general equation *D5* as the following:

$$\text{ATAC}_k(u, v) = \frac{1}{2} \cdot \sum_{i=1}^{|V|} \sum_{j=1}^{|V|} \delta(i; u) \cdot \delta(j; v) \cdot \delta(d_{ij}; k)$$

where u and v represent two different atom types and $\delta(i; u)$, $\delta(j; v)$, and $\delta(d_{ij}; k)$ are three Kronecker delta functions equal to one if atom i is of type u , atom j is of

type v , and the topological distance d_{ij} is equal to k , respectively, zero otherwise. Atom types can be defined using one or more atomic properties; typical properties are the element type, connectivity, hybridization state, and pharmacophore properties.

Atom pair descriptors are usually calculated on a predefined set of atom types and a defined range of topological distances, the number of which determines the length of the resulting descriptor vector. Since atom pair descriptors are alignment-free, they enable fast calculation of similarity even on large databases.

Carhart descriptors are atom pair descriptors where atom types are defined using the chemical element of atoms, their connectivity calculated on the H-depleted molecular graph and the number of π electrons (Carhart et al. 1985). Alternatively, atoms can be labeled as the center of a predefined list of functional groups or structural features; the resulting descriptor vector will represent the frequencies of the occurrences of every defined atom pair on the defined range of topological distances.

Another type of well-known atom pair descriptors are the CATS (Chemically Advanced Template Search) descriptors (Schneider et al. 1999; Renner et al. 2006). The atom-type definition in CATS is related to the concept of “potential pharmacophore points” (PPP), where a PPP is a generalized atom type defined considering some physicochemical aspects. Five PPP have been originally proposed for the calculation of CATS: hydrogen-bond donor (D), hydrogen-bond acceptor (A), positive (P), negative (N), and lipophilic (L). Any atom of the molecule can be assigned to none, one, or two atom types, resulting in 15 possible atom pairs (DD, DA, DP, DN, DL, AA, AP, AN, AL, PP, PN, PL, NN, NL, LL). CATS are calculated on a topological distance varying from 0 to 9 resulting into a vector of 150 frequencies. CATS descriptors can be used as a vector of occurrences or they can be scaled. Scaling is usually performed dividing the occurrences by the number of non-hydrogen atoms or dividing each of the 15 possible PPP pairs by the sum of the occurrences of the two PPP belonging to the atom pair. The relative scaling, using the sum of the occurrences of the PPP, highlights rare PPP considering them as possible important contributors of the pharmacophoric aspect of the molecules with respect to abundant PPP pairs like the lipophilic atom types.

Autocorrelation descriptors can be extended to be calculated on a three-dimensional structure. Three-dimensional autocorrelation descriptors are based on Euclidean interatomic distances, while the property is still defined as an atomic property.

An example of three-dimensional autocorrelation descriptors is the CATS3D descriptors (Fechner et al. 2003) that are the three-dimensional counterpart of the CATS descriptors. Since distances are not calculated on the topological structure but in the real 3D space, they are not integer values, and consequently, Euclidean interatomic distances have to be partitioned into a defined set of distance bins. CATS3D have been defined using 20 distance bins from 0 to 20 Å in steps of 1 Å. The formula for these descriptors is the following:

$$\text{CATS3D}_k(u, v) = \frac{1}{2} \cdot \sum_{i=1}^{|\mathbf{V}|} \sum_{j=1}^{|\mathbf{V}|} \delta(i; u) \cdot \delta(j; v) \cdot \delta(r_{ij}; k)$$

where k is the considered distance bin, u and v represent two different atom types, and $\delta(i; u)$, $\delta(j; v)$, and $\delta(r_{ij}; k)$ are three Kronecker delta functions equal to one respectively if atom i is of type u , atom j is of type v , and their interatomic distance r_{ij} is comprised in the k -th bin, zero otherwise.

Geometrical Descriptors

Geometrical descriptors are those molecular descriptors calculated taking into account the three-dimensional structure of the molecule, i.e., the position of the atoms in the three-dimensional space. Geometrical descriptors derived from a three-dimensional graph representation of the molecule, taking into account not only the positions of the atoms but also the connections among them, are usually named as the topographic indices. The calculation of geometrical descriptors can be performed only for molecular structures whose three-dimensional description is available. Three-dimensional structure can be retrieved from crystallographic data or as the result of molecular optimization.

Geometrical descriptors potentially have higher information content than two-dimensional descriptors, but they have to be treated carefully since their values vary modifying the molecule conformation. Their application is therefore bound to those cases for which a specific conformation is selected or diversity among conformations has to be evaluated. Despite this, many geometrical descriptors have been satisfactorily used to model physicochemical and toxicological properties.

The geometrical distance matrix \mathbf{G} is the common representation of the distance relationships between all of the atom pairs in the 3D space; this is a symmetric $|\mathbf{V}| \times |\mathbf{V}|$ matrix whose entries r_{ij} are the Euclidean distances between every pair of atoms i and j in the molecule:

$$\mathbf{G} \equiv \begin{bmatrix} 0 & \cdots & r_{1|\mathbf{V}|} \\ \vdots & \ddots & \vdots \\ r_{|\mathbf{V}|1} & \cdots & 0 \end{bmatrix}$$

The 3D-Wiener index (${}^3D W_H$) is a topographic index calculated by analogy with the Wiener index from the geometrical distance matrix as:

$${}^3D W_H = \frac{1}{2} \cdot \sum_{i=1}^{|\mathbf{V}|} \sum_{j=1}^{|\mathbf{V}|} r_{ij}$$

Other topographic indices can be derived from the geometrical distance matrix applying the proper matrix operators.

A local vertex invariant that can be calculated from the geometrical distance matrix is the geometric distance degree, which is calculated as the row sum of the geometrical distance matrix as:

$$G_{\sigma_i} = \sum_{j=1}^{|V|} r_{ij}$$

where r_{ij} is the Euclidean distance between atoms i and j .

In analogy to the connectivity indices, the 3D connectivity indices ($\chi\chi$) have been defined substituting the vertex degree δ with the geometric distance degree G_{σ} :

$${}^m\chi\chi = \sum_{k=1}^K \left(\prod_{i=1}^{n_k} G_{\sigma_i} \right)_k^{-1/2}$$

where K is the number of m -th order subgraph constituted by n_k vertices.

Two additional examples of descriptors calculated using the geometrical distance matrix as a basis are the Euclidean connectivity (χ^E) index (Balasubramanian 1995):

$$\chi^E = \sum_{i=1}^{A-1} \sum_{j=i+1}^A (G_{\sigma_i} \cdot G_{\sigma_j})^{-\frac{1}{2}}$$

and the 3D Balaban index (Mihalic et al. 1992):

$${}^{3D}J = \frac{|E|}{c+1} \cdot \sum_{i=1}^{A-1} \sum_{j=i+1}^A a_{ij} (G_{\sigma_i} \cdot G_{\sigma_j})^{-1/2}$$

where c is the cyclomatic number, that is, the number of independent cycles in the structure.

Matrices used to encode information about a molecular structure can be combined in different ways in order to obtain novel matrices taking into account different features of the structure. An example of matrix combination is the distance/distance matrix $\mathbf{G/D}$, which is a $|V| \times |V|$ symmetric vertex matrix whose values are defined as the ratio between the geometrical distance and the topological distance:

$$[\mathbf{G/D}]_{ij} = \begin{cases} \frac{r_{ij}}{d_{ij}} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}$$

where r_{ij} and d_{ij} are the Euclidean and the topological distance between atoms i and j , respectively.

Two of the most known classes of three-dimensional descriptors are the WHIM (Weighted Holistic Invariant Molecular) descriptors (Todeschini et al. 1994; Todeschini and Gramatica 1997) and the GETAWAY (Geometry, Topology, and Atom-Weights Assembly) descriptors (Consonni et al. 2002a, b).

Unlike the topographic indices, which combine information on the topological and 3D structure of molecules, WHIM descriptors do not take into account the connections among atoms but only their position in the three-dimensional space. WHIM descriptors encode three-dimensional information on size, shape, symmetry, and atomic property distribution. Their calculation exploits the eigenvalues and eigenvectors of a weighted covariance matrix of the centered Cartesian atomic coordinates derived from a three-dimensional representation of the molecular structure: eigenvectors define the principal axes of the molecule and, accordingly, the unique reference frame for the considered molecular structure. WHIM descriptors are invariant to roto-translation: they are invariant to translation as a result of the centering of the atomic coordinates and to rotation as long as the principal axes are unique. These characteristics are fundamental in order to define molecular descriptors independent of prior molecule alignment.

The weighted covariance matrix for a generic atomic weighting scheme w is defined as:

$$s_{jk}^2 = \frac{\sum_{i=1}^A w_i (q_{ij} - \bar{q}_j) (q_{ik} - \bar{q}_k)}{\sum_{i=1}^A w_i}$$

where the summations run over the molecule atoms, s_{jk} is the weighted covariance between the two sets of atomic coordinates j and k ($j, k = x, y, z$), w_i is the atomic property of atom i , and q_i , q_j , and \bar{q} represent the coordinates of the i -th and j -th atoms and the corresponding average value, respectively.

Different weighting schemes have been proposed and accordingly different covariance matrices and hence different principal axes were obtained. WHIM descriptors have been divided in two classes, directional WHIM descriptors and global WHIM descriptors.

The three eigenvalues of the weighted covariance matrix play a crucial role in the WHIM descriptor calculation. Each eigenvalue represents a dispersion measure of the projected atoms along the considered principal axis. Directional WHIM descriptors consider the three eigenvalues separately providing information on the molecular size, shape, density, and symmetry along the principal axis, while global WHIM descriptors consider the relationships among the eigenvalues to describe molecular properties along the three principal directions in the molecule.

GETAWAY descriptors are three-dimensional descriptors that encode information about the influence that each atom has in determining the whole shape in the molecule and evaluate the interactions among atoms with respect to their

geometrical position in the three-dimensional space. For their calculation molecular matrix, representations are considered.

The first matrix is called molecular influence matrix \mathbf{H} , which is an $A \times A$ symmetric matrix, defined as:

$$\mathbf{H} = \mathbf{M} \times (\mathbf{M}^T \times \mathbf{M})^{-1} \times \mathbf{M}^T$$

where \mathbf{M} is the molecular matrix of the centered Cartesian coordinates (x, y, z) calculated on the H-filled three-dimensional molecular graph for a defined conformation. Centering the atomic coordinates is done with respect to the geometrical center of the molecule and assures the translational invariance. Influence matrix \mathbf{H} is invariant to rotation of the molecular coordinates, granting that GETAWAY descriptors are independent of any alignment.

Elements of the influence matrix can be split, with respect to their meaning, in two sets. The former comprises the diagonal elements of \mathbf{H} , denoted as h_i and called leverages, whose range is from 0 to 1. Leverages represent the contribution of each atom in defining the molecule shape; atoms located near the molecular surface have higher leverages than atoms near the molecule center. The latter includes the off-diagonal elements h_{ij} that reflect the capability of the j -th atom to interact with the i -th atom. Positive values of h_{ij} mean that atoms i and j are positioned in the same molecular region revealing possible atomic interactions, while negative values of h_{ij} mean that the two considered atoms i and j are positioned in opposite molecular region with respect to the center of the molecule reflecting low possibility of interaction.

The second matrix to be considered for GETAWAY calculation is the influence/distance matrix \mathbf{R} , an $A \times A$ symmetric matrix determined combining the elements of the influence matrix \mathbf{H} with the elements of the geometrical distance matrix \mathbf{G} :

$$[\mathbf{R}]_{ij} = \begin{cases} \frac{\sqrt{h_i \cdot h_j}}{r_{ij}} & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases}$$

where h_i and h_j are the leverages of atoms i and j , respectively, and r_{ij} is the Euclidean distance between them. Diagonal elements are set to zero by definition while off-diagonal elements represent the interactions between pairs of atoms, which decrease as the interatomic distance increases. This matrix has been defined to emphasize the interactions between atoms that are near each other and located on the molecular surface.

Unlike WHIM descriptors that allow a holistic description of the molecular properties, GETAWAY descriptors have been developed in order to exploit local information of molecular structures.

Structural Keys and Molecular Fingerprints

All the molecular descriptors presented in the previous sections derive from a mathematical manipulation of a molecular structure representation, typically a matrix-based representation, using a well-defined set of mathematical operators. Those descriptors have been defined in order to highlight different chemical aspects of a molecular structure starting from a holistic representation of the chemicals.

On the contrary, structural keys and molecular fingerprints try to describe a chemical capturing different local aspects of a molecular structure. Specifically, these approaches identify a set of structural fragments participating in the molecule composition.

Typically, structural keys are a predefined set of structural features, atom pairs, functional groups, ring systems, and atom-centered fragments that have been prepared in order to be able to discriminate among molecules and eventually be able to highlight chemical properties of a molecule. Structural keys can be defined as occurrence frequency or binary vector depending on whether the structural features identified in the molecules are counted or denoted only as present or absent.

Several structural key sets have been defined; the most known are the two sets of MACCS keys (Durant et al. 2002) comprising 960 and 166 structural features, respectively, and the PubChem structural keys, which include 881 bits reflecting the presence or absence of 881 structural features. PubChem service uses PubChem fingerprints for similarity neighboring and similarity searching (Bolton et al. 2008).

Unlike the structural keys that are a well-defined set of molecular features (i.e., fragment dictionary), the molecular fingerprints are identified exploring the molecular structure in all possible substructure patterns following a set of rules. Depending on the rules used to identify the molecular patterns, different types of fingerprint can be calculated.

Substructure patterns can be identified increasing or lowering the discriminatory power just defining the rules to be used to recognize and differentiate the patterns. The first rule to be considered is related to the kind of patterns that can be identified during the fingerprint calculation process. Two major types of patterns can be identified: paths and atom centered, also known as circular patterns. Then two classes of fingerprints can be calculated, path and extended connectivity fingerprint, also known as Morgan or circular fingerprint (Rogers and Hahn 2010).

Exhaustively enumerating all linear fragments of a molecular graph up to a given size generates a path fingerprint, while exhaustively enumerating all circular patterns grown radially from each heavy atom of the molecule up to the given radius size generates a circular fingerprint.

The first parameter to be set once the type of fingerprint has been defined is the maximum path length or the maximum radius of the patterns to be identified. Then, additional options can be set in order to increase or decrease the discrimination power of substructures. Patterns can be differentiated defining the atom properties encoded in the atom identifiers; typical properties that can be considered during the pattern identification are the atomic number, the number of neighbor atoms (i.e., atom connectivity), the number of bonded hydrogens, the formal charge, and the atom aromaticity.

Since recognized substructures are not included in a defined list, the number of identified substructures varies depending on the considered molecules affecting the fingerprint vector length. Typically, molecular fingerprint is a Boolean vector coding the presence or absence of structural features. On the contrary, some molecular fingerprints have been defined as a set of occurrences of different molecular fragments generated by breaking down all considered molecules; an example of this kind of fingerprint is the ISIDA descriptors (Ruggiu et al. 2010).

Alternatively, in order to obtain a fixed-length vector of molecular fragments, the Boolean vector constituting the fingerprint is processed using hash functions. A hash function is any function that can be used to reduce a variable-length Boolean vector to a fixed-length one. Specifically, each identified pattern is used as pseudorandom number generator seed, and the output of the hash function is a set of bits. The set of bits produced by hashing all the patterns identified in the molecule is added using the logical OR operator to produce the fingerprint of the whole molecule. This means that every bit associated to a substructure pattern will be set in the molecule fingerprint enabling Boolean operations on fingerprints to screen molecule databases; Boolean operations allow fingerprint comparison extremely fast for substructure searching. Fingerprint hashing is introduced when it is not possible to assign a particular bit to each pattern, as for structural keys, due to the huge number of possible identifiable patterns.

Hashing functions enable the definition of a fixed-length vector but introduce the so-called bit collision; this means that, under a predefined set of rules, a specific fragment will always be associated to a defined set of bits in the fingerprint, but two different fragments may share one or more bits among their bit sets.

Unlike structural keys that reveal with certainty the presence or absence of a particular pattern, fingerprints are not so accurate since if a fingerprint indicates the absence of a pattern then that particular substructure is certainly missing; on the contrary, a fingerprint can only indicate a pattern's presence with some probability. Nevertheless, fingerprints encode an almost exhaustive set of patterns with respect to structural keys, resulting in a more detailed description of a molecular structure in almost all situations.

Molecular fingerprints have been originally introduced for fast database searching since they are able to identify common substructures; however, molecular fingerprints are currently more and more used for building QSAR models, especially those based on local similarity as in the kNN approach (Cassotti et al. 2014b).

Future Perspectives

Despite the research on molecular descriptors and QSAR is quite recent, it is now able to provide solid answers to real problems. Molecular descriptor theory is facing new challenges requiring both chemical and mathematical skills. The description of structures not considered before, like disconnected structures and nanomaterials, is one of these challenges.

Molecular Descriptors for Disconnected Structures

Molecular descriptors have been widely used in order to deploy accurate models for the prediction of toxicological and physicochemical properties of organic compounds. The increasing number of available chemical data together with the progress in molecular descriptor theory and chemoinformatic techniques are the reasons for this successful approach.

One of the challenges that molecular modeling is facing is the capability to be applied not only on small organic molecules but also on different kinds of chemicals, like disconnected structures. Disconnected structures comprise salts, mixture, and ionic liquids. Recently, several studies approached this problem focusing on the prediction of nonadditive physicochemical properties like density, bubble temperature, and azeotropic behavior (Ajmani et al. 2006). Toxicological properties of mixtures have been studied too (Roy et al. 2014).

Besides data availability, one of the relevant issues concerning modeling of disconnected structures is the representation of these structures through appropriate molecular descriptors (Oprisiu et al. 2013). Most of the existing molecular descriptors have been developed to catch molecular structure information of small, full connected, organic molecules. Recent studies are evaluating the feasibility to apply these classical descriptors to disconnected structures. Considering the calculation algorithm, the classical descriptors can be divided into two classes. The former comprises all the descriptors whose mathematical definition does not require that the molecule is full connected; the latter includes those descriptors that can be calculated only on a fully connected graph. Trivial examples of molecular descriptors that can be calculated, preserving their chemical meaning on disconnected structures, are most of the constitutional descriptors; molecular weight is a molecular descriptor whose meaning is preserved on disconnected structures; a cyclomatic number is another descriptor whose definition comprises the possibility to be calculated on disconnected structures. Functional groups, atom-centered fragments, fingerprints, and structural keys are additive descriptors whose interpretation is identical for fully connected and non-fully connected structures.

While this first class of descriptors can be directly calculated on disconnected structures, the second class requires a deeper evaluation on how these descriptors can be calculated on disconnected structures. One approach is to calculate molecular descriptors on every full connected constituent of the compound and then perform a linear combination of the calculated values. Linear combinations typically are the average, the sum, or the absolute difference of the descriptor values (Oprisiu et al. 2013).

Obviously, the characterization of disconnected structures using molecular descriptors is a complex field since, in addition to the traditional issues, interaction among mixture constituents should be taken into account.

Data availability will be, as for traditional QSAR, the drive that will permit to evaluate the possibility to describe disconnected structures with molecular descriptors. Simultaneously, the advance in molecular descriptor theory could give

the instruments to develop and test molecular descriptors devoted to the description of disconnected structures.

Molecular Descriptors for Nanostructures

Nanomaterials are usually defined as those materials of which a single unit size is between 1 and 1000 nanometers. Besides the need to evaluate physicochemical properties of this class of chemicals, an emerging issue, related to the increasing use of engineered nanomaterials, is the assessment of their potential toxicological effects toward biological and environmental systems (Haasch et al. 2005; Buzea et al. 2007).

As well as for disconnected structures, the chance to use a quantitative structure–activity relationship approach is being widely evaluated during the last years (Puzyn et al. 2009; Oksel et al. 2015). QSAR applied to nanomaterials, named nano-QSAR or QNAR (quantitative nanostructure–activity relationships), is an emerging research field whose preliminary steps have been moved (Fourches et al. 2010, 2011).

As well as for disconnected structures also for nanomaterials, the fascinating aspect that needs to be studied is the definition of proper molecular descriptors able to properly delineate a nanostructure. High structural complexity and diversity among nanomaterials increase the difficulties to develop quantitative parameters capable to encode their structural information.

Due to the lack of theoretical descriptors, currently most of the molecular descriptors used to provide nanostructure description are experimental descriptors. Experimental molecular descriptors are usually derived from images captured by electron microscopy, transmission electron microscopy, and atomic force microscopy. These descriptors define size distribution, agglomeration state, shape, porosity, and surface area (Puzyn et al. 2009). Besides experimental descriptor usage, some researches attempt to describe nanomaterials using traditional theoretical descriptors (Kar et al. 2014).

So far, the data availability is one of the major bottlenecks for the advancement in nano-QSAR studies. The study and definition of theoretical molecular descriptors accomplishing the requirements defined in the introduction, such as structural interpretation, good correlation with at least one experimental property, and especially a well-defined applicability on nanostructures, is one of the biggest challenges that nano-QSAR should face.

Conclusions

In this chapter an overview of the different molecular structure representations along with the theoretical basis of molecular descriptors has been presented. The intention was to provide the readers not only with a detailed description of some of the most known molecular descriptors proposed during the last 40 years but also, above all,

with the theoretical fundamentals to understand the generation process of molecular descriptors.

The molecular graph is the most common representation of the molecular structure, and from it, applying a finite set of operators, a huge number of descriptors can be derived. Most of them have clear chemical interpretation but others rather seem to be the outcome of some mathematical trick. In conclusion, when dealing with molecular descriptors, one should keep in mind that molecular descriptors are numbers that can be easily generated through different combinations of structure representations, atom/bond weighting schemes, and mathematical functions, but in order to produce useful descriptors, some basic rules should be fulfilled not only related to their mathematical definition but also to their chemical interpretation and the correlation with one or more experimental property.

Bibliography

- Ajmani, S., Rogers, S. C., Barley, M. H., & Livingstone, D. J. (2006). Application of QSPR to mixtures. *Journal of Chemical Information and Modeling*, *46*, 2043–2055.
- Balaban, A. T. (1982). Highly discriminating distance-based topological index. *Chemical Physics Letters*, *89*, 399–404.
- Balaban, A. T. (1985). Applications of graph theory in chemistry. *Journal of Chemical Information and Computer Sciences*, *25*, 334–343.
- Balasubramanian, K. (1995). Geometry-dependent connectivity indices for the characterization of molecular structures. *Chemical Physics Letters*, *235*, 580–586.
- Basak, S. C., Gute, B. D., & Grunwald, G. D. (1997). Use of topostructural, topochemical, and geometric parameters in the prediction of vapor pressure: A hierarchical QSAR approach. *Journal of Chemical Information and Computer Sciences*, *37*, 651–655.
- Bobra, A., Shiu, W. Y., & Mackay, D. (1985). Quantitative structure-activity relationships for the acute toxicity of chlorobenzenes to *Daphnia magna*. *Environmental Toxicology and Chemistry*, *4*, 297–305.
- Bolton, E. E., Wang, Y., Thiessen, P. A., & Bryant, S. H. (2008). PubChem: Integrated platform of small molecules and biological activities. *Annual Reports in Computational Chemistry*, *4*, 217–241.
- Boyle, N. M. O., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *Journal of Chemical Information and Modeling*, *3*, 33.
- Broto, P., Moreau, G., & Vanduycke, C. (1984). Molecular structures: Perception, autocorrelation descriptor and sar studies. *European Journal of Medicinal Chemistry*, *19*, 66–70.
- Burden, F. R. (1989). Molecular identification number for substructure searches. *Journal of Chemical Information and Computer Sciences*, *29*, 225–227.
- Buzea, C., Pacheco, I. L., & Robbie, K. (2007). Nanomaterials and nanoparticles: Sources and toxicity. *Biointerphases*, *2*, MR17–MR71.
- Carhart, R. E., Smith, D. H., & Venkataraghavan, R. (1985). Atom pairs as molecular features in structure-activity studies: Definition and applications. *13*, 8–11.
- Cassotti, M., Ballabio, D., Consonni, V., Mauri, A., Tetko, I. V., & Todeschini, R. (2014a). Prediction of acute aquatic toxicity toward *Daphnia magna* by using the GA-kNN method. *ATLA, Alternatives to Laboratory Animals*, *42*, 31–41.
- Cassotti, M., Consonni, V., Mauri, A., & Ballabio, D. (2014b). Validation and extension of a similarity-based approach for prediction of acute aquatic toxicity towards *Daphnia magna*. *SAR and QSAR in Environmental Research*, *25*, 1013–1036.

- Cherkasov, A., Muratov, E. N., Fourches, D., Varnek, A., Baskin, I. I., Cronin, M., Dearden, J. C., Gramatica, P., Martin, Y. C., Todeschini, R., Consonni, V., Kuz, V. E., Cramer, R. D., Benigni, R., Yang, C., Rathman, J. F., Terfloth, L., Gasteiger, J., Richard, A. M., & Tropsha, A. (2014). QSAR modeling: Where have you been? Where are you going to? *Journal of Medicinal Chemistry*, *57*, 4977–5010.
- Consonni, V., Todeschini, R., & Pavan, M. (2002a). Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 1. Theory of the novel 3D molecular descriptors. *Journal of Chemical Information and Computer Sciences*, *42*, 682–692.
- Consonni, V., Todeschini, R., Pavan, M., & Gramatica, P. (2002b). Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 2. Application of the novel 3D molecular descriptors to QSAR/QSPR studies. *Journal of Chemical Information and Computer Sciences*, *42*, 693–705.
- Corbett, P. T., Leclaire, J., Vial, L., West, K. R., Wietor, J. L., Sanders, J. K. M., & Otto, S. (2006). Dynamic combinatorial chemistry. *Chemical Reviews*, *106*, 3652–3711.
- Cormen, T. H., Leiserson, C. E., & Rivest, R. L. (2001). *Introduction to algorithms* (2nd ed.). The MIT Press/McGraw-Hill.
- Durant, J. L., Leland, B. A., Henry, D. R., & Nourse, J. G. (2002). Reoptimization of MDL keys for use in drug discovery. *Journal of Chemical Information and Computer Sciences*, *42*, 1273–1280.
- Elyashberg, M., Williams, A. J., Blinov, K. (2011). Contemporary computer-assisted approaches to molecular structure elucidation. *Royal Society of Chemistry*.
- Estrada, E. (1995). Edge adjacency relationships and a novel topological index related to molecular volume. *Journal of Chemical Information and Computer Sciences*, *35*, 31–33. doi:[10.1021/ci00023a004](https://doi.org/10.1021/ci00023a004).
- Fechner, U., Franke, L., Renner, S., Schneider, P., & Schneider, G. (2003). Comparison of correlation vector methods for ligand-based similarity searching. *Journal of Computer-Aided Molecular Design*, *17*, 687–698.
- Fouches, D., Muratov, E. N., & Tropsha, A. (2010). Trust but verify: On the importance of chemical structure curation in cheminformatics and QSAR modeling research. *Journal of Chemical Information and Modeling*, *50*, 1189–1204.
- Fourches, D., Pu, D., Tassa, C., Weissleder, R., Shaw, S. Y., Mumper, R. J., & Tropsha, A. (2010). Quantitative nanostructure – Activity relationship modeling. *ACS Nano*, *4*, 5703–5712. doi:[10.1021/nn1013484](https://doi.org/10.1021/nn1013484).
- Fourches, D., Pu, D., & Tropsha, A. (2011). Exploring quantitative nanostructure-activity relationships (QNAR) modeling as a tool for predicting biological effects of manufactured nanoparticles. *Combinatorial Chemistry & High Throughput Screening*, *14*, 217–225. doi:[10.2174/138620711794728743](https://doi.org/10.2174/138620711794728743).
- Geary, R. C. (1954). The contiguity ratio and statistical mapping. *Incorporated Statistician*, *5*, 115–127, 129–146. doi:[10.2307/2986645](https://doi.org/10.2307/2986645)
- Gissi, A., Lombardo, A., Roncaglioni, A., Gadaleta, D., Mangiardi, G. F., Nicolotti, O., & Benfenati, E. (2015). Evaluation and comparison of benchmark QSAR models to predict a relevant REACH endpoint: The bioconcentration factor (BCF). *Environmental Research*, *137C*, 398–409.
- Guha, R., & Willighagen, E. L. (2012). A survey of quantitative descriptions of molecular structure. *Current Topics in Medicinal Chemistry*, *12*, 1946–1956. doi:[10.1016/j.biotechadv.2011.08.021](https://doi.org/10.1016/j.biotechadv.2011.08.021). [Secreted](#).
- Haasch, M. L., McClellan-Green, P., & Oberdörster, E. (2005). Consideration of the toxicity of manufactured nanoparticles. *AIP Conference Proceedings*, *786*, 586–590.
- Hansch, C., Leo, A., & Livingstone, D. J. (1996). Exploring QSAR fundamentals and applications in chemistry and biology. *Pesticide Biochemistry and Physiology*, *56*, 78.
- Harary, F. (1969). *Graph theory*. Reading: Addison-Wesley.
- Hollas, B. (2003). An analysis of the autocorrelation descriptor for molecules. *Journal of Mathematical Chemistry*, *33*, 91–101.

- Hughes, K., Paterson, J., & Meek, M. E. (2009). Tools for the prioritization of substances on the Domestic Substances List in Canada on the basis of hazard. *Regulatory Toxicology and Pharmacology*, 55, 382–393.
- Irwin, J. J., & Shoichet, B. K. (2005). ZINC – A free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling*, 45, 177–182.
- Ivanciuc, O., & Balaban, A. T. (1994). Design of topological indices. Part 8. Path matrices and derived molecular graph invariants. *MATCH Communications Mathematical and in Computer Chemistry*, 30, 141–152.
- Ivanciuc, O., & Ivanciuc, T. (2000). Matrices and structural descriptors computed from molecular graph distances. In A. T. Balaban & J. Devillers (Eds.), *Topological indices and related descriptors in QSAR and QSPR* (pp. 221–277). Amsterdam: Gordon and Breach Science Publishers.
- Jurs, P. C., Dixon, J. S., & Egolf, L. M. (1995). Representations of molecules. In H. Van Waterbeemd (Ed.), *Chemometrics methods in molecular design* (Vol. 2, pp. 15–38). New York: VCH Publishers.
- Kar, S., Gajewicz, A., Puzyn, T., & Roy, K. (2014). Nano-quantitative structure-activity relationship modeling using easily computable and interpretable descriptors for uptake of magnetofluorescent engineered nanoparticles in pancreatic cancer cells. *Toxicology in Vitro*, 28, 600–606.
- Kier, L. B., & Hall, L. H. (1977). The nature of structure-activity relationships and their relation to molecular connectivity. *European Journal of Medicinal Chemistry*, 12, 307–375.
- Kind, T., & Fiehn, O. (2010). Advances in structure elucidation of small molecules using mass spectrometry. *Bioanalytical Reviews*, 2, 23–60.
- Kühne, R., Ebert, R. U., Vonderohe, P. C., Ulrich, N., Brack, W., & Schüürmann, G. (2013). Read-across prediction of the acute toxicity of organic compounds toward the water flea *Daphnia magna*. *Molecular Informatics*, 32, 108–120.
- Lehn, J.-M. (1999). Dynamic combinatorial chemistry and virtual combinatorial libraries. *Chemistry A European Journal*, 5, 2455–2463.
- Lovasz, L., & Pelikan, J. (1973). On the eigenvalues of trees. *Periodica Mathematica Hungarica*, 3, 175–182.
- Mannhold, R., Poda, G. I., Ostermann, C., & Tetko, I. V. (2009). Calculation of molecular lipophilicity: State-of-the-art and comparison of log P methods on more than 96,000 compounds. *Journal of Pharmaceutical Sciences*, 98, 861–893.
- Mauri, A., Manganaro, A., Todeschini, R., Consonni, V., & Ballabio, D. (2014). Dragon software for molecular descriptor calculation.
- Merris, R. (1994). Laplacian matrices of graphs: A survey. *Linear Algebra and its Applications*, 197–198, 143–176.
- Mihalic, Z., Nikolić, S., & Trinajstić, N. (1992). Comparative study of molecular descriptors derived from the distance matrix. *Journal of Chemical Information and Modeling*, 32, 28–37.
- Mohar, B., Babic, D., & August, R. (1993). A novel definition of the Wiener index for trees. *Journal of Chemical Information and Computer Sciences*, 33, 153–154.
- Moran, P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37, 17–23.
- Moreau, J. L., & Broto, P. (1980). Autocorrelation of molecular structures: Application to SAR studies. *Nouveau Journal de Chimie*, 4, 757–764.
- Murray-Rust, P. (1999). Chemical markup, XML, and the Worldwide Web. 1. Basic principles. *Journal of Chemical Information and Computer Sciences*, 39, 928–942.
- Murray-Rust, P., & Rzepa, H. S. (2001). Chemical markup, XML and the World-Wide Web. 2. Information objects and the CMLDOM. *Journal of Chemical Information and Computer Sciences*, 41, 1113–1123.
- Oksel, C., Ma, C. Y., Liu, J. J., Wilkins, T., & Wang, X. Z. (2015). (Q)SAR modelling of nanomaterial toxicity: A critical review. *Particuology*, 21, 1–19.
- Oprisiu, I., Novotarskyi, S., & Tetko, I. V. (2013). Modeling of non-additive mixture properties using the Online CHEmical database and Modeling environment (OCHEM). *Journal of Chemical Information and Modeling*, 5, 4. doi:10.1186/1758-2946-5-4.

- Pence, H. E., & Williams, A. (2010). Chemspider: An online chemical information resource. *Journal of Chemical Education*, 87, 1123–1124.
- Puzyn, T., Leszczynska, D., & Leszczynski, J. (2009). Toward the development of “Nano-QSARs”: Advances and challenges. *Small*, 5, 2494–2509. doi:10.1002/smll.200900179.
- Randić, M. (1975). On characterization of molecular branching. *Journal of the American Chemical Society*, 97, 6609–6615. doi:10.1021/ja00856a001.
- Randić, M. (1992). Similarity based on extended basis descriptors. *Journal of Chemical Information and Modeling*, 32, 686–692. doi:10.1021/ci00010a016.
- Randić, M. (1996). Molecular bonding profiles. *Journal of Mathematical Chemistry*, 19, 375–392. doi:10.1007/BF01166727.
- Randić, M. (2001). The connectivity index 25 years after. *Journal of Molecular Graphics and Modelling*, 20, 19–35. doi:10.1016/S1093-3263(01)00098-5.
- Renner, S., Fechner, U., & Schneider, G. (2006). Alignment-free pharmacophore patterns – A correlation vector approach. In T. Langer & R. D. Hoffmann (Eds.), *Pharmacophores and pharmacophore searches* (pp. 49–79). Weinheim: Wiley-VCH.
- Rogers, D., & Hahn, M. (2010). Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, 50, 742–754. doi:10.1021/ci100050t.
- Roy, K., Das, R. N., & Popelier, P. L. a. (2014). Quantitative structure-activity relationship for toxicity of ionic liquids to *Daphnia magna*: Aromaticity vs. lipophilicity. *Chemosphere*, 112, 120–127. doi:10.1016/j.chemosphere.2014.04.002.
- Ruggiu, F., Marcou, G., Varnek, A., & Horvath, D. (2010). ISIDA property-labelled fragment descriptors. *Molecular Informatics*, 29, 855–868.
- Schneider, G., Neidhart, W., Giller, T., & Schmid, G. (1999). “Scaffold-Hopping” by topological pharmacophore search: A contribution to virtual screening. *Angewandte Chemie International Edition in English*, 38, 2894–2896.
- Testa, B., & Kier, L. B. (1991). The concept of molecular structure in structure-activity relationship studies and drug design. *Medicinal Research Reviews*, 11, 35–48.
- Todeschini, R., & Consonni, V. (2009). *Molecular descriptors for chemoinformatics* (Vol. 2). Weinheim: Wiley-VCH.
- Todeschini, R., & Gramatica, P. (1997). The Whim theory: New 3D molecular descriptors for QSAR in environmental modelling. *SAR and QSAR in Environmental Research*, 7, 89–115.
- Todeschini, R., Lasagni, M., & Marengo, E. (1994). New molecular descriptors for 2D and 3D structures. Theory. *Journal of Chemometrics*, 8, 263–272.
- Trinajstić, N., Nikolic, S., Lucic, B., Amic, D., & Mihalic, Z. (1997). The Detour matrix in chemistry. *Journal of Chemical Information and Modeling*, 37, 631–638.
- Tropsha, A. (2010). Best practices for QSAR model development, validation, and exploitation. *Molecular Informatics*, 29, 476–488.
- Vighi, M., & Calamari, D. (1985). QSARs for organotin compounds on *Daphnia magna*. *Chemosphere*, 14, 1925–1932.
- Weininger, D. (1988). SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Modeling*, 28, 31–36.
- Wiener, H. (1947). Structural determination of paraffin boiling points. *Journal of the American Chemical Society*, 69, 17–20.
- Williams, A., & Tkachenko, V. (2014). The Royal Society of Chemistry and the delivery of chemistry data repositories for the community. *Journal of Computer-Aided Molecular Design*, 28, 1023–1030.
- Worth, A. P. (2010). Chapter 13: The role of QSAR methodology in the regulatory assessment of chemicals. Media. pp 367–382.
- Young, D., Martin, T., Venkatapathy, R., & Harten, P. (2008). Are the chemical structures in your QSAR correct? *QSAR and Combinatorial Science*, 27, 1337–1345.
- Zeeman, M., Auer, C. M., Clements, R. G., Nabholz, J. V., & Boethling, R. S. (1995). U.S. EPA regulatory perspectives on the use of QSAR for new and existing chemical evaluations. *SAR and QSAR in Environmental Research*, 3, 179–201.

Katarzyna Odziomek, Anna Rybinska, and Tomasz Puzyn

Contents

Introduction	2096
Data	2097
Data Preparation	2097
Clustering	2098
Similarity and Distance	2099
Hierarchical Clustering	2102
<i>k</i> -means	2113
PCA	2118
SOM	2124
Building SOMs	2124
Summary	2130
Bibliography	2131

Abstract

In this chapter, we present an overview of various chemometric methods, appropriate for analyzing and interpreting data from social media, industry, academia, medicine, and other sources. We discuss unsupervised machine-learning techniques used for grouping (hierarchical cluster analysis, *k*-means) and exploring (principal component analysis, self-organizing Kohonen maps) all types of data, both quantitative and qualitative. For each method described in this chapter, we explain the basic concepts, provide a rudimentary algorithm, and present practical applications. All the examples are based on a set of molecular descriptors calculated for a selected group of persistent organic pollutants (POPs).

K. Odziomek (✉) • A. Rybinska • T. Puzyn
Laboratory of Environmental Chemometrics, Faculty of Chemistry, University of Gdańsk,
Gdańsk, Poland
e-mail: k.odziomek@qsar.eu.org; rybinska@qsar.eu.org; puzi@qsar.eu.org

Introduction

The vast amount of digital information generated every day in social media, industry, and academia necessitates the use of advanced techniques appropriate for processing, analysis, and interpretation of data.

There are two main types of data analysis algorithms, namely, supervised and unsupervised machine-learning methods. A simplified schematic of machine-learning methods is presented in Fig. 1. Supervised learning is used for modeling, i.e., making predictions with the help of a calibration, discrimination, or classification model, depending on the research problem. Unsupervised learning Principal component analysis (PCA), used for exploring the hidden data structures and relationships between variables, helps us find groups (clusters) of objects (samples) similar to each other or, conversely, significantly dissimilar from the rest, as defined by a selected metric. The main difference between the two methods is that when constructing a model, the unsupervised learning method utilizes only the explanatory (independent) variable matrix (\mathbf{X}), while the supervised learning method takes also the response (dependent) variable (y) into account (Brown et al. 2009).

This chapter discusses methods of unsupervised machine learning, or pattern recognition, which are often used in problem solving in various fields of research such as chemistry, economics, forensic science, and medicine (Skwarzec et al. 2011; Li et al. 2013; Kountchev and Iantovics 2013; Golebiowski et al. 2014; Petushkova et al. 2014; Schnegg et al. 2015). It is meant as a brief overview of the most popular methods, along with examples, so as to facilitate an easier understanding of the topic.

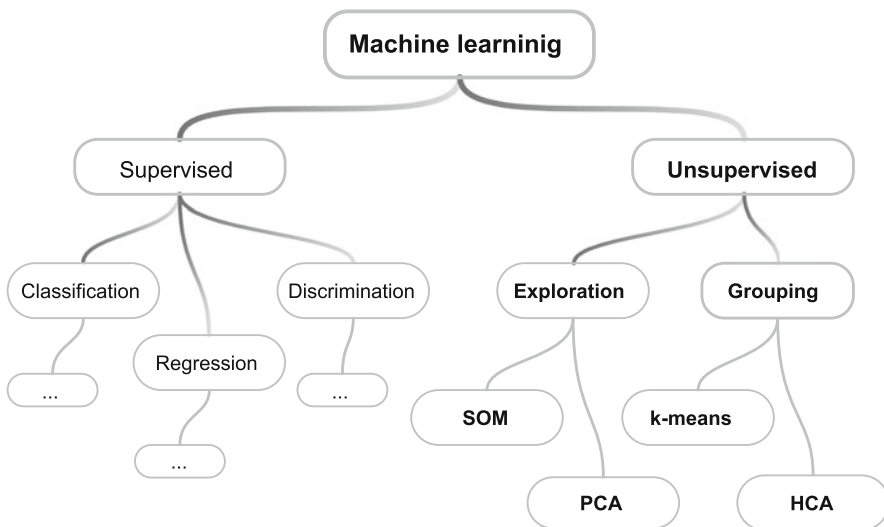


Fig. 1 An overview of data mining methods

Data

All examples given in this chapter are based on a data set on persistent organic pollutants (POPs). This data set contains 21 molecular descriptors (Table 1) calculated for 1436 chloro (“C”) and bromo (“B”) analogues of dibenzo-*p*-dioxins (PXDDs, where X stands either for C or B), dibenzofurans (PXDFs), biphenyls (PXBs), naphthalenes (PXNs), diphenyl ethers (PXDEs), and benzenes (XBs) (Gajewicz et al. 2010). Most of the examples in this chapter will be based on a reduced set of POPs, namely, polychlorinated naphthalenes PCNs, polychlorinated dibenzo-*p*-dioxins PCDDs, and polychlorinated dibenzofurans PCDFs (Fig. 2).

As the POP class membership of each sample is known beforehand, the authors chose to use this data as a tool for comparing the effectiveness and the performance of each clustering method presented in this chapter.

Data Preparation

In order to ensure the reliability and accuracy of the results, we must first evaluate the quality of the raw data. The purpose of such evaluation is identifying possible mistakes (errors) made during data collection and reducing the risk of propagating

Table 1 Molecular descriptors

Symbol	Definition of molecular descriptor	Unit
nAT	Number of atoms	—
nX	Number of halogen substituents	—
nO	Number of oxygen atoms	—
MW	Molecular weight	u
HOF	Standard heat of formation	kcal·mol ⁻¹
EE	Electronic energy	eV
Core	Core-core repulsion energy	eV
TE	Total energy	eV
TEp	Total energy of the corresponding cation	eV
VIP	Vertical ionization potential	eV
HOMO	Energy of the highest occupied molecular orbital	Hartree
LUMO	Energy of the lowest unoccupied molecular orbital	Hartree
D	Dipole moment	Debye
SAS	Solvent accessible surface	Å ²
MV	Molecular volume	Å ³
Qm	Lowest negative Mulliken partial charge on the molecule	—
Qp	Highest positive Mulliken partial charge on the molecule	—
P	Polarizability derived from the dipole moment	Å ³
EN	Mulliken electronegativity	Hartree
Hard	Parr and Pople’s absolute hardness	Hartree
Shift	Schuurman MO shift alpha	Hartree

		DESCRIPTORS			
No.	Compound	X1	X2	X3	
m	1	PCN-01	~	~	~
	2	PCN-02	~	~	~
	3	PCN-03	~	~	~

	1436	PCDD-74	~	~	~

n

Fig. 2 A schematic representation of the example POP data set used in this chapter with m samples (molecules) and n variables (descriptors)

those errors during the next stages of analysis. Data control also allows us to assess whether any variable transformations (preprocessing) are necessary. In order to use pattern recognition methods presented in this chapter, we must first perform a specific type of data transformation, called standardization, autoscaling, or Z-transformation. Standardization is a way of centering and scaling data in such a way (Eq. 1) that the resulting variables have a mean value equal to 0 and the standard deviation equal to 1.

$$z_i = \frac{x_i - \mu}{\sigma} \quad (1)$$

where x_i is the sample value, z_i is the standardized sample value, μ is the mean value for all samples in the column, and σ is the standard deviation of all the values in the column.

It should be noted that, after this transformation, all the variables have mean value equal to 0 and variance equal to 1 (see Fig. 3). Thusly, the effect of unequal value ranges, caused by differing variable units, and has been nullified, as the standardized data are unitless. Each variable has equal significance (weight) and influence on the analysis results (Livingstone 2009).

Clustering

In order to explore the data, to identify and visualize their underlying structure, and to understand the relationships between objects (samples), we should employ clustering methods. Clustering is a way of looking for natural patterns or groups in the data, in other words, a way of determining the relative positions of

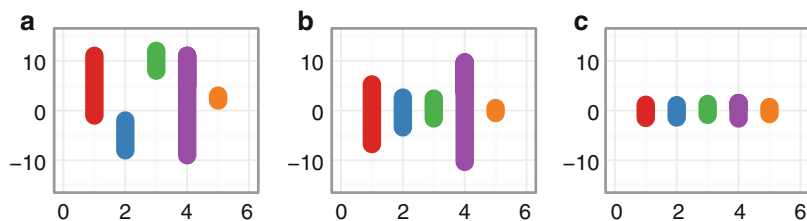


Fig. 3 Data transformation during autoscaling (a) raw data; (b) centered data, where $\mu = 0$; and (c) centered and scaled data, where $\mu = 0$, $\sigma = 1$

all objects in the multidimensional variable space. It is an unsupervised machine-learning method, that is to say, the size and membership of the groups are not known in advance. The groups are formed as a result of a clustering algorithm and are extracted from the data set accordingly to criteria selected by the user.

As all clustering methods are built on the concept of similarity, not only do they allow us to determine group membership but they also help us identify outliers. Ways of measuring similarity will be discussed in depth later in this chapter.

There are two main approaches to clustering. The first one, called *hierarchical*, produces levels of rank-ordered clusters. The second one, called *partitive*, sorts the data into a predefined number of clusters of equal importance (Everitt et al. 2011).

Similarity and Distance

Similarity can be expressed in terms of distance between two objects in the variable space. This space can be one-, two-, three-, or multidimensional, proportionally to the number of variables. It is quite intuitive to interpret similarity and distance as inverse concepts: the greater the distance between objects, the lesser their similarity. Two objects are considered to have similar properties and, consequently, belong to the same group, if the distance between them is sufficiently small (Brereton 2003).

The perception of similarity through distance, the way we interpret objects as similar or not, depends on the distance metric selected, variables used, and data transformation method applied.

Examples of different approaches to defining distances between objects, that is, different *similarity metrics*, are shown in Fig. 4. The *Euclidean distance* between the two objects, $i(1)$ and $j(2)$, denoted by line C, can be calculated according to Eq. 2:

$$D_{\text{Euclidean}} = \sqrt{\sum_{j=1}^J (x_{ij} - x_{kj})^2} \quad (2)$$

where i, j are sample indices, n is the variable number, x_i is the i -th sample value of n -th variable, and x_j is the j -th sample value of n -th variable.

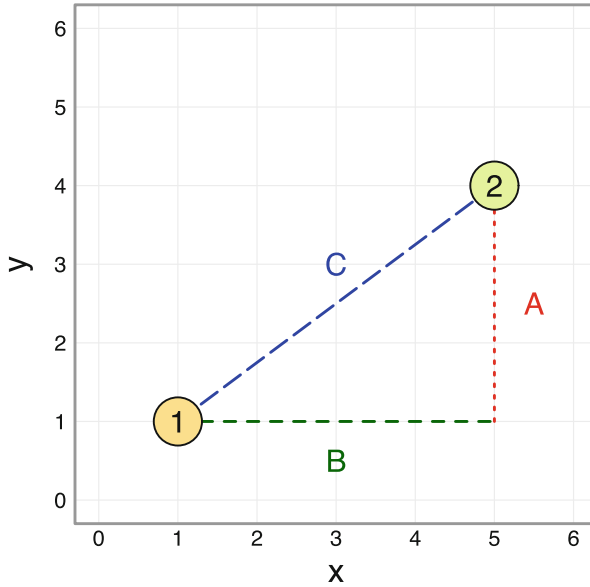


Fig. 4 Different approaches to defining distance metrics

In the presented example, the Euclidean distance between objects 1 and 2 is given as:

$$C = \sqrt{B^2 + A^2}$$

$$C = \sqrt{(5 - 1)^2 + (4 - 1)^2} = 5$$

The second distance metric represented in Fig. 4, known as *Manhattan* or *city block distance*, is calculated as follows (Eq. 3):

$$D_{\text{Manhattan}} = \sum_{j=1}^J |x_{ij} - x_{kj}| \tag{3}$$

In the example:

$$D_{\text{Manhattan}} = 4 + 3 = 7$$

The last distance metric, *Chebyshev*, can be determined according to Eq. 4:

$$D_{\text{Chebyshev}} = \max |x_{ij} - x_{kj}| \tag{4}$$

In our example, $|A| < |B|$; thus,

$$D_{\text{Chebyshev}} = |B| = 4$$

Table 2 Most popular distance metrics

No.	Metric	Equation	Type	Data type
1	Euclidean	$\sqrt{\sum_{j=1}^J (x_{ij} - x_{kj})^2}$	Dissimilarity	Interval/binary
2	Manhattan (city block)	$\sum_{j=1}^J x_{ij} - x_{kj} $	Dissimilarity	Interval
3	Chebyshev	$\max x_{ij} - x_{kj} $	Dissimilarity	Interval
4	Minkowski(order n)	$\sqrt[n]{\sum_{j=1}^J (x_{ij} - x_{kj})^n}$	Dissimilarity	Interval
5	Canberra	$\sum_{j=1}^J \frac{ x_{ij} + x_{kj} }{x_{ij} + x_{kj}}$	Dissimilarity	Interval
6	Mahalanobis	$\sqrt[n]{\sum_{j=1}^J (x_{ij} - x_{kj})^n S^{-1}(x_{ij} - x_{kj})^n}$	Dissimilarity	Interval
7	Cosine(Ochiai)	$\frac{\sum_{j=1}^J x_{ij} x_{kj}}{\sqrt{\sum_{j=1}^J x_{ij}^2 \sum_{j=1}^J x_{kj}^2}}$	Similarity	Interval/binary
8	Pearson correlation coefficient	$\frac{\sum_{j=1}^J (x_{ij} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{j=1}^J (x_{ij} - \bar{x}_i)^2 \sum_{j=1}^J (x_{kj} - \bar{x}_j)^2}}$	Similarity	Interval/binary
9	Squared Pearson correlation coefficient	$\left[\frac{\sum_{j=1}^J (x_{ij} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{j=1}^J (x_{ij} - \bar{x}_i)^2 \sum_{j=1}^J (x_{kj} - \bar{x}_j)^2}} \right]^2$	Similarity	Interval/binary

Distance measures can emphasize similarity, as well as dissimilarity between the objects. The abovementioned metrics highlight the differences between samples and quantify dissimilarity. It should be noted that the choice of the distance measure depends also on the type of data available. A list of the most popular (dis)similarity metrics can be found in Table 2.

The measure of proximity between two clusters, or an object and an already formed cluster, is called *linkage* (Varmuza and Filzmoser 2009). Let us assume there are three clusters called C_1 (with n_1 objects), C_2 (with n_2 objects), and C_3 (with n_3 objects). If clusters C_2 and C_3 are aggregated to form a new single cluster called C_4

Table 3 Linkage types

Linkage type	Equation
Single [nearest neighbor]. Distance between two clusters is the minimum distance between any single observation in one cluster and any single observation in the other cluster	$D_{C_1C_4} = \min(D_{C_1C_2}, D_{C_1C_3})$
Complete [furthest neighbor]. Distance between two clusters is the maximum distance between any single observation in one cluster and any single observation in the other cluster	$D_{C_1C_4} = \max(D_{C_1C_2}, D_{C_1C_3})$
Centroid [unweighted pair group method centroid, UPGMC]. Euclidean distance between two clusters is the distance between the cluster centroids (means, middle points)	$D_{C_1C_4} = \ c_1 - c_4\ $
Average [unweighted pair group method with arithmetic mean, UPGMA]. Distance between two clusters is the mean distance between all observations in one cluster and all observations in the other cluster	$D_{C_1C_4} = \frac{n_2}{n_2+n_3} (D_{C_1C_2}) + \frac{n_3}{n_2+n_3} (D_{C_1C_3})$
McQuitty [weighted pair group method with arithmetic mean, WPGMA]. Distance between the newly formed cluster and any other cluster is the mean value of the distances from each of the two merged clusters to that cluster	$D_{C_1C_4} = \frac{1}{2} (D_{C_1C_2}, D_{C_1C_3})$
Ward [minimum variance]. The proximity of two clusters is calculated as the Euclidean distance between their centroids multiplied by a correction factor, thus minimizing the within-cluster sum of squares	$D_{C_1C_4} = \ c_1 - c_4\ \frac{\sqrt{2n_1n_4}}{n_1+n_4}$

(with $n_2 + n_3 = n_4$ objects), the distance between cluster C_1 and the new cluster C_4 is calculated according to one of the approaches listed in Table 3.

The *single linkage* method is a good choice for identifying the most homogenous (similar) groups. The *complete linkage* method tends to highlight the differences between samples – the resulting groups are similar in size but very diverse. The *furthest neighbor* method can be sensitive to outliers. While the single or complete linkage methods are based on single-pair distances, the *average* and *centroid linkage* methods use a more central measure of location.

Hierarchical Clustering

Hierarchical clustering is an iterative method of grouping objects into a tiered, ordered structure, where all individuals are assigned to a specific, mutually exclusive subgroups. There are two approaches to hierarchical clustering:

- (a) *Agglomerative* (bottom-up), which starts with each of the n objects forming separate, single-member clusters and recursively merges the two most similar

groups or individuals into one of a higher level until all objects have been combined into a single cluster

- (b) *Divisive* (top-down), which starts with all n objects forming one single cluster and with each iteration splits it into two smaller groups until all objects form separate, single-member clusters

The agglomerative methods are widely used in various areas of science: medicine, environmental science, computer vision, and analytical chemistry, which is why the authors choose to focus on them in this section.

Advantages

- Wide range of similarity metrics and cluster linkage techniques
- Applicability to a variety of data types (interval, binary, and count data)

Disadvantages

- Ambiguity in selecting the final number of clusters
- Irreversibility of the merge at each level

Clustering Algorithm

The workflow of agglomerative hierarchical clustering is presented below. In this example, the variables were all measured on different scales; therefore, prior to the analysis, a data transformation is necessary. In order to convert data variables into ones with comparable units, all the sample values were standardized, that is, mean centered and variance scaled (Fig. 5).

Agglomerative hierarchical clustering consists of the following steps:

1. Computing the pairwise distance matrix (Fig. 6). Using a selected metric, the distances between each point and all other objects in the data set are calculated. Here, the Euclidean distance is used.
2. Forming clusters (Fig. 7). The first step is to find two objects with smallest distance to each other (A and B). These objects will be merged into the first cluster (C_1).
3. Determining intercluster distances. Using a selected linkage technique, the relative proximity of the new cluster C_1 to all the remaining objects (C, D, E, and F) is calculated. When estimating linkage, all unmerged objects are considered to be single-member clusters or *singletons*. Here, the single linkage (nearest neighbor) approach has been used.
4. Merging the previously formed cluster (C_1) and the closest object C, into a new cluster (C_2) (Fig. 8).
5. Iterating step 3 thru 4 until all objects have been incorporated into a single cluster with nested subclusters (Fig. 9). In our example, cluster C_1 contains only objects A and B. Cluster 3 contains only objects E and D. One tier up in the hierarchy, cluster C_2 contains cluster C_1 and object C. Up another tier, Cluster C_4 contains cluster C_2 and C_3 . The final tier the cluster C_5 comprises all objects, formed by merging cluster C_4 and object F.

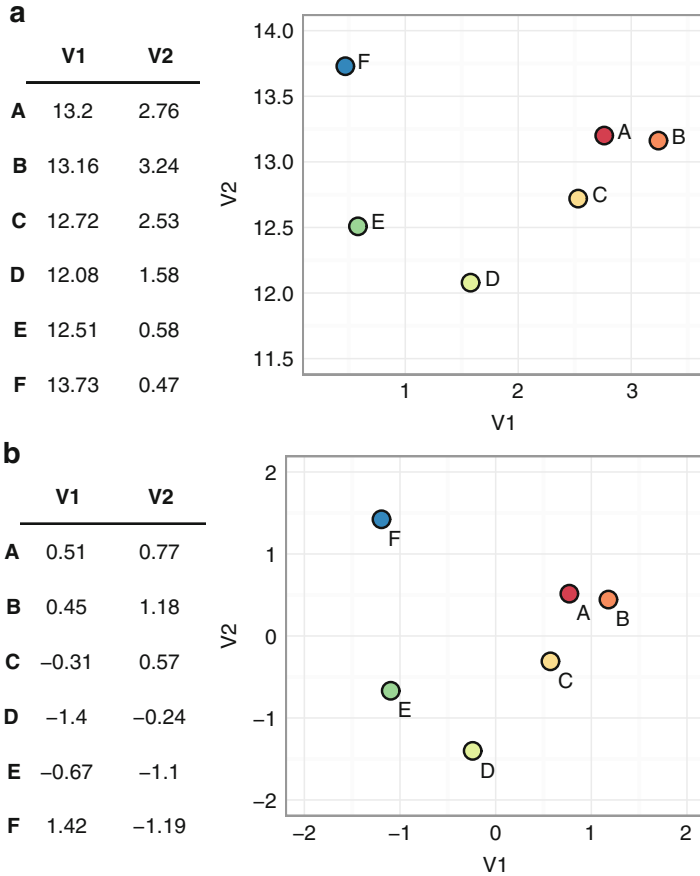


Fig. 5 The effect of autoscaling on a randomly generated data set; (a) original (raw) data, (b) autoscaled data

The hierarchical clustering results can be visualized as a tree-like diagram, called the *dendrogram* plot (Fig. 10). Typically, the y-axis indicates the distance at each successive split (or join) or branch height. On the x-axis, the objects are shown as leaves.

There are three main types of dendrograms to choose from: rectangular, triangular, or circular, although the rectangular is most popular (see Fig. 11). Regardless of the choice of the dendrogram type, the interpretation remains the same. Objects belonging to the same cluster are always more similar to each other than to the objects from other clusters.

Depending on the distance metric and linkage method used, the tree diagram will be different (Fig. 12). For example, using the Canberra distance metric instead of the

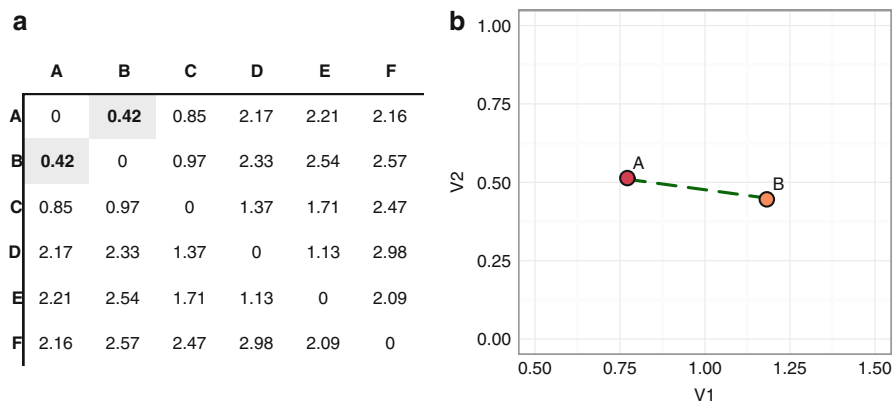
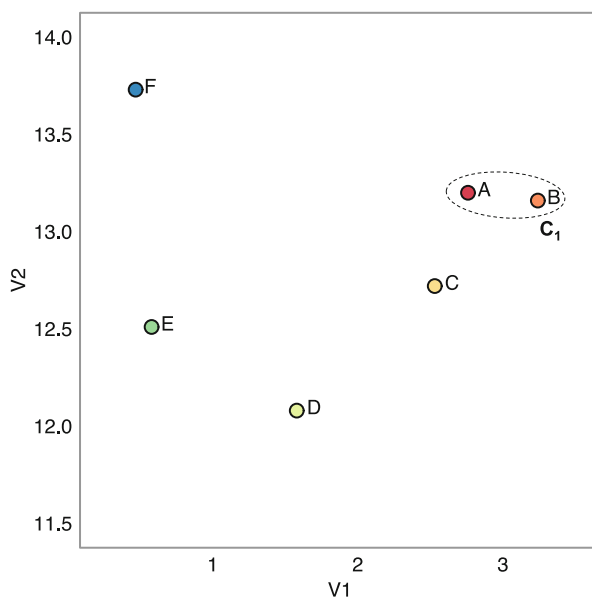


Fig. 6 Example of (a) Euclidean distance matrix, where the smallest distance between the two has been highlighted, (b) visualization of the Euclidean distance between the closest objects – A and B

Fig. 7 Forming the first cluster C_1 by merging two closest objects A and B, based on the single linkage approach



standard Euclidean distance will result in greater distance value ranges and distinct cluster separation; see Fig. 12a.

Selecting appropriate distance and linkage parameters will help answer the questions about the data structure and facilitate the correct interpretation of results. When looking for similarities between samples, one should use the single linkage method instead of the complete linkage approach (Fig. 12b).

Let us now turn to the POP data set. As we are trying to cluster the objects into most diverse groups in order to find the common characteristics for POP samples

Fig. 8 Forming the second cluster by merging cluster C_1 with the closest object C, based on the single linkage approach

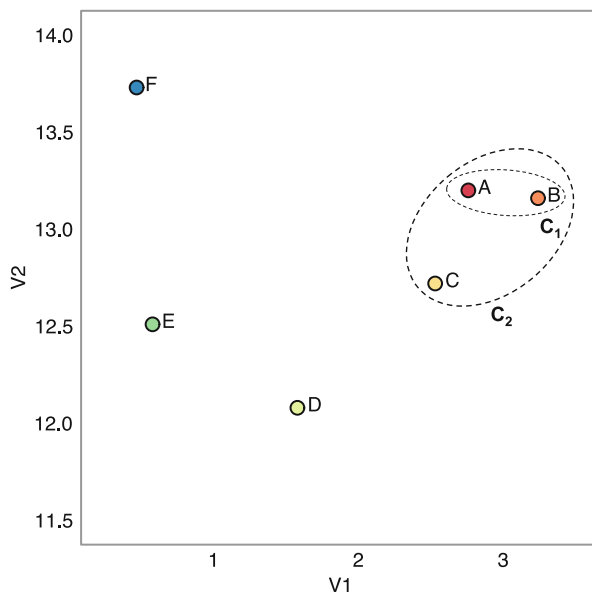
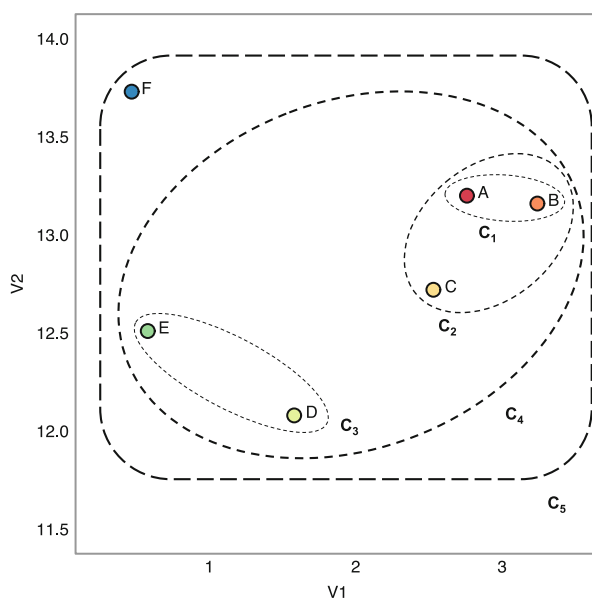


Fig. 9 Final clustering results, with showing hierarchical (nested) cluster structure



from the chemical groups of polychlorinated naphthalenes PCNs, polychlorinated dibenzo-*p*-dioxins PCDDs, and polychlorinated dibenzofurans PCDFs, we used the Euclidean distance metric. For comparison's sake, we applied the three types of linkage metrics to build a dendrogram: complete (Fig. 13a), Ward's (Fig. 13b), and single (Fig. 13c). Depending on the linkage method used, the number of main

Fig. 10 An example of a rectangular dendrogram, created using Euclidean distance and complete linkage

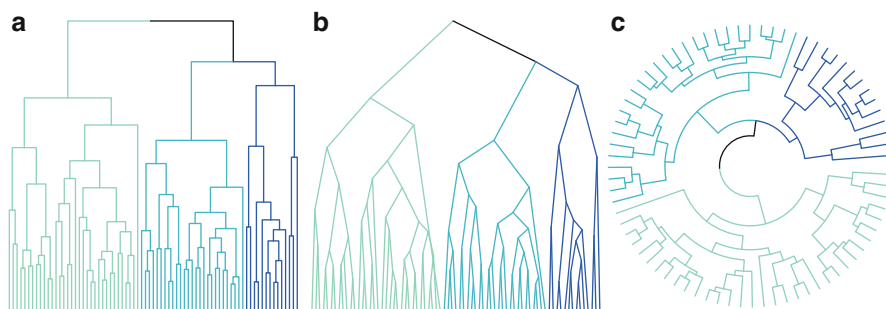
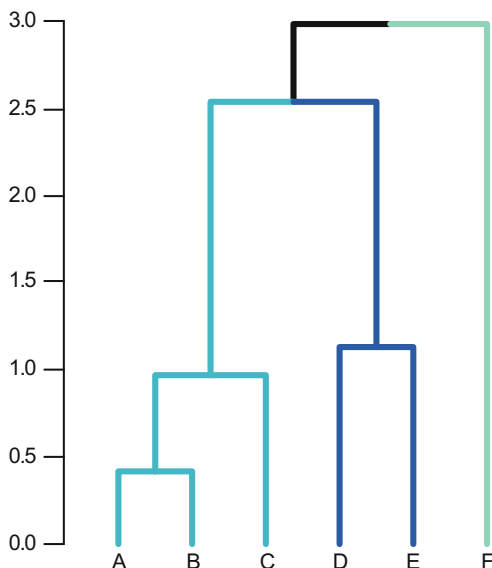


Fig. 11 Different types of dendrograms (a) rectangular, (b) triangular, and (c) circular

branches varies significantly, e.g., for the complete method, four main clusters can be identified; for Ward's method, three main clusters are visible; and for the single linkage method, two main cluster groups are noticeable. The preliminary assumptions on cluster number are very subjective and depend on the person analyzing the plots. Another scientist may see a completely different ramification to the one presented by the authors.

Choosing Number of Clusters

The dendrogram resulting from agglomerative hierarchical clustering does not explicitly specify group structure. Based only on the tree diagram, we are unable to identify the number of groups and their memberships or size. That information can, however, be obtained by “cutting” the dendrogram at a certain height (h) or by specifying the desired number of groups (k). Choosing the cutoff point or a default

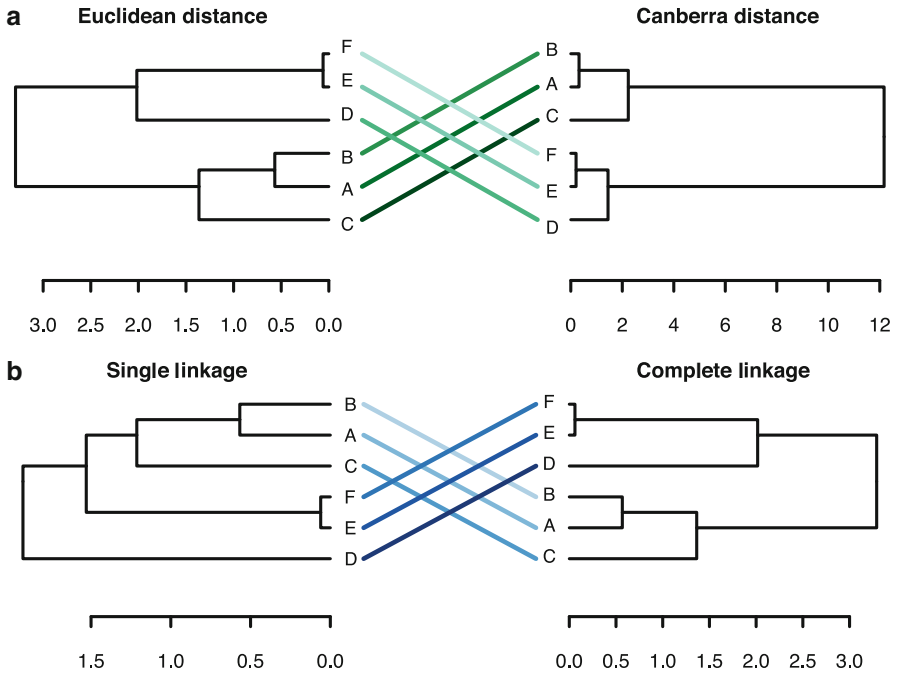


Fig. 12 Comparison of distance metrics and linkage methods: (a) Euclidean versus Canberra distance, complete linkage; (b) single versus complete linkage, Euclidean distance

cluster number is not an easy task, especially considering the variety of data types, units, and ranges, as well as differing needs and expectations of individual analysts.

A substantial number of approaches to determining the branch cut height can be found in the literature (Milligan and Cooper 1985). The simplest one, a *rule of thumb* really, is to use the square root of half of the sample number (Eq. 4):

$$g = \sqrt{\frac{n}{2}}$$

where n is the number of samples and g is the cutoff value.

In order to verify the number of main groups (clusters) identified through different linkage methods, we use the rule of thumb approach. The number of POP samples n equals to 288; therefore, the cutoff height $g = 12$. The resulting groups for all three linkage methods, colored for better identification, are shown in Fig. 14. When using the complete linkage method, after visual inspection, we identified four main clusters. This verified when using the rule of thumb approach, as shown in Fig. 14a. Curiously, the three main clusters initially identified in the second dendrogram (Ward’s method) were later disproved by the rule of thumb, which showed 11 subclusters (Fig. 14b). Interestingly, while we initially identified

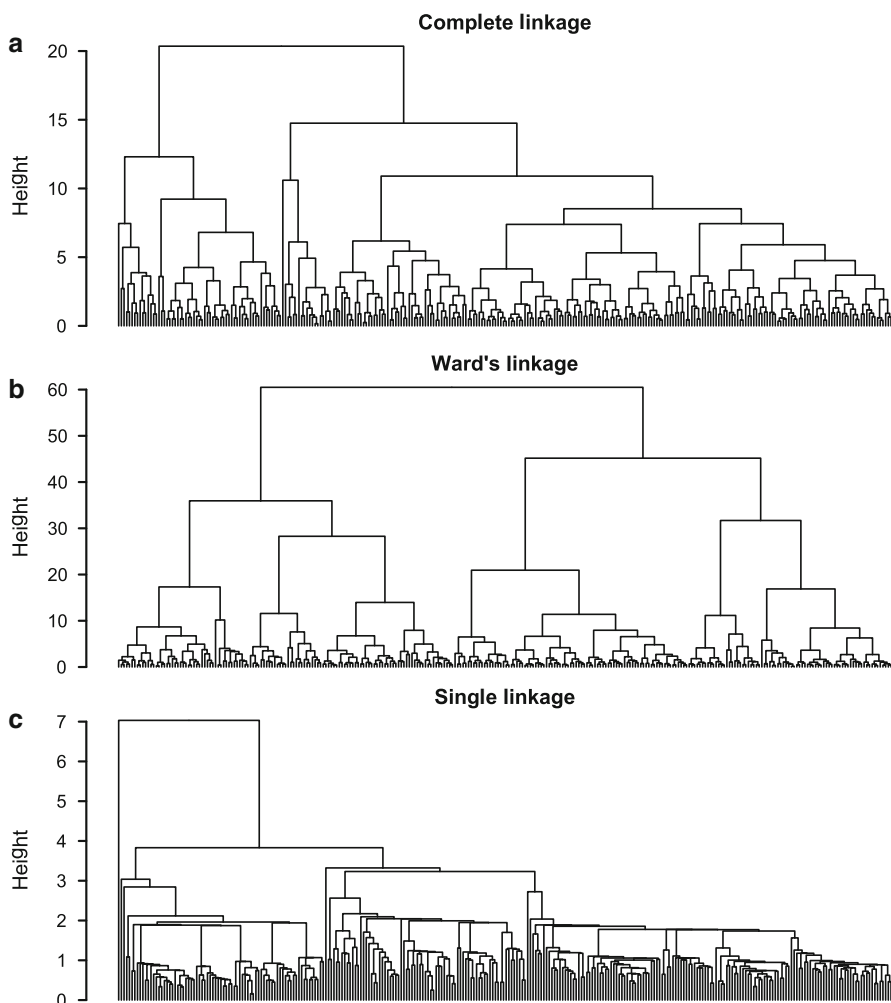


Fig. 13 Hierarchical clustering results for three POP groups (PCNs, PCDDs, and PCDFs) using the Euclidean distance: (a) complete linkage, (b) Ward's linkage method, (c) single linkage

two main subgroups in the third dendrogram (single linkage), according to the rule of thumb, there is only one, all-encompassing cluster (Fig. 14c).

Other, more complex means of selecting the number of clusters are also available. Most of them are automated in nature (i.e., those for which user input is not necessary), such as the C_{index} (Eq. 5):

$$C_{index} = \frac{S - S_{min}}{S_{max} - S_{min}} \quad (5)$$

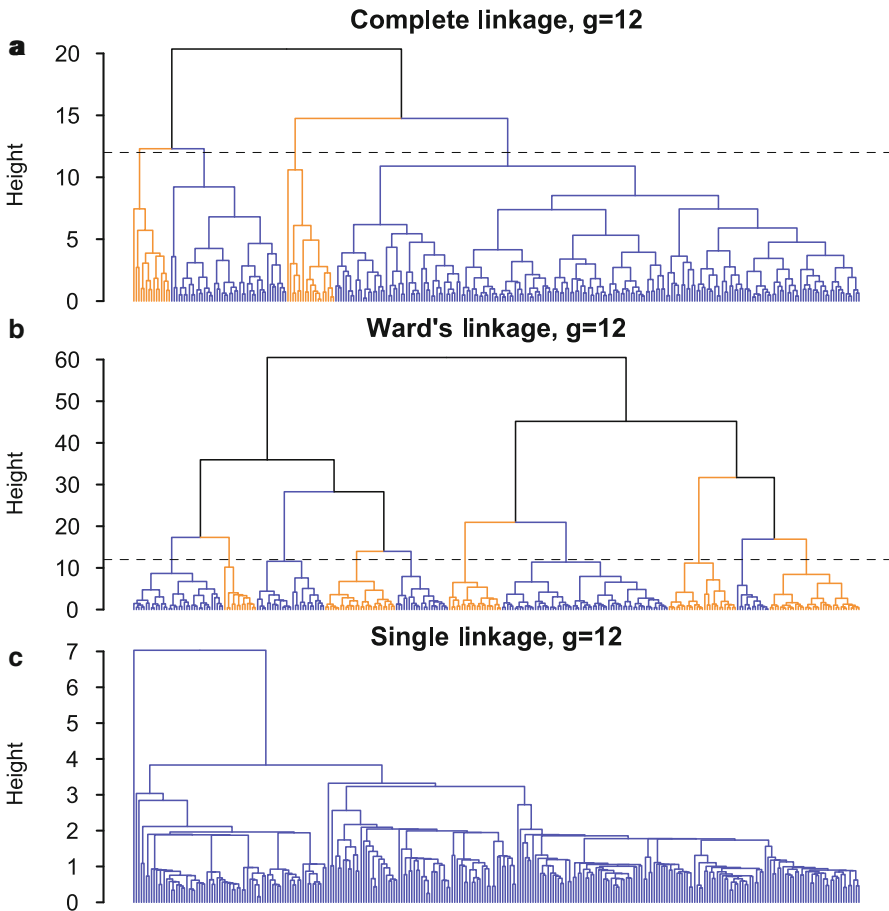


Fig. 14 Varying number of clusters depending on the linkage method used: (a) complete linkage, (b) Ward's linkage method, (c) single linkage. The dashed vertical line shows the cutoff tree height ($g = 12$)

where NC is the total number of pairs of objects from the same cluster, S is the sum of distances for the total number of object pairs in the same cluster (NC), S_{min} is the sum of the NC smallest distances between all the pairs of points in the entire data set, and S_{max} is the sum of the NC largest distances between all the pairs of points in the entire data set.

The C_{index} takes values between 0 and 1, where a small value indicates good clustering. In the case of the POP data set, there are 287 cluster sets ($k = n - 1$ values) possible, ranging from 2 to 287. When comparing C_{index} values for all the possibilities (Fig. 15), we can see that the lowest C_{index}^{287} value is 0, for a grouping consisting practically only of 287 one-element clusters. Such a grouping is, of course, meaningless and unusable. Therefore, in practice, we look for a global

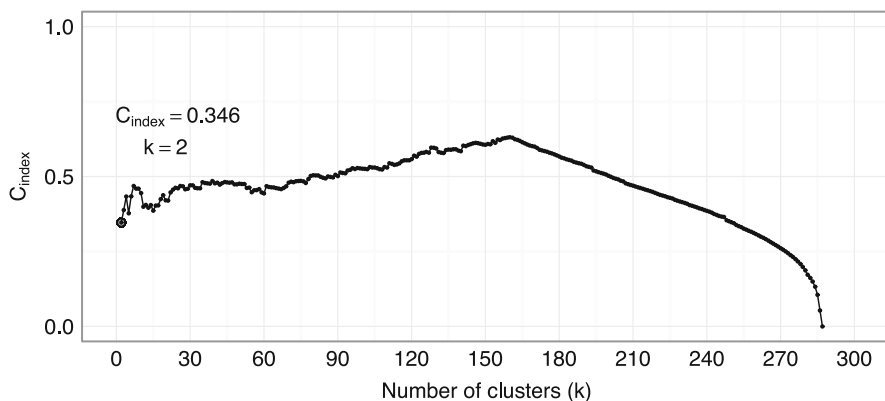


Fig. 15 Number of clusters versus C_{index} value

minimum and the lowest possible number of clusters at the same time. In our case, when taking both of those condition into account, we arrive at $C_{index}^2 = 0.346$, found for $k = 2$. Therefore, two is the optimal cluster number.

Another numerical technique to determine the optimal cluster number is the silhouette index S_i , which is calculated individually for each sample (Eq. 6):

$$S_i = \frac{b_i - a_i}{\max \{a_i, a_i\}} \quad (6)$$

where a_i – is the average dissimilarity of i -th object to all other objects in the same cluster, and b_i – is the minimum of average dissimilarity of i -th object to all objects in other, closest cluster.

The silhouette width S_i can take values between -1 and 1 . For objects with high S_i value (close to 1), the cluster assignment is correct and accurate. For samples with silhouette value close to zero, the cluster assignment is ambiguous and imprecise. That is, the sample may have very well been assigned to another cluster instead, as it is equidistant from both clusters. Objects with S_i , value close to -1 , were wrongly classified (assigned cluster membership) and lie somewhere in between all the possible clusters.

The overall silhouette width is simply the average S_i value over all objects in the data set. When assessing the optimal number of clusters, the one with the maximum average silhouette width should be chosen. The silhouette index should be visualized in the form of a bar graph (silhouette plot), which, in addition to S_i values, shows cluster membership for each object (Fig. 16).

We used the silhouette plot to verify and compare the clustering for two different linkage methods: complete (Fig. 16a) and Ward's (Fig. 16b). In the case of complete linkage, 71 out of 288 POP samples (24.65% of the data) have S_i values below 0, meaning they could be considered as "misclustered" samples. Such a high number

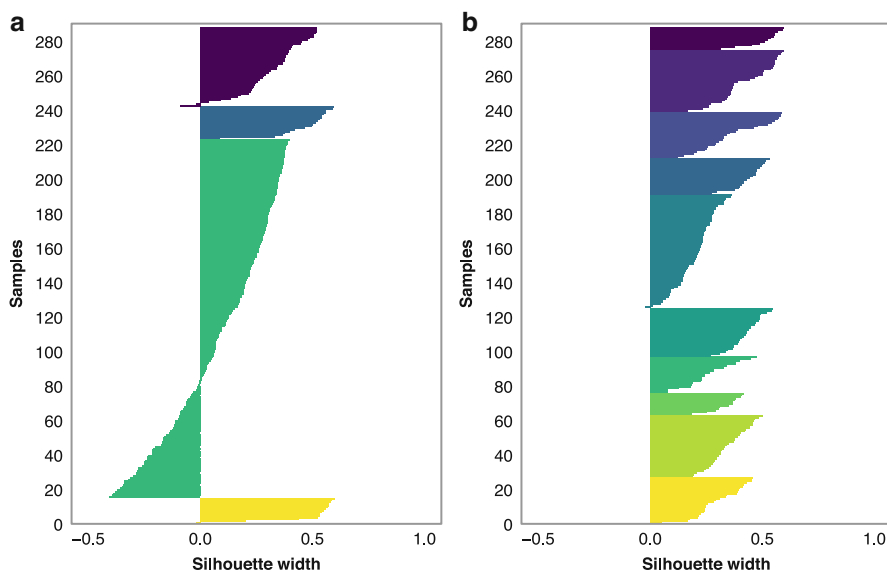


Fig. 16 Silhouette plots for three POP groups (PCN, PDDD, and PCDF). Comparison of the cluster number resulting from cutting dendrograms at height $g = 12$; (a) complete linkage, (b) Ward's linkage method. Groups have been color-coded for easier interpretation

of samples with negative silhouette values may suggest that the selected linkage method may not be the most suitable for the data at hand. This conclusion is further confirmed when looking at the silhouette plot for Ward's method (Fig. 16b). There is only one POP sample with the S_i value below 0, suggesting that this linkage method gives a more accurate separation of samples (grouping). It is worth noting that, if a data set contains a significant number of samples with S_i values below 0, it may also suggest that the selected features (variable) do not describe the underlying data structure in an appropriate manner.

There is also another way of comparing the cluster membership with the original POP group, illustrated in Fig. 17. Cluster 1 consists of several PCNs and two PCDFs. Cluster 2 contains the bulk of the data set, that is, the majority of samples from all three POP groups (PCNs, PCDDs, and PCDFs). Cluster 3 and 4 contain a small number of PCDDs and PCDFs. We see now why the silhouette analysis yielded 71 "misclustered" samples. The reason for such misclustering might be the use of too many variables (in this case, all 21 molecular descriptors), some of which might be redundant.

Despite the great number of cluster selection methods available, there is no single, universally applicable technique regarded as a golden standard. The silhouette index is a potential candidate for an effective and simple way of verifying the quality and performance of the selected clustering algorithm. Sadly though, even the most sophisticated cluster selection and identification method needs to be augmented by

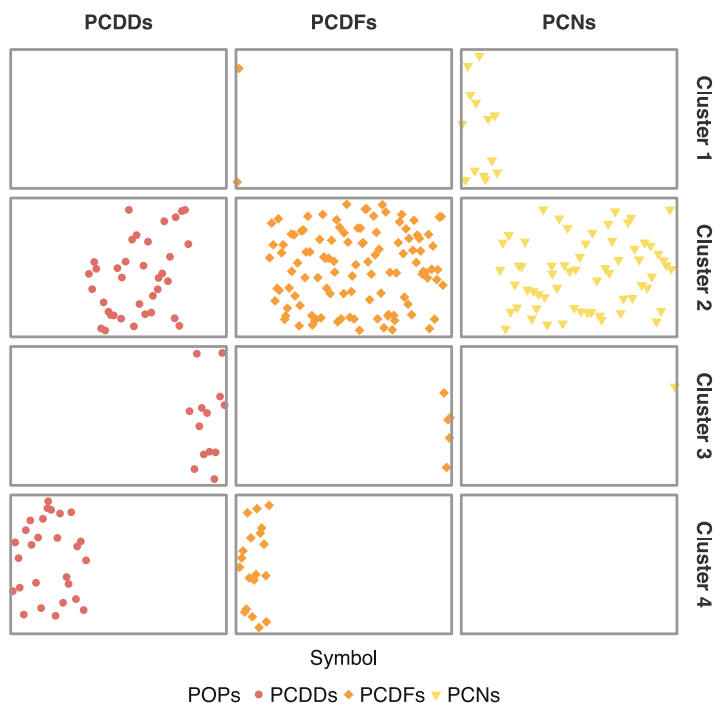


Fig. 17 Hierarchical clustering grouping versus the original chemical groups. Euclidean distance and complete linkage method were used

user expertise and experience. No two research problems are alike; therefore, no universal, unambiguous approach to all of them can be found.

***k*-means**

The *k*-means algorithm is one of the most popular partitioning methods, mainly because of its ease of use. It groups data into a user-defined number of clusters, *k*. Each observation is assigned to the nearest cluster, the center of which is defined by the arithmetic mean (average) value of its points, the centroid (Myatt 2007). Among many available distance metrics (see Table 1) used in *k*-means, the Euclidean distance is the most popular one. The main advantage of this method is its clarity and simplicity, making its implementation an easy task.

Partitioning Algorithm

The algorithm consists of the following steps:

1. Setting up initial positions of the *k* cluster centers. This can be by supplying a set of predefined centroid coordinates or by random generation of cluster centers.

2. Calculating distances from each i -th element to each of the initial cluster centers using a selected distance metric.
3. Based on sample-center distances, assigning each i -th element to the nearest cluster.
4. Calculation of new cluster means (centroids) after object reassignment.
5. If the distance from the sample to the new cluster center is smaller than the distance to the prior cluster center, assignment of the sample to a new cluster with the closer centroid.
6. Iterating steps 2 thru 5 until permanent cluster assignment for all elements, i.e., until no element changes its cluster membership. It is possible to define the maximum number of iterations regardless of group reassignment.

When applying the k -means method to the POP data set, we observed some cluster overlap, as was the case with HCA results (Fig. 18).

In order to compare the clustering results from both the hierarchical and partitioning method, we calculated the silhouette for each sample (Fig. 19). We found that the overall performance of the k -means algorithm was better in terms of cluster homogeneity; only 7 out of all 288 samples (2.43 % of all data) have S_i below 0, which means only seven of them have been “misclustered.” It is a great improvement over the HCA clustering, which yielded 71 “misclustered” samples.

All partitioning methods divide the data into a set number clusters in a way that optimizes an objective function, that is, meets specific criteria set for the results prior

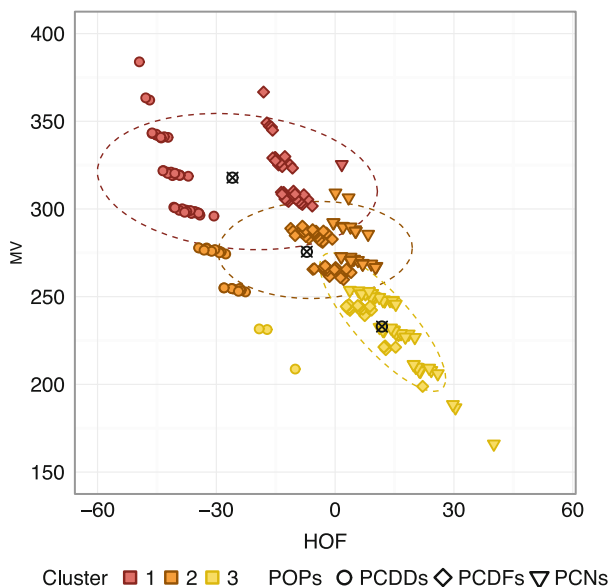


Fig. 18 The results of k -means clustering for selected POP groups (PCNs, PCDDs, PCDFs)

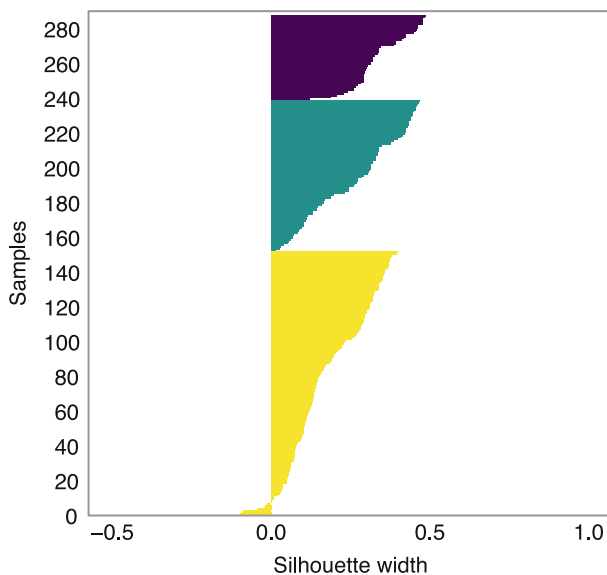


Fig. 19 The silhouette plot of k -means clustering results for the selected POP groups (PCNs, PCDDs, PCDFs)

to the analysis. In case of k -means, the algorithm minimizes the sum of squared errors, SSE (Han et al. 2012), the within-cluster sum of squared distances between each sample and cluster center (Eq. 7).

$$\text{SSE} = \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - c_k\|^2 \quad (7)$$

where k is the number of clusters, C_k is the k -th cluster, c_k is the centroid of the k -th cluster, and x_i is the sample.

It is important to note that a different choice of initial centroid positions may result in a different final assignment of points to clusters. Therefore, it is advisable to run the k -means algorithm a number of times, each with a different set of starting parameters, and then compare the results. This will ensure that the algorithm will reach the global optimal cluster configuration and will not get “stuck” at a local minimum of the objective function.

Advantages

- Effective when working with well-separated, compact clusters
- Computationally faster and more appropriate for large data sets than agglomerative hierarchical clustering
- Intuitive use and ease of interpretation, simplicity of implementation

Disadvantages

- The number of groups must be prespecified prior to running the algorithm
- Not robust to noisy data and outliers – might create nonoptimal clusters (solution: use k -medoids)
- Not applicable for categorical data (solution: use k -modes)

k -medoids

While the k -means algorithm is a very simple and useful tool, it is quite susceptible to extreme values or noise, which distort the mean value of the cluster. The k -medoids algorithm is a modified version of the k -means approach, where rather than averaging all objects in the cluster to find the center, the middle point, or medoid, is selected from among the samples in the cluster (Everitt et al. 2011). Thusly, the clusters are represented by their most centric element, instead of a numerical value that may not belong to the cluster at all. Medoids are defined as data points, whose average dissimilarity to all the objects in a cluster is minimal.

Both the k -means and k -medoids algorithms minimize the distance between all points inside a cluster and its designated center. The difference being that, unlike the k -means algorithm, k -medoids work with pairwise dissimilarities instead of Euclidean distances.

The k -medoids algorithm is analogical to that of the k -means method:

1. Selection of an initial set of k -medoids among the n objects in the data set.
2. Assignment of each object in the data set to the nearest (least dissimilar) medoid.
3. Searching for more optimal medoids. Iteratively for every medoid m , swapping (replacing) it with each one of the non-medoid objects o_i to check if it improves the total distance of the resulting clustering.

The suitability of the medoid “candidate” o_i is assessed by computing its average dissimilarity to all remaining non-medoid data points $o_{n \neq i}$. The objective function is the total cost of the new configuration, which is the sum of candidate-no-medoid differences E (Eq. 8).

$$E = \sum_{i=1}^n d(o_i, m) \quad (8)$$

where d is the dissimilarity measure, o_i is the non-medoid object, m is the medoid, and n is the number of objects in the data set.

4. Selection of new medoids o_i with the lowest cost of the configuration (Fig. 20).

k -modes

The k -means clustering algorithm cannot be applied to categorical data, as it relies on means to represent cluster centers, and nonnumerical data have no defined mean value. The k -modes, a variant of k -means, enables grouping categorical data by using the mode as the cluster center point. Moreover, it replaces the Euclidean

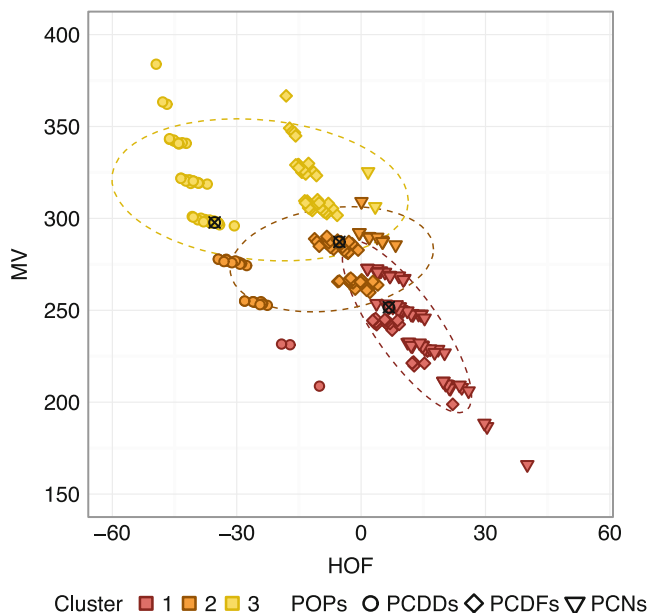


Fig. 20 The results of k -medoid clustering for the selected POP groups (PCNs, PCDDs, PCDFs)

distance metric from k -means with matching dissimilarity measure and a frequency-based approach to update cluster modes.

The k -modes algorithm consists of the following steps (Khan and Kant 2007):

1. Selecting k initial modes, one for each of the cluster.
2. Assigning every data object to the cluster whose mode is nearest to it. The proximity (dissimilarity) measure is based on the total number of mismatches d (Eq. 9). The smaller the number of mismatches, the more similar the two elements are.

$$d(X, Y) = \sum_{j=1}^m \delta(x_j, y_j) \quad (9)$$

$$\begin{cases} 0, & x_j = y_j \\ 1, & x_j \neq y_j \end{cases}$$

where X and Y are two categorical objects and m is the number of categorical attributes; x_j, y_j are categories for attribute j .

3. Computing new modes for all clusters.
4. Repeating step 2 thru step 3 until no data object has changed its cluster membership (Fig. 21).

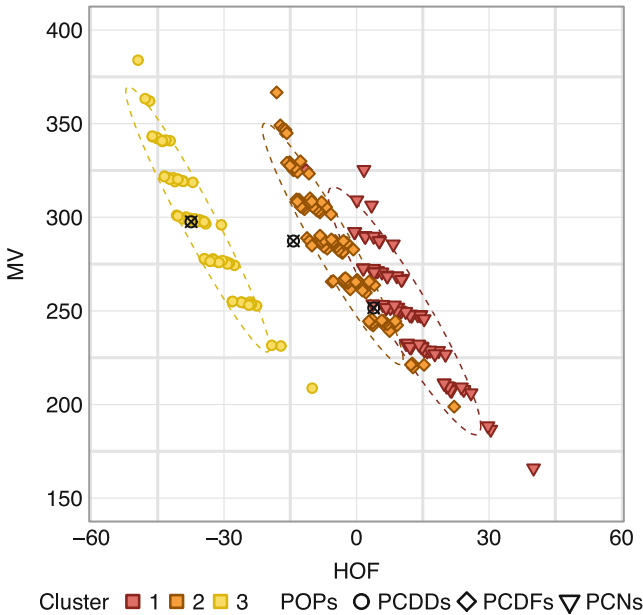


Fig. 21 The results of *k*-modes clustering for the selected POP groups (PCNs, PCDDs, CDFs)

PCA

When faced with a large, multidimensional set of variables, the sheer volume may make it difficult to see patterns and relationships hidden in its structure. Principal component analysis allows us to visualize and describe this structure by means of reducing the dimensionality of the data. This method is based on principal components (PCs), which are linear combinations of the original variables (PCs) and form a new multidimensional space onto which the original data set is projected.

The main applications of principal component analysis are:

- Projecting multidimensional (multivariate) data onto 2D and 3D scatterplots
- Wining out important information from data noise
- Converting data with high inter-variable correlation into a set of uncorrelated latent variables suitable for predictive modeling

In data mining, information content is measured in terms of data variance (Eq. 10).

$$\sigma^2(x) = \frac{\sum_{i=1}^m (x_i - \mu)^2}{n - 1} \tag{10}$$

where σ is the variance of variable x , μ is the mean value of variable x , n is the number of samples, and x_i is the i -th sample of variable x .

Variables with very little difference in values (low variability) are in fact non-informative and have little bearing on the clusters present in the data set. These variables are inadequate to the task of describing and explaining the underlying data structure, as they have very little influence on it.

In order to describe the linear relationship of two variables, we can calculate their covariance:

$$\text{cov}(x_k, x_l) = \frac{\sum_{j=1}^n (x_{jk} - \mu_k)(x_{jl} - \mu_l)}{n - 1}$$

where x_k and x_l are k -th and l -th variables, μ_k and μ_l are variable means, and n is the number of samples.

The covariance parameter is the measure of how much the values of two variables change together. If values of one variable change (e.g., increase) in the same direction as values of the other, i.e., the variables tend to show similar behavior, the covariance value is positive. Although, the covariance parameter, while a good indicator of whether two variables correspond to each other, is not suitable for a pairwise comparison of multiple variables with differing units and value ranges. In other words, the same numerical value of covariance might mean a highly proportional relationship for one pair of variables, whereas for another pair, it might signify a completely random one.

In order to unambiguously assess and compare the relationships between variables, we must calculate the correlation coefficient r (Eq. 11):

$$r(x_k, x_l) = \frac{\text{cov}(x_k, x_l)}{\sqrt{\text{var}(x_k)}\sqrt{\text{var}(x_l)}} \quad (11)$$

where x_k , x_l are k -th and l -th variables, σ_k^2 , σ_l^2 are variable variances, and cov is variable covariance.

The Workflow

Prior to the analysis, the data matrix \mathbf{X} requires preprocessing and the preferred data transformation methods is standardization (autoscaling). In the resulting autoscaled matrix \mathbf{Z} , all the variables have mean value μ equal to 0 and the standard deviation value σ equal to 1. Thus, any distorting effects caused by differing variable units have been negated.

The principal component analysis consists of the following steps (Jolliffe 2002; Brereton 2009):

1. Calculating the correlation-covariance (*corr-cov*) matrix \mathbf{C} . The diagonal elements of a covariance matrix are the variable variances, and the non-diagonal

elements are their covariance. For standardized (autoscaled) data, the covariance matrix is equivalent to the correlation matrix, and the diagonal elements are in fact the Pearson's correlation coefficients (Eq. 12):

$$\begin{aligned} \text{var}(x_k) &= 1 \\ \text{var}(x_l) &= 1 \\ r(x_k, x_l) &= \text{cov}(x_k, x_l) \end{aligned} \tag{12}$$

where x_k, x_l are k -th and l -th variables, cov is the variable covariance, and r is the variable correlation coefficient.

2. Computing the eigenvalues and eigenvectors of the correlation-covariance matrix. The eigenvalues λ indicate the amount of variance in data explained by their corresponding principal component. Each eigenvector contains a set of variable coefficients, i.e., loadings, \mathbf{P} . The resulting eigenvectors \mathbf{P} are arranged in descending order of their eigenvalues. The loadings express the variables' contributions to the principal components.
3. Calculating the scores. In order to obtain the positions of all data objects in the new principal component space, we must calculate the principal component scores, \mathbf{T} . The scores matrix is a product of autoscaled matrix \mathbf{Z} and the eigenvector matrix \mathbf{P} .
4. Determining the number of significant principal components, k . All principal components are mutually orthogonal – each one contains unique information that none of the others represent. Yet not all of the PCs are equally important. Considering that they are sorted in a descending order of information content, only a first selected few principal components will be useful in representing the data set.

The essence of the mathematical operations performed in order to obtain the principal components can be distilled into the scheme shown in Fig. 22. The rows in matrix \mathbf{Z} represent samples and the columns represent variables. The product \mathbf{TP} is an approximation of to the original data set, i.e., a model, the error of which is being represented by matrix \mathbf{E} .

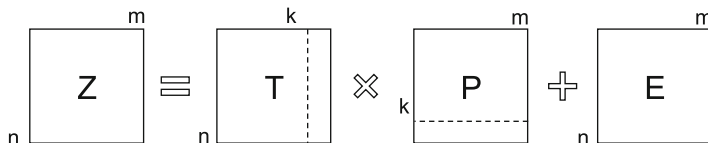


Fig. 22 A schematic of PCA decomposition. k is the number of significant principal components, \mathbf{Z} is the $n \times m$ standardized data matrix, \mathbf{T} is the $n \times k$ scores matrix, \mathbf{P} is the $k \times m$ loadings matrix, and \mathbf{E} is the $n \times m$ error matrix

Table 4 Eigenvalues and explained variance of the principal components – POP data set

PC	Eigenvalue	Explained variance [%]	Cumulative explained variance [%]
1	11.36	54.07	54.07
2	5.75	27.40	81.48
3	2.25	10.70	92.18
4	0.93	4.42	96.59
5	0.44	2.07	98.67
6	0.11	0.53	99.20
7	0.08	0.40	99.60
8	0.06	0.28	99.88
9	0.02	0.07	99.96
10	0.01	0.02	99.98
11	0.00	0.01	99.99
12	0.00	0.00	100.00
13	0.00	0.00	100.00
14	0.00	0.00	100.00
15	0.00	0.00	100.00
16	0.00	0.00	100.00
17	0.00	0.00	100.00
18	0.00	0.00	100.00
19	0.00	0.00	100.00
20	0.00	0.00	100.00
21	0.00	0.00	100.00

Let us review the results of the principal component analysis of the POP data set. There are various methods of selecting the significant number of principal components (Jolliffe 2002):

1. The Kaiser criterion, which states that the significant PCs have eigenvalue greater than or equal to 1. In our case (Table 4), we would choose the first three principal components (PC1 thru PC3).
2. The minimum cumulative variance criterion is PC significance only for values over a certain arbitrary threshold, here 90%. Again, according to this rule, we should use the first three principal components (PC1 thru PC3, Table 4).
3. The “elbow method” based on the scree plot (Fig. 23a). The “elbow” in question is the point of the plot where the line reaches a plateau. We would make use of the first four principal components (PC1 thru PC4).

We choose to obey the Kaiser criterion, that is, to consider PC1, PC2, and PC3 as the most significant.

The next step is to visualize the data set in the space created by the principal components (Fig. 24), i.e., the *scores*. The principal component scores show the position of all samples projected onto the PC hyperspace. When analyzing the scores for the first three principal components, we noticed that PC1 and PC2 separate the

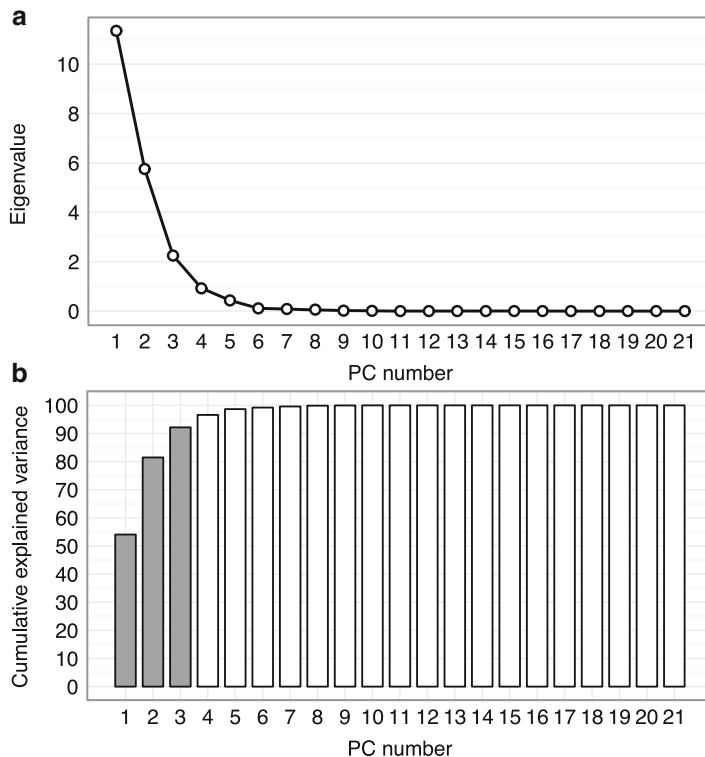


Fig. 23 Explained variance of principal components for selected POP groups (PCNs, PCDDs, PCDFs): (a) scree plot of the principal components (b) cumulative explained variance

three selected POP groups (PCNs, PCDDs, PCDFs) quite well (Fig. 24a), whereas PC1 in combination with PC3 (Fig. 24b) seems to highlight possible outliers. When juxtaposing PC2 and PC3, we see highly condensed group made up of PCN and PCDF samples as well as highlight one possible outlier (Fig. 24c).

Knowing that each principal component is a linear combination of all variables (descriptors) in the data set, we can use loadings to investigate which variables influence them most. A variable is considered to be highly influential on the principal component if the absolute value of its standardized loading is greater than 0.7 (Jolliffe 2002). What this means is that in reality, only the “highly influential” variables dictate the internal data structure.

According to PC1’s loading vectors presented in Fig. 25, the first principal component is determined by the size and bulk of the molecules. The contributing structural descriptors (nX, MW, MV, and SAS) are all correlated with PC1, that is the greater the PC1 values, the higher the number of chlorine atoms, molecular mass, volume, and surface. The polarizability (P) of a molecule and the core-core repulsion energy (Core) between two atoms are proportional to molecular size, which is expressed by the high loading values of these descriptors. Another characteristic dependent on the size of the molecular is its energy. Here, the proportion is inverse,

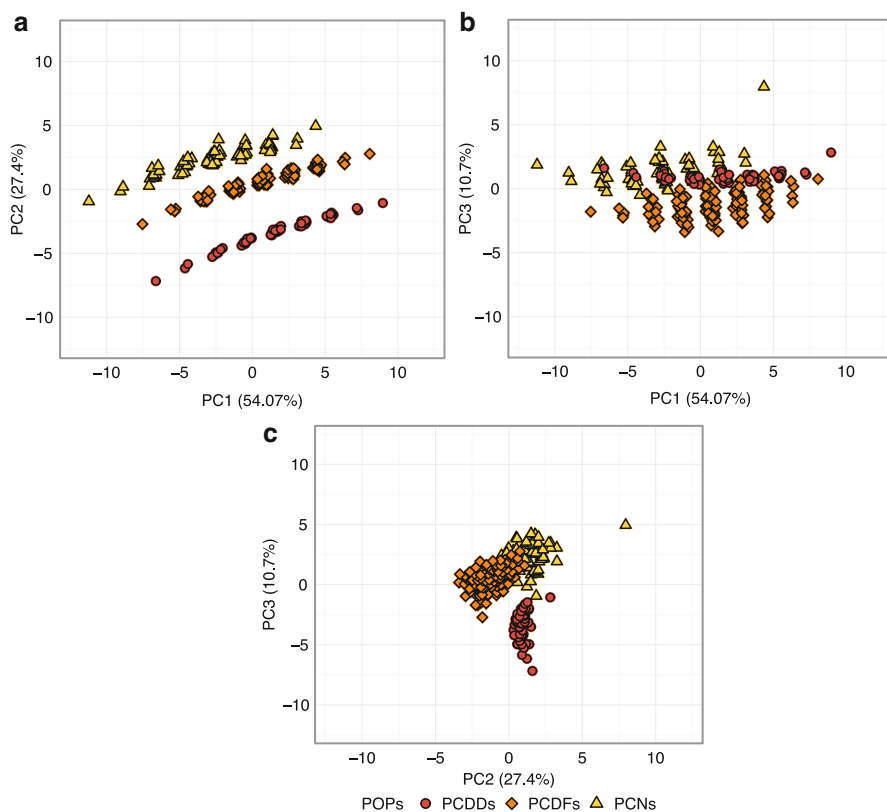


Fig. 24 Scatterplots of the first three principal component scores (PC1, PC2, PC3) – POP data set

as values of energy and heat are expressed in negative numbers, so the bigger the molecule, the lower (“more negative”) the energy value. This is expressed through a set of descriptors: the heat of formation (HoF), the electronic energy (EE), the total energy (TE) and the total energy of the corresponding cation (TEp), and the LUMO energy and Parr and Pople’s absolute hardness (Hard). All of them have negative loading values and are inversely correlated with PC1.

PC2 is determined by the molecule’s ability to donate and accept electrons. The higher the number of oxygen atoms (nO) and the overall number of atoms in the molecule (nAT), the more difficult it is to detach an electron, which is expressed by the inverse proportion between those two descriptors and the vertical ionization potential (VIP). VIP itself is calculated as the difference between the energy of a neutral molecule and the energy of the corresponding cation. According to the Koopmans’ theorem, the negative value of HOMO can be used as an approximation of the first ionization energy, so in fact, both descriptors express the same molecular feature. Conversely, the bigger the number of electronegative atoms (i.e., oxygen) in the molecule, the greater the overall electronegativity (EN). The chemical shift (Shift) is the equivalent of the electronegativity but with an opposite sign.

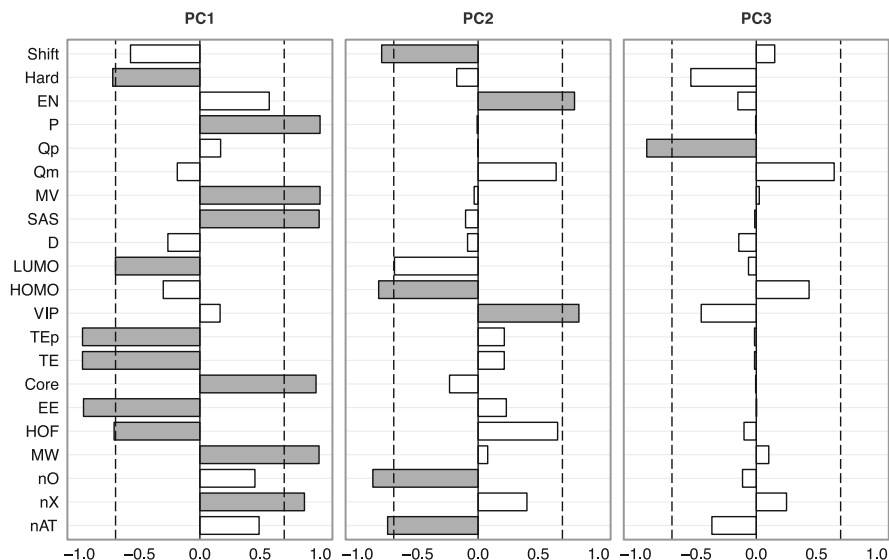


Fig. 25 First three principal component loadings (PC1, PC2, PC3) – POP data set

PC3 represents the symmetry of a molecule, expressed through the highest positive Mulliken charge on the molecule (Qp). The more symmetrical the structure of a molecule, the lower the partial charge.

SOM

Self-organizing (or Kohonen) maps, SOMs (Kohonen 2001), are a type of unsupervised partitive clustering methods. They are a quick and easy way of simultaneous data grouping and visualization, inspired by the biological information transport and processing system – neurons and synapses. SOMs are an ideal tool for analyzing large, complex data sets for which the standard scatterplot would be unreadable due to overlapping data points, the so-called overplotting.

SOMs preserve the topology, i.e., the relative distances between objects, of a high-dimensional variable space while mapping it onto a low-dimensional representation, usually a two-dimensional grid or plane.

Building SOMs

Determining the Structure

A Kohonen neural network consists of two layers (Kohonen 2001; Vesanto and Alhoniemi 2000):

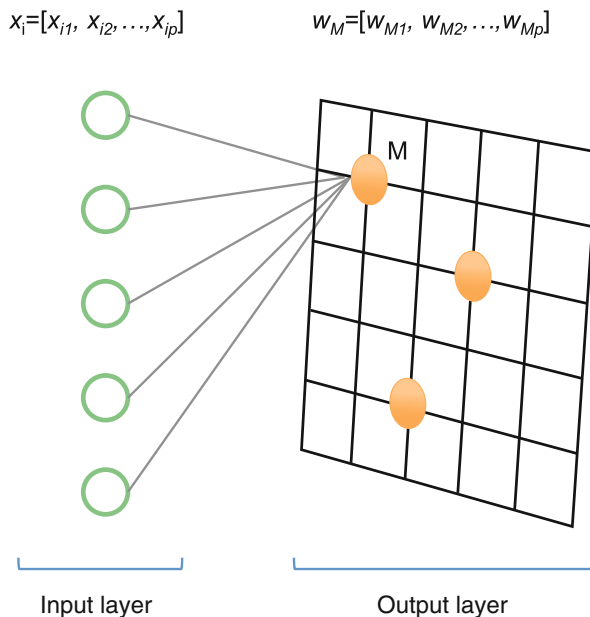


Fig. 26 A schematic representation of a SOM network

1. *Input layer* – an array of input vectors, each of length p , based on the original data matrix; every sample (object) is represented by an input vector containing its coordinates in the original p -dimensional variable space (Fig. 26).
2. *Output layer* – a grid with M nodes (neurons), where M is a number of nodes determined by the user; each node is represented by a *weight vector*, or *codebook vectors*, of length p , containing *weights*, i.e., values defining its location in the original p -dimensional variable space (Fig. 26).

In order to set up the computations, one must first define and select a number of parameters describing the shape, topology, and size of the network:

- (i) Number of neurons in the output layer
- (ii) Network shape: planar (1D line, 2D square or rectangle), cylindrical, and toroidal
- (iii) Neighborhood type: square, rectangular, or hexagonal
- (iv) Neighborhood radius
- (v) Neighborhood function: rectangular, Gaussian, cut Gaussian, triangular, and exponential
- (vi) Learning rate
- (vii) Number of iterations during the training process

All of these parameters influence the network's learning capabilities, quality of grouping, and computation time.

The *number of neurons* (i) defines the number of grid nodes present in the output layer. The greater the number of neurons in the network, the more complex problems it can solve and sharper the divide between resulting groups (clusters) – at the cost of computational resources. Choosing too many neurons may mean that a part of them will go unused when applied to low-complexity problems. Conversely, employing an insufficient number of neurons produces overlapping clusters and loss of information about class structure (Gemperline 2006).

The *network shape* (ii) describes the arrangement of the nodes in the network. The simplest possible shape resembles a 1D chain, where each link is a node. The more complex form can be a 2D map (arranged in a square or rectangular grid), 3D cylinder to a toroid. To ensure the equal number of neighbors for every neuron, it is best to use infinite, edgeless, toroidal form (Brown et al. 2009).

The *neighborhood type* (iii) is a method of counting the neighboring nodes. A square neighborhood means that a neuron has four nearest neighbors, a hexagonal neighborhood means that the neuron has six nearest neighbors, and a rectangular neighborhood means that a neuron has nine nearest neighbors.

The *neighborhood radius* (iv) is the number of surrounding nodes associated with the *winner neuron* or the *best matching unit*, *BMU*. This group of neurons, along with the BMU, is called a *neighborhood set*. The radius value decreases during each training cycle and is ultimately limited only to the winning neuron (Maimon and Rokach 2005).

The *neighborhood function* (v) controls the intensity of change of the weights in the codebook vectors from the neighborhood set, that is, the vectors closest to the winner neuron. The most popular is the Gaussian function, through which the winner neuron's weight adjustment is greater than that of the remaining neurons (Kohonen 2001; Brereton 2009; Brown et al. 2009).

The *learning rate* (vi) is responsible for the network's performance. It's a coefficient which can take values between 0 and 1 and determines how similar the neuron weight vectors will be to the *input pattern* or *sample vector*. The sample vector is a randomly chosen object (sample) from the original data matrix, serving as a reference point for the training vector at each iteration. The initial value of the learning rate decreases with each iteration, the increment determined by a linear or exponential function. Choosing a too small number may result in an insufficient weight correction and therefore limited adjustment of codebook vectors in the p -dimensional space. As a consequence, some of the map units may never get a chance to fully learn the input pattern and sufficiently represent all the samples (Brereton 2009; Brown et al. 2009; Hastie et al. 2009).

The *number of iterations* (vii) defines the amount of repetitions in the training cycle. Too small a number may result in an undertrained network, which in turn leads to incorrect clustering. On the other hand, a too large number may lead to waste of computational time and resources on redundant cycles when an optimal solution had been found a few repetitions prior. The recommended number of iteration samples is the number of map units multiplied by a factor of 500 (Brereton 2009).

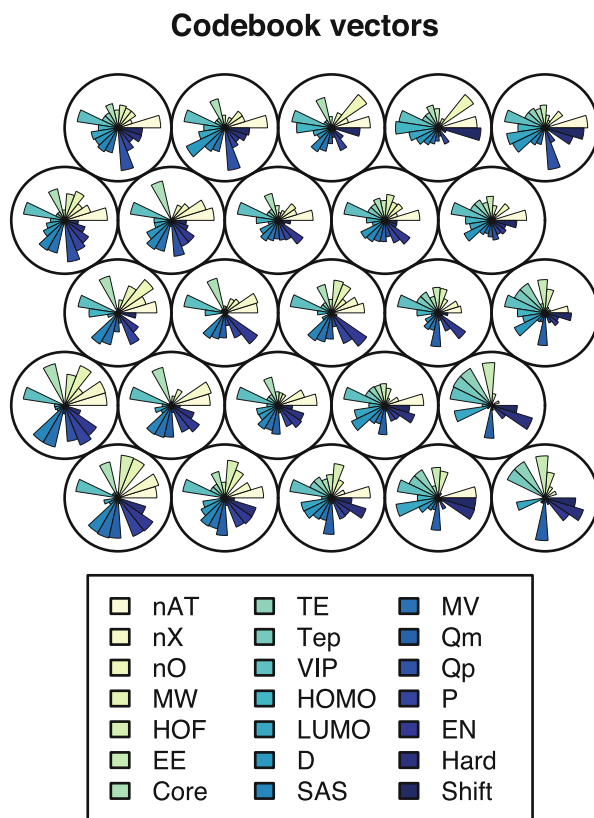


Fig. 27 Optimal weight vectors of the output neurons (*codebook vectors*) for the POP data set

Training

After establishing all the necessary parameters, it is time to begin training the network, that is, performing an iterative search of optimal weight vectors of the output neurons (Fig. 27). The steps of this iterative process are as follows (Vesanto and Alhoniemi 2000; Brereton 2009; Brown et al. 2009):

1. Initializing the network – selecting the initial weights for M neuron. The most common method is a random assignment of a value from a min-max range. The number of weights in the codebook vector is equal to the number of variables (features) in the original data set.
2. Selecting, usually at random, a single input vector.
3. Calculating the distances between the input vector and the neurons described by the codebook vector in the p -dimensional variable space. The choice of distance metric is left to the user – a list of the most commonly used one can be found elsewhere in this chapter.

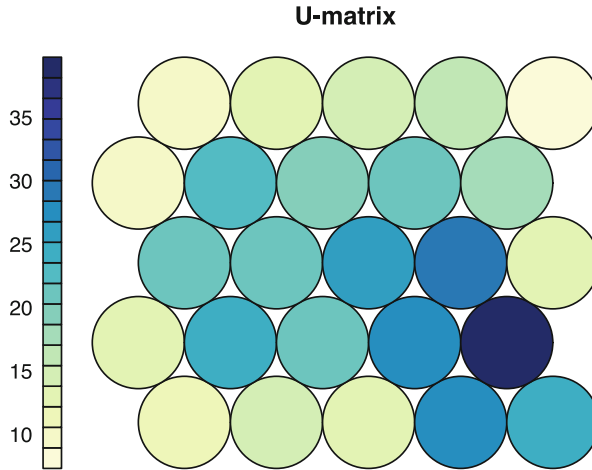


Fig. 28 SOM U-matrix for the POP data set

4. Identifying the winner neuron, or best matching unit, the most similar, i.e., the closest, to the input vector.
5. Adjustment of neuron positions. There are two methods of repositioning the nodes in the variable space:
 - (a) *Winner takes all, WTA* – where only the winner neuron is relocated in the direction of the input vector
 - (b) *Winner takes most, WTM* – where the winner neuron and the associated neurons are repositioned in the vector space
6. Until the user-defined number of cycles has been reached, repeat the cycle starting from step 2.

Visualization

Once we have determined the optimal weights of the output layer neurons, we can graphically represent the relationships between them and the original data. There are several aspects of SOMs to visualize (Brereton 2009):

1. Unified distance matrix, or U-matrix, which illustrates the similarity between neurons (map units), color-coded by distance (Fig. 28). Based on the U-matrix, it is possible to identify outliers as well as clusters, as neurons representing similar samples are positioned at adjacent regions of the map.
2. Hit histogram, where the size of the map unit is proportional to the number of points (objects) represented by that neuron (Fig. 29).
3. Component planes, which demonstrate how a specific variable influences the SOMs (Fig. 30). Each variable from the data set can be represented on a separate component plane.

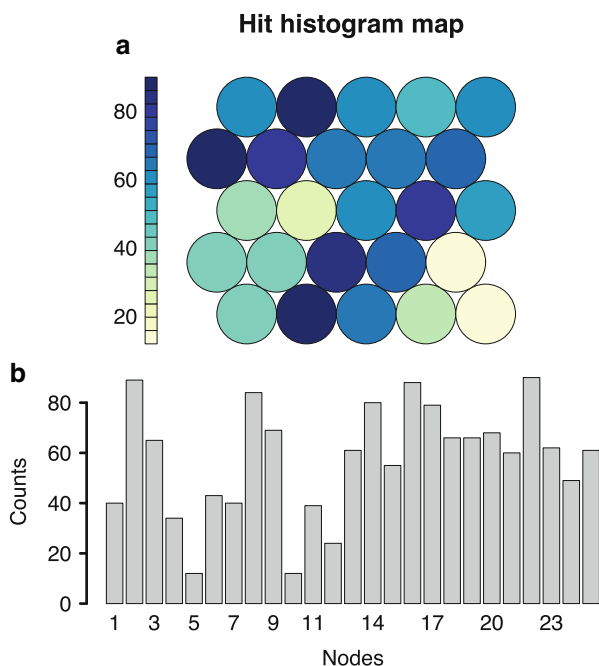


Fig. 29 Two ways of representing SOM hit histogram: (a) a 2D histogram, (b) a “classical” barplot

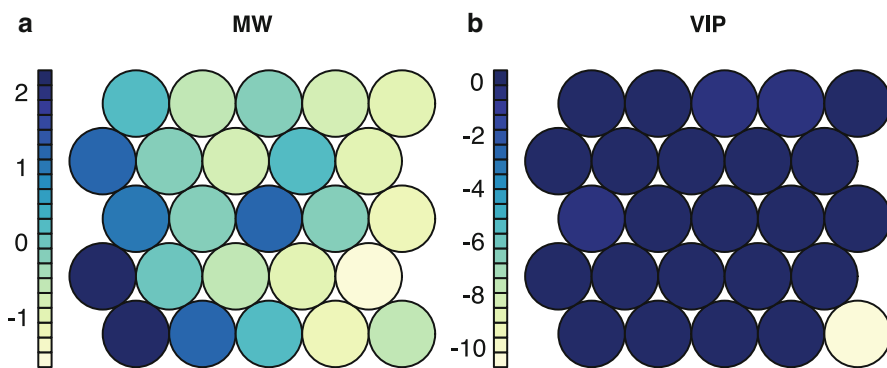


Fig. 30 SOM component planes: (a) molecular weight, MW, (b) vertical ionization potential, VIP

Quality Assessment

In order to assess the quality of the SOM training, there are a number of measures that can be employed. Among the most popular are the (Brereton 2009):

- MQE, mean quantization error. The average distance between each input vector and its best matching unit, calculated according to Eq. 13:

$$\text{MQE} = \frac{\sum_{i=1}^n d(x_i, w_c)}{n} \quad (13)$$

where $d(x_i, w_c)$ is the distance between input vector (x_i) and weight vector of the winner neuron; n is a number of input patterns.

- TE, topographic error. The ratio of all input vectors for which the first and second BMUs are not adjacent, calculated according to Eq. 14:

$$\text{TE} = \frac{\sum_{i=1}^n l(x_i)}{n} \quad (14)$$

where $l(x_i)$ is a function equal to 1 when the first and second most similar codebook vector of a particular sample is from adjacent units; otherwise, the function is 0.

Summary

In this chapter, we present selected unsupervised methods of data grouping and exploration. Each of them allows the user to focus on different aspects of the data. Hierarchical cluster analysis enables the initial exploration of the data structure and determining which objects are similar to each other in the context of various complexity levels. The k -means approach allows the user to more precisely identify group membership by an iterative search of the variable space and real-time improvement of clustering results. Thanks to modified versions of k -means, such as k -medoids and k -modes, we are able to handle outliers and categorical data with ease. With the help of robust exploratory methods, such as the principal component analysis and self-organizing Kohonen maps, it is possible to examine large data sets and determine which of the variable is crucial in describing underlying data structure.

Each of the presented techniques has some drawbacks and limitations, which is why user expertise and experience, as well as extensive knowledge on the inner workings of these approaches, are essential for effective data analysis. We hope that this chapter will serve as an overview and a starting point for further pursuit of knowledge in this field.

Bibliography

- Brereton, R. G. (2003). *Chemometrics: Data analysis for the laboratory and chemical plant*. Chichester/Hoboken: Wiley.
- Brereton, R. G. (2009). *Chemometrics for pattern recognition*. Chichester: Wiley.
- Brown, S. D., TaulerFerre, R., & Walczak, B. (2009). *Comprehensive chemometrics: Chemical and biochemical data analysis*. Amsterdam/London: Elsevier.
- Everitt, B., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). Oxford: Wiley-Blackwell.
- Gajewicz, A., Haranczyk, M., & Puzyn, T. (2010). Predicting logarithmic values of the subcooled liquid vapor pressure of halogenated persistent organic pollutants with QSPR: How different are chlorinated and brominated congeners? *Atmospheric Environment*, *44*(11), 1428–1436.
- Gemperline, P. (2006). *Practical guide to chemometrics* (2nd ed.). Boca Raton: CRC/Taylor & Francis.
- Golebiowski, M., Sosnowska, A., Puzyn, T., Bogus, M. I., Wieloch, W., Włóka, E., & Stepnowski, P. (2014). Application of two-way hierarchical cluster analysis for the identification of similarities between the individual lipid fractions of *Lucilia sericata*. *Chemistry and Biodiversity*, *11*, 733–748.
- Han, J., Kamber, M., & Pei, J. P. D. (2012). *Data mining: Concepts and techniques* (3rd ed.). Waltham/Oxford: Morgan Kaufmann/Elsevier Science, distributor.
- Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). New York: Springer.
- Jolliffe, I. T. (2002). *Principal component analysis* (Springer series in statistics 2nd ed.). New York: Springer.
- Khan, S. S., & Kant, S. (2007). Computation of initial modes for K-modes clustering algorithm using evidence accumulation. Paper presented at the Proceedings of the 20th international joint conference on artificial intelligence, Hyderabad.
- Kohonen, T. (2001). *Self-organizing maps* (3rd ed.). Berlin/London: Springer.
- Kountchev, R., & Iantovics, B. (2013). *Advances in intelligent analysis of medical data and decision support systems* (Studies in Computational Intelligence, Vol. 473). Springer International Publishing Switzerland.
- Li, Y., Pang, G.-F., Fan, C.-L., & Chen, X. (2013). Hierarchical cluster analysis of matrix effects on 110 pesticide residues in 28 tea matrixes. *Journal of AOAC International*, *96*(6), 1453–1465.
- Livingstone, D. (2009). *A practical guide to scientific data analysis*. Chichester: Wiley.
- Maimon, O. Z., & Rokach, L. (2005). *Data mining and knowledge discovery handbook*. Ramat-Aviv: Springer.
- Milligan, G., & Cooper, M. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, *50*(2), 159–179.
- Myatt, G. J. (2007). *Making sense of data: A practical guide to exploratory data analysis and data mining*. Hoboken: Wiley-Interscience.
- Petushkova, N. A., Pyatnitskiy, M. A., Rudenko, V. A., Larina, O. V., Trifonova, O. P., Kisrieva, J. S., Samenkova, N. F., Kuznetsova, G. P., Karuzina, I. I., & Lisitsa, A. V. (2014). Applying of hierarchical clustering to analysis of protein patterns in the human cancer-associated liver. *PLoS One*, *9*(8), e103950.
- Schnegg, M., Massonnet, G., & Gueissaz, L. (2015). Motorcycle helmets: What about their coating? *Forensic Science International*, *252*, 114–126.
- Skwarzec, B., Kabat, K., Puzyn, T., & Astel, A. (2011). Inflow of polonium, uranium and plutonium radionuclides in Odra River catchment area assessment by environmental expertise. *Journal of Radioanalytical and Nuclear Chemistry*, *292*(2), 519–529.

-
- Varmuza, K., & Filzmoser, P. (2009). *Introduction to multivariate statistical analysis in chemometrics*. CRC Press: Boca Raton, p xiii, 321 p.
- Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on Neural Networks/A Publication of the IEEE Neural Networks Council*, 11(3), 586–600.

Recent Developments in 3D QSAR and Molecular Docking Studies of Organic and Nanostructures

51

Bakhtiyor Rasulev

Contents

Introduction	2134
Ligand-Based 3D QSAR Methods	2135
Recent Advances in 3D QSAR Studies of Organic Compounds	2141
3D QSARs and Combined Docking Studies of Nanostructured Materials	2151
Concluding Remarks	2156
Bibliography	2156

Abstract

The development of quantitative structure–activity relationship (QSAR) methods is going very fast for the last decades. OSAR approach already plays an important role in lead structure optimization, and nowadays, with development of big data approaches and computer power, it can even handle a huge amount of data associated with combinatorial chemistry. One of the recent developments is a three-dimensional QSAR, i.e., 3D QSAR. For the last two decades, 3D-OSAR has already been successfully applied to many datasets, especially of enzyme and receptor ligands. Moreover, quite often 3D QSAR investigations are going together with protein–ligand docking studies and this combination works synergistically. In this review, we outline recent advances in development and applications of 3D QSAR and protein–ligand docking approaches, as well as combined approaches for conventional organic compounds and for nanostructured materials, such as fullerenes and carbon nanotubes.

B. Rasulev (✉)

Center for Computationally Assisted Science and Technology, North Dakota State University, Fargo, ND, USA

e-mail: rasulev@icnanotox.org; bakhtiyor.rasulev@ndsu.edu

Introduction

The methodology of quantitative structure–activity relationship (QSAR) is very well described in various publications (Hansch et al. 1995; Kubinyi 1997a, b; Eriksson et al. 2003). In short, QSAR is a method to find correlations and mathematical models for congeneric series of compounds, affinities of ligands to their binding sites, rate constants, inhibition constants, toxicological effect, and many other biological activities, based on structural features, as well as group and molecular properties, such as electronic properties, polarizability, or steric properties (Klebe et al. 1994; Hansch et al. 1995; Karelson et al. 1996; Kubinyi 1997a, b; Perkins et al. 2003; Isayev et al. 2006; Martin 2009; Rasulev et al. 2010; Puzyn et al. 2011).

Thus, QSAR approaches have been used for many types of biological activities to describe correlations for series of drugs and drug candidates (Kubinyi 1997a, b; Veber et al. 2002). In addition, in case of available crystallographic data on the proteins, the QSAR models can be developed with the additional information from the three-dimensional (3D) structures of these proteins, interacting with drug candidates, by applying protein–ligand docking data for further QSAR analysis, or, if there is no data on 3D structure of protein, then developing QSAR based on three-dimensional features of investigated molecules (Moro et al. 2005; Ragno et al. 2005; Gupta et al. 2009; Hu et al. 2009; Sun et al. 2010; Araújo et al. 2011; Ahmed et al. 2013). The second approach was named as 3D3D QSAR approach (Wise et al. 1983; Cramer and Bunce 1987; Cramer et al. 1988; Clark et al. 1990). There are also many other multidimensional approaches, including 4D QSAR and 5D QSAR, but all of them are just extension of QSAR analysis to a number of conformations (orientations, tautomers, stereoisomers, or protonation states) per molecule, number of concentrations (dosages) per compound, etc (Lill 2007). In overall, when talking about 3D QSAR, computational chemists usually assume that the QSAR analysis takes into account a three-dimensional structure of the compound in minimal energy conformation and builds QSAR model based on various 3D fields generated (Kubinyi 1997a, b).

A first similar to 3D QSAR approach was developed by Cramer in 1983, which was the predecessor of 3D approaches called dynamic lattice-oriented molecular modeling system (DYLOMMS) that involves the use of PCA to extract vectors from the molecular interaction fields, which are then correlated with biological activities (Wise et al. 1983). Later authors improved this approach and by combining the two existing techniques, GRID and PLS, has developed a powerful 3D QSAR methodology, so-called comparative molecular field analysis (CoMFA) (Cramer et al. 1988; Clark et al. 1990). Soon after, CoMFA has become a prototype of 3D QSAR methods (Kim et al. 1998; Todeschini and Gramatica 1998; Podlogar and Ferguson 2000). CoMFA approach was then implemented in the Sybyl software (Tripos 2006) from Tripos Inc.

As it was mentioned before, a good and fruitful approach is a combination of molecular docking and 3D QSAR pharmacophore methods (Patel et al. 2008; Gupta et al. 2009; Araújo et al. 2011; Ahmed et al. 2013). Molecular docking and 3D QSAR model are the two potent methods in drug discovery process. Thus,

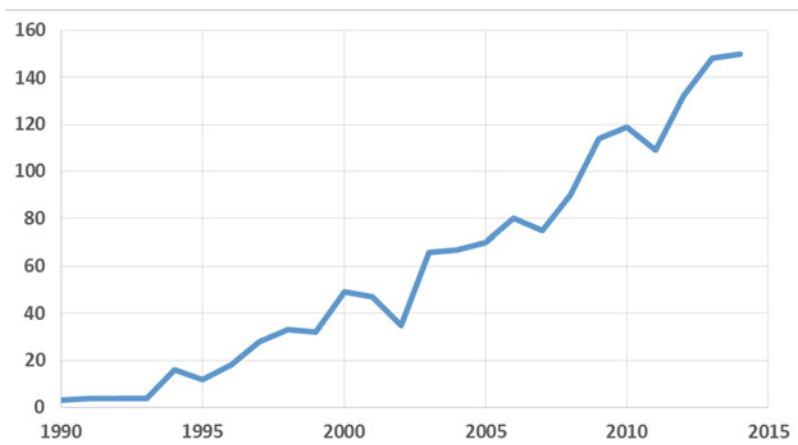


Fig. 1 A number of 3D QSAR-related papers for a period of 1990–2014 (Source – Web of Science)

virtual screening using 3D QSAR approaches followed by docking has become one of the reputable methods for drug discovery and enhancing the efficiency in lead optimization (Oprea and Matter 2004). The main advantage of this combined approach of 3D QSAR and pharmacophore-based docking is to focus on specific key interaction for protein–ligand binding to improve drug candidates. Ameliorate the selection of active compounds; it is optimal to use both methods like molecular docking and 3D QSAR modeling (Gopalakrishnan et al. 2005; Klebe 2006; Perola 2006; Pajeva et al. 2009; Yang 2010).

Since the time of development of 3D QSAR approach, a number of papers and methods' developments were published within 3D QSAR methodology. Let's briefly list and explain these methods here and then discuss recent developments and applications of these 3D QSARs in the assessment of the properties of biologically active compounds and development of drugs and drug candidates. As can be seen from Fig. 1, the number of publications related to 3D QSAR approach is increasing every year, from 3 to 5 publications in the beginning of 1990s to about 150 publications per year in 2014. It confirms the increasing importance of the method and successful application in many drug design projects.

Ligand-Based 3D QSAR Methods

To give some view on a number of 3D QSAR methods developed for the last three decades, we listed below the ligand-based 3D QSAR methods and very short description per each of them.

CoMFA – Comparative molecular field analysis is the method which correlates the field values of the structure with biological activities. CoMFA generates an

equation correlating the biological activity with the contribution of interaction energy fields at every grid point (Cramer et al. 1988). The method was developed in the 1988 and still one of the most popular ones for 3D QSAR modeling.

CoMSIA – Comparative molecular similarity indices analysis (CoMSIA) method, where the molecular similarity indices calculated from steric and electrostatic alignment (SEAL) similarity fields and applied as descriptors to encode steric, electrostatic, hydrophobic, and hydrogen bonding properties (Klebe et al. 1994). This is a development of CoMFA method and also gets very popular in drug design.

GRID – This method and program was designed as an alternative to the original CoMFA approach. It is actually a force field which calculates the interaction energy fields in molecular-field analysis and determines the energetically favorable binding sites on molecules of known structure. The method to some extent is similar to CoMFA, and it computes explicit nonbonded (or non-covalent) interactions between a molecule of known 3D structure and a probe (i.e., a chemical group with certain user-defined properties). The probe is located at the sample positions on a lattice throughout and around the macromolecule. The method offers two distinct advantages, one of them is the use of a 6–4 potential function for calculating the interaction energies, which is smoother than the 6–12 form of the Lennard-Jones type in CoMFA, and another advantage is the availability of different types of probes (Goodford 1985). Moreover, the program in addition of computing the regular steric and electrostatic potentials also calculates the hydrogen bonding potential using a hydrogen bond donor and acceptor, as well as the hydrophobic potential using a “DRY probe.” Later on, a water probe was included to calculate hydrophobic interactions (Kim et al. 1998; Kim 2001).

MSA – Molecular shape analysis (MSA) is a ligand-based 3D QSAR method which attempts to merge conformational analysis with the classical Hansch approach. The method deals with the quantitative characterization, representation, and manipulation of molecular shape in the construction of a QSAR model (Hopfinger 1980).

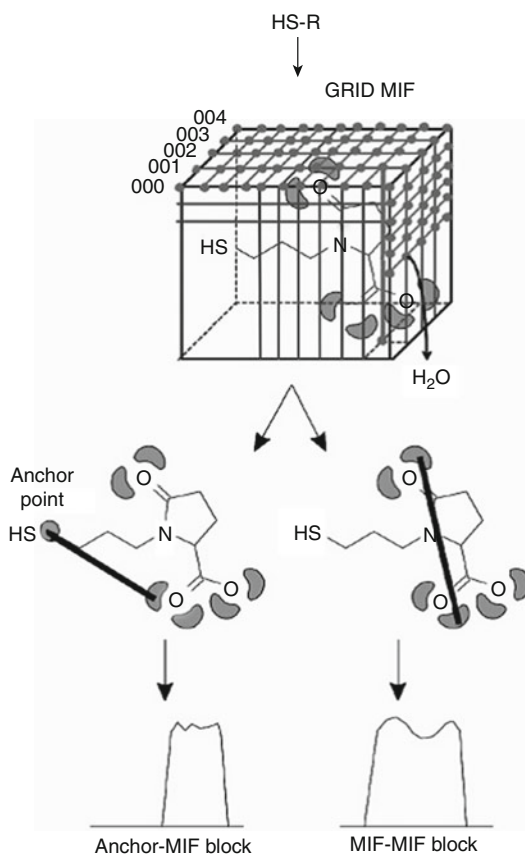
HASL – Inverse grid-based approach represents the shapes of the molecules inside an active site as a collection of grid points (Doweyko 1988). The methodology of this approach begins with the intermediate conversion of the Cartesian coordinates (x, y, z) for superposed set of molecules to a 3D grid consisting of the regularly spaced points that are (1) arranged orthogonally to each other, (2) separated by a particular distance termed as the resolution (which determines the number of grid points representing a molecule), and (3) all sprawl within the van der Waals radii of the atoms in the molecule.

Thus, the resulting set of points is referred to as the molecular lattice and represents the receptor active site map (like in CoMFA). Quite important that the overall lattice dimensions are dependent on the size of the molecules and the resolution chosen.

GRIND – This is the method that uses *grid-independent descriptors* (GRIND) which encode the spatial distribution of the molecular interaction fields (MIF) of the studied compounds (Pastor et al. 2000). In the anchor-GRIND method (Fontaine

et al. 2005), to compare the MIF distribution of different compounds, the user defines a single common position in the structure of all the compounds in the series, so-called anchor point. This anchor point does not provide enough geometrical constraints to align the compounds studied; however, it is used by the method as a common reference point, making it possible to describe the geometry of the MIF regions in a more precise way than GRIND does. The anchor point is particularly easy to assign in datasets having some chemical substituents well known as being crucial for the activity. In the anchor-GRIND approach, the R groups are described with two blocks of variables: the anchor-MIF and the MIFMIF blocks (Fig. 2). The first one describes the geometrical distribution of the R MIF relative to the anchor point, while the second one describes the geometrical distribution of the MIF within each R group. These blocks are obtained after the following steps: (i) every R group is considered as attached to the scaffold, (ii) the anchor point is set automatically on an atom of the scaffold, (iii) a set of MIF are calculated with the program GRID (Goodford 1985) as well as the shape field implemented in the program Almond (Cruciani et al. 2004), and (iv), as the last step, the blocks of descriptors

Fig. 2 Calculation of the anchor-GRIND descriptors for an ACE inhibitor with the anchor point set on the zinc binder sulfur atom (Reproduced with permission from reference (Fontaine et al. 2005) Copyright, American Chemical Society, 2005)



are computed from the anchor point and the filtered MIF. Authors also were able to incorporate a molecular shape into the GRIND descriptors (Fontaine et al. 2004).

GERM – Genetically evolved receptor model (GERM) is a method for 3D QSAR and also for constructing 3D models of protein-binding sites in the absence of a crystallographically established or homology-modeled structure of the receptor (Walters and Hinds 1994). As for many 3D QSAR datasets, the primary requirement for GERM is a structure–activity set for which a sensible alignment of realistic conformers has been determined. The methodology is the following: it encloses the superimposed set of molecules in a shell of atoms (analogous to the first layer of atoms in the active site) and allocates these atoms with explicit atom types (aliphatic H, polar H, etc., to match the types of atoms found in the investigated proteins).

CoMMA – Comparative molecular moment analysis (CoMMA) is one of the unique alignment-independent 3D QSAR methods, which involves the computation of molecular similarity descriptors (similar to CoMSIA) based on the spatial moments of molecular mass (i.e. shape) and charge distributions up to second-order as well as related quantities (Silverman and Platt 1996).

COMBINE – Comparative binding energy analysis (COMBINE) method was developed to make use of the structural data from ligand–protein complexes, within a 3D QSAR methodology. The method is based on the hypothesis where free energy of binding can be correlated with a subset of energy components calculated from the structures of receptors and ligands in bound and unbound forms (Ortiz et al. 1995; Lushington et al. 2007).

CoMSA – Comparative molecular surface analysis (CoMSA) is a non-grid 3D QSAR method that utilizes the molecular surface to define the regions of the compounds which are required to be compared using the mean electrostatic potentials (MEPs) (Polanski et al. 2002, 2006). In overall, the methodology proceeds by subjecting the molecules in the dataset to geometry optimization and assigning them with partial atomic charges.

AFMoC – Adaptation of fields for molecular comparison (AFMoC) is a close to 3D QSAR method that involves fields derived from the protein environments (i.e. not from the superimposed ligands as in CoMFA); therefore, it is also known as a “reverse” CoMFA (=AFMoC) approach or protein-dependent 3D QSAR (Gohlke and Klebe 2002). The methodology is the following: a regularly spaced grid is placed into the receptor binding site, followed by mapping of the knowledge-based pair potentials between protein atoms and ligand atom probes onto the grid intersections resulting in the potential fields. Thus, based on these potential fields, specific interaction fields are generated by multiplying distance-dependent atom-type properties of actual ligands docked into the active site with the neighboring grid values. In result, these atom-type-specific interaction fields are then correlated with the binding affinities using PLS technique, which assigns individual weighting factors to each field value.

CoRIA – Comparative residue interaction analysis (CoRIA) is a 3D QSAR approach which utilizes the descriptors that describe the thermodynamic events involved in ligand binding, to explore both the qualitative and the quantitative features of the ligand–receptor recognition process. The CoRIA methodology is

the following: initially it simply consisted of calculating the nonbonded (van der Waals and coulombic) interaction energies between the ligand and the individual active site residues of the receptor that are involved in interaction with the ligand (Datar et al. 2006; Dhaked et al. 2009). By employing the genetic algorithm-supported PLS technique (G-PLS), these energies then correlated with the biological activities of molecules, along with the other physiochemical variables describing the thermodynamics of binding, such as surface area, lipophilicity, molar refractivity, molecular volume, strain energy, etc.

SOMFA – Self-organizing molecular-field analysis, where firstly the mean activity of training set is subtracted from the activity of each molecule to get their mean-centered activity values. The methodology is the following:

- A 3D grid around the molecules with values at the grid points signifying the shape or electrostatic potential is generated.
- The shape or electrostatic potential value at every grid point for each molecule is multiplied by its mean-centered activity.
- The grid values for each molecule are summed up to give the master grids for each property.
- Then the so-called $SOMFA_{property,i}$ descriptors from the master grid values are calculated and correlated with the log-transformed molecular activities (Robinson et al. 1999).

kNN-MFA – This relatively new method was developed and reported in 2006 by Ajmani et al. (2006). kNN-MFA is a k-nearest neighbor molecular-field analysis. kNN-MFA adopts a k-nearest neighbor principle for generating relationships of molecular fields with the experimentally reported activity. This method utilizes the active analogue principle that lies at the foundation of medicinal chemistry. Like many 3D QSAR methods, kNN-MFA requires suitable alignment of a given set of molecules. This is followed by generation of a common rectangular grid around the molecules. The steric and electrostatic interaction energies are computed at the lattice points of the grid using a methyl probe of charge +1. These interaction energy values are considered for relationship generation and utilized as descriptors to decide nearness between molecules.

3D-HoVAIFA – This method based on three-dimensional holographic vector of atomic interaction field analysis (Zhou et al. 2007). Initially the holographic vector for 3D QSAR methods was developed by Zhou et al. in 2007 (Zhou et al. 2007). Proceeding from two spatial invariants, namely, atom relative distance and atomic properties on the bases of three common nonbonded (electrostatic, van der Waals, and hydrophobic) interactions which are directly associated with bioactivities, 3D-HoVAIF method derives multidimensional vectors to represent molecular steric structural characteristics.

CMF – This is a recently introduced continuous molecular-field approach (Baskin and Zhokhova 2013). This is a novel approach that consists in encapsulating continuous molecular fields into specially constructed kernels. It is based on the application of continuous functions for the description of molecular fields instead

of finite sets of molecular descriptors (such as interaction energies computed at grid nodes) commonly used for this purpose. The feasibility of using molecular-field kernels in combination with the support vector regression (SVR) machine learning method to build 3D QSAR models has been demonstrated by the same authors earlier (Zhokhova et al. 2009). Authors claim that by combining different types of molecular fields and methods of their approximation, different types of kernels with different types of kernel-based machine learning methods, it is possible not only to present lots of existing methods in chemoinformatics and medicinal chemistry as particular cases within a single methodology but also to develop new approaches aimed at solving new problems (Baskin and Zhokhova 2013). The example of application of this approach is described later in this chapter.

PHASE – This is a flexible system (engine) (Dixon et al. 2006) for common pharmacophore identification and assessment, 3D QSAR model development, and 3D database creation and searching (within Schrodinger Suite, Schrodinger, LLC). It includes some subprograms, for example, LigPrep, which attaches hydrogens, converts 2D structures to 3D, generates stereoisomers, and, optionally, neutralizes charged structures or determines the most probable ionization state at a user-defined pH. It also includes MacroModel conformational search engine to generate a series of 3D structures that sample the thermally accessible conformational states. For purposes of 3D modeling and pharmacophore model development, each ligand structure is represented by a set of points in 3D space, which coincide with various chemical features that may facilitate non-covalent binding between the ligand and its target receptor. PHASE provides six built-in types of pharmacophore features: hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic (H), negative ionizable (N), positive ionizable (P), and aromatic ring (R). In addition, users may define up to three custom feature types (x, y, z) to account for characteristics that don't fit clearly into any of the six built-in categories. To construct a 3D QSAR model, a rectangular grid is defined to encompass the space occupied by the aligned training set molecules. This grid divides space into uniformly sized cubes, typically 1 Å on each side, which are occupied by the atoms or pharmacophore sites that define each molecule.

APF – In 2008, Totrov M (Totrov 2008) introduced atomic property fields (APF) for 3D QSAR analysis. APF concept is introduced as a continuous, multicomponent 3D potential that reflects preferences for various atomic properties at each point in space (Fig. 3). The approach is extended to multiple flexible ligand alignments using an iterative procedure, Self-Consistent atomic Property Fields by Optimization (SCAPFOld). The application of atomic property fields and SCAPFOld for virtual ligand screening and 3D QSAR is tested on published benchmarks. The new method is shown to perform competitively in comparison to current state-of-the-art methods (CoMFA and CoMSIA).

Thus, there are studies with comparative analysis of these two methods, PHASE and Catalyst (HypoGen). In 2007, Evans et al. (2007) provided a comparative study of PHASE and Catalyst methods for their performance in determining 3D QSARs and concluded that the performance of PHASE is better than or equal to that of Catalyst HypoGen, with the datasets and parameters used. Authors found that within

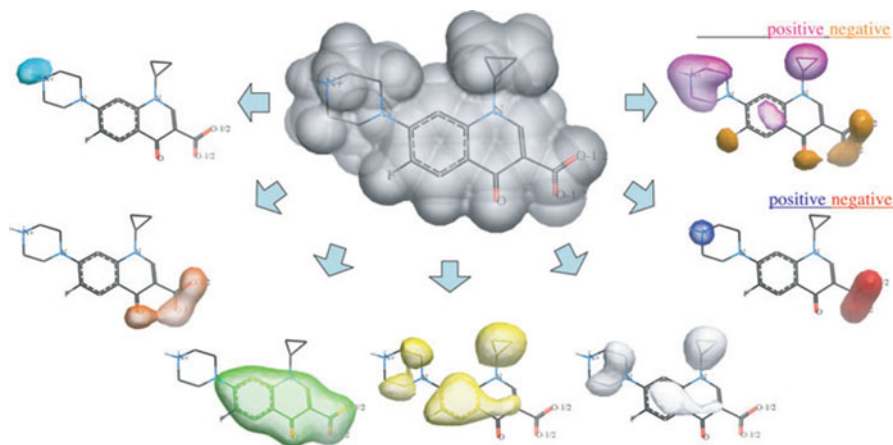


Fig. 3 An overall scheme of APF potentials visualized for an example of a drug molecule. Transparent blobs are the equipotential contours of various APF components (Reproduced with Permission from reference (Totrov 2008). Copyright John Wiley and Sons, 2008)

PHASE, the atom-based grid QSAR model generally performed better than the pharmacophore-based grid, and by using overlays from Catalyst to build PHASE grid QSAR models, they found evidence that better performance of PHASE on these datasets was due to the use of the grid technique.

Recent Advances in 3D QSAR Studies of Organic Compounds

In this part of the review, we discuss the new developments in the methods and in applications of 3D QSARs for various chemicals, including nanostructured materials.

For the last 10 years, there was mainly improvement of the 3D QSAR approaches which were developed before 2005. As it was discussed above, all these methods as CoMFA, CoMSIA, GRID, SOMFA, etc., were developed in late 1990s and early 2000s. Some of the recently introduced methods are just improvements of old approaches. However, if we take a look at applications, we can see many interesting publications and novel ligand developments which were designed by 3D QSAR and docking methods.

Recently, a quite interesting study was performed by Virsodia et al. (2008) on antitubercular activity of 23 substituted *N*-phenyl-6-methyl-2-oxo-4-phenyl-1,2,3,4-tetrahydropyrimidine-5-carboxamides by application of 3D QSAR using CoMFA and CoMSIA methods. Authors synthesized and assessed the antitubercular activity of all investigated compounds followed by comprehensive 3D QSAR modeling. Authors were able to get good models with $r^2 = 0.98$ and 0.95 , with cross-validated $q^2 = 0.68$ and 0.58 , respectively. Authors stated that CoMFA and

CoMSIA contours helped them to design some new molecules with improved activity (Virasodia et al. 2008).

Another CoMFA and CoMSIA study was performed by Ravichandran et al. (2009) by analysis of anti-HIV activity of 96 1,3,4-thiazolidine derivatives. Authors were able to get good models by CoMFA and CoMSIA with r^2 values 0.931 and 0.972, respectively. The predictive model was evaluated using a test set comprising of 17 molecules, and the predicted r^2 values of CoMFA and CoMSIA models were 0.861 and 0.958, respectively.

With the use of CoMSIA method, Kumar et al. (2009) were able to investigate novel Bignelli dihydropyrimidines with potential anticancer activity. The developed model based on 32 compounds showed a good statistical data – for training set $q^2 = 0.51$, while for the test set $r^2 = 0.93$.

Raparti et al. in 2009 (Raparti et al. 2009) reported a study based on a novel kNN-MFA 3D QSAR which was discussed above, where authors synthesized, assessed for antimycobacterial activity, and investigated by 2D and 3D QSAR approaches a series of ten compounds (4-(morpholin-4-yl)-N0-(arylidene)benzohydrazides). Authors were able to get satisfactory for this size of dataset statistical results for 3D QSAR model against *M. tuberculosis* (Raparti et al. 2009), with $r^2 = 0.910$ and $q^2 = 0.507$, respectively.

Another kNN-MFA 3D QSAR study was conducted and published by Kishore Jha et al. in 2010 (Jha et al. 2010). Authors evaluated the antimicrobial activity of 21 compounds by kNN-MFA combined with various selection procedures. As selection methods, authors were using simulated annealing (SA), genetic algorithms (GA), and stepwise (SW) forward–backward methods. The developed model showed satisfactory results for this kind of studies, with $q^2 = 0.696$ and $r^2_{pred} = 0.615$. Authors concluded that the 3D QSAR study has shown that less electronegative substituent would be favorable for the activity, and therefore the future molecules should be designed with less electronegative group to result in potentially active molecules.

Thus, recently, Araújo et al. (2011) studied acetylcholine inhibitors (AChEIs) by application of combined approach, so-called receptor-dependent 3D QSAR (RD 3D QSAR) where they investigated a series of 60 benzylpiperidine inhibitors of human acetylcholinesterase. They received two models with $r^2 = 0.86$, $q^2 = 0.74$ and $r^2 = 0.90$, $q^2 = 0.75$, which were validated by a combined GA-PLS approach. Based on those models, authors have proposed four new benzylpiperidine derivatives and predicted the pIC_{50} for each molecule. The good predicted potency of one of the benzylpiperidine derivatives indicated a promising potency for this candidate as a new HuAChE inhibitor (Araújo et al. 2011).

In another similar study, in 2009 Gupta et al. (2009) conducted an interesting combined study with protein–ligand docking-based 3D QSAR study of HIV-1 integrase inhibitors. They were using protein–ligand docking to identify a potential binding mode for 43 inhibitors at HIV-1 IN active site, and best docked conformation of certain molecule was used as a template for alignment. The docking was followed by CoMFA and CoMSIA modeling, and authors developed very good models with r^2_{cv} values of 0.728 and 0.794, respectively, and non-cross-

validated ones $r_{ncv}^2 = 0.934$ and 0.928 . This combined docking-based 3D QSAR methodology showed really good predictive abilities and can be employed further in the development of better inhibitors for various proteins.

One more study is worth to discuss where authors applied a combination of docking and 3D QSAR to reveal the most important structural factors for the activity. Here, Hu et al. (2009) applied a receptor- and ligand-based 3D QSAR study for a series of 68 non-nucleoside HIV-1 reverse transcriptase inhibitors (2-amino-6-arylsulfonylbenzotrioles and their thio and sulfinyl congeners). Authors were applying docking simulations to position the inhibitors into RT active site to determine the most probable binding mode and most reliable conformations. This complex receptor-based and ligand-based alignment procedure and different alignment modes allowed authors to obtain reliable and predictive CoMFA and CoMSIA models with cross-validated q^2 value of 0.723 and 0.760 , respectively. Authors concluded that the CoMFA steric and CoMSIA hydrophobic fields support the idea that bulkier and hydrophobic groups are favorable to bioactivity in the 3- and 5-positions of the B (benzene)-ring. At the same time, these groups are unfavorable in the 4-position. Also, the CoMSIA H-bond donor and acceptor fields suggest that the sulfide and sulfone inhibitors are more active than the sulfoxide ones due to H-bonding with protein residues.

It is good to mention here another interesting study where combination of methods is used, including molecular docking and 3D QSAR to develop a predictive QSAR model. Moro et al. (2006) suggested the use a combination of molecular electrostatic potential (MEP) surface properties (autocorrelation vectors) with the conventional partial least-square (PLS) analysis to produce a robust ligand-based 3D structure–activity relationship (autoMEP/PLS). They applied this approach to predict human A3 receptor antagonist activities. First of all, the approach was suggested as an efficient and alternative pharmacodynamic-driven filtering method for small-size virtual libraries. For this, authors generated a small-sized combinatorial library (841 compounds) that was derived from the scaffold of the known human A3 antagonist pyrazolo-triazolo-pyrimidines (Moro et al. 2005). This is another successful example of combined approach of docking and 3D QSAR to investigate and design active analogue compounds. Authors were using MULTIDOCK code that is part of MOE suite (Molecular Operating Environment (MOE) 2016) to get a conformational sampling and then calculate interaction energies using MMFF94 (Halgren 1996) and use it for further steps. The MEPs were derived from a classical point charge model: the electrostatic potential for each molecule is obtained by moving a unit positive point charge across the van der Waals surface, and it is calculated at various points j on this surface (Moro et al. 2006). Authors were able to test the approach by synthesizing several predicted potent compounds, and they found that all the newly synthesized compounds are correctly predicted as potent human A3 antagonists (Moro et al. 2006).

As a continuation of development of pharmacophore- and docking-based methods for QSAR, the novel PHASE code was developed. This updated code then was used by Amnerkar and Bhusari (2010) to investigate by 3D QSAR approach the anticonvulsant activity of some prop-2-eneamido and 1-acetyl-pyrazolin derivatives

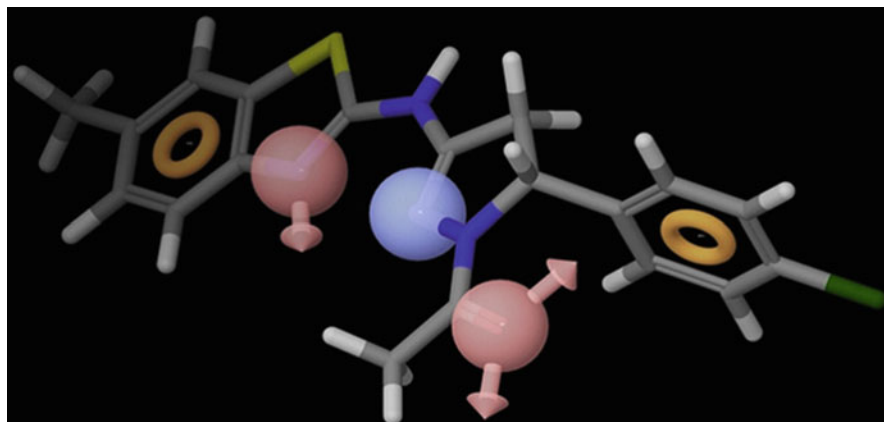


Fig. 4 Common pharmacophore generated from the best PHASE hypothesis. Pharmacophore features are *red sphere* for hydrogen bond acceptors (A) with the *arrows* pointing in the direction of lone pair, *blue sphere* for positively charged group (P), and *orange torus* for aromatic rings (R). Compound 52 aligned to the pharmacophore for which *blue* indicates nitrogen, *yellow* refers to sulfur, *green* indicates chlorine, *gray* indicates carbon, and *white* refers to hydrogen (Reproduced with Permission from reference (Amnerkar and Bhusari 2010). Copyright Elsevier, 2010)

of aminobenzothiazole. They received a statistically significant 3D QSAR model with r^2 of 0.922 and q^2 of 0.814. The model was analyzed in order to understand the trends of investigated molecules for their anticonvulsant properties. Authors found the influence of electron withdrawing, hydrogen bond donor, and negative/positive ionic and hydrophobic groups to anticonvulsant activity. Authors believe that the derived 3D QSAR as well as clues for possible structural modifications will be of interest and significance for the strategic design of more potent molecules in the benzothiazoles as anticonvulsant agents. The pharmacophore hypothesis generated from PHASE-based 3D QSAR analysis can be seen in Fig. 4.

Another PHASE application for 3D QSAR study is published by Pulla et al. (2016). Authors applied a 3D QSAR approach to investigate silent mating-type information regulation 2 homologue 1 (SIRT1) which is the homologous enzyme of silent information regulator-2 gene in yeast. SIRT1 was believed to be overexpressed in many cancers (prostate, colon) and inflammatory disorders (rheumatoid arthritis); that is why it has good therapeutic importance. Authors conducted both structure-based and ligand-based drug design strategies with utilizing high-throughput virtual screening of chemical databases. Then an energy-based pharmacophore was generated using the crystal structure of SIRT1 bound with a small molecule inhibitor and compared with a ligand-based pharmacophore model that showed four similar features. A 3D QSAR model was developed and applied to generated structures. Among the designed compounds, Lead 17 emerged as a promising SIRT1 inhibitor with IC_{50} of 4.34 μ M and, at nanomolar concentration (360 nM), attenuated the proliferation of prostate cancer cells (LnCAP) (Pulla et al. 2016). The 3D QSAR model was developed using PHASE 3.4 module in Maestro 9.3 software

package developed by Schrodinger, LLC (Dixon et al. 2006). Docking studies were executed using Glide 5.8 module (Halgren et al. 2004). Authors were validating the pharmacophore model by set composed of 1055 compounds, consisting of 1000 decoys and 55 known inhibitors. The drug-like decoy set of 1000 compounds was obtained from the Glide module (Halgren et al. 2004). A final 3D QSAR model was developed based on dataset of 79 molecules reported as SIRT1 inhibitors in various literatures. To develop QSAR model, PHASE module relied on PLS regression applied to a large set of binary-valued variables. Each independent variable in the model originated from the grid of cubic volume elements spanning the space occupied by the training set ligands. Each training ligand in the training set was represented by binary code consisting of set of bit values (0 or 1) indicating the volume of elements occupied by van der Waals model of the ligand. Authors were able to get a very good 3D QSAR model with $r^2 = 0.953$, $q^2 = 0.908$, and $r^2_{ext} = 0.941$. A validated 3D QSAR model (for ADHRR 802) authors used to generate contour maps could help in understanding the importance of functional groups at specific positions toward biological activity. These insights could be known by comparing the contour maps of the most and least active compounds, as shown in Fig. 5 represented in Pulla et al. (2016). The blue and red cubes indicated the favorable and unfavorable regions, respectively, of the hydrogen bond donor effect, while light-red and yellow cubes indicated favorable and unfavorable regions, respectively, of the hydrophobic effect, and the cyan and orange cubes indicated favorable and unfavorable regions, respectively, of the electron-withdrawing effect. From Fig. 5a, it can be seen that the blue favorable regions of the hydrogen bond donor effect were nearer to the donor feature (D5) of the active molecule; however, it could also be observed that blue boxes were also concentrated at the amide group beside thiophen, thus illustrating that additional donor groups at these regions (blue cubes) could increase biological activity. At the same time, in the inactive molecule, red unfavorable boxes were observed around the donor feature (D5), inferring the biological inactiveness of the molecule. In the case of the hydrophobic effect, the light-red color cubes were seen surrounding the hydrophobic feature (H8, piperidine) of the active molecule, whereas the presence of few yellow unfavorable cubes indicated that these hydrophobic groups were not in the right position in the inactive molecule, illustrating the weak biological activity. Next, in the case of the electron-withdrawing effect of the active molecule, the favorable cyan cubes were seen around the acceptor feature (A3), and cyan cubes were also seen near the pyrimidine ring. It inferred that the acceptor features near the pyrimidine ring could further increase the bioactivity of the molecule. However, in the case of the inactive molecule, mostly unfavorable orange cubes were observed around the acceptor feature (A3), illustrating the importance of the electron-withdrawing group in the activity of lead molecules.

Thus, another new combined docking-based 3D QSAR study (Sun et al. 2010) was published with the analysis of influenza neuraminidase inhibitors. The study was based on novel 3D-HoVAIFA method which is based on three-dimensional holographic vector of atomic interaction field analysis (Zhou et al. 2007). As it was mentioned above, initially the holographic vector for 3D QSAR was developed by

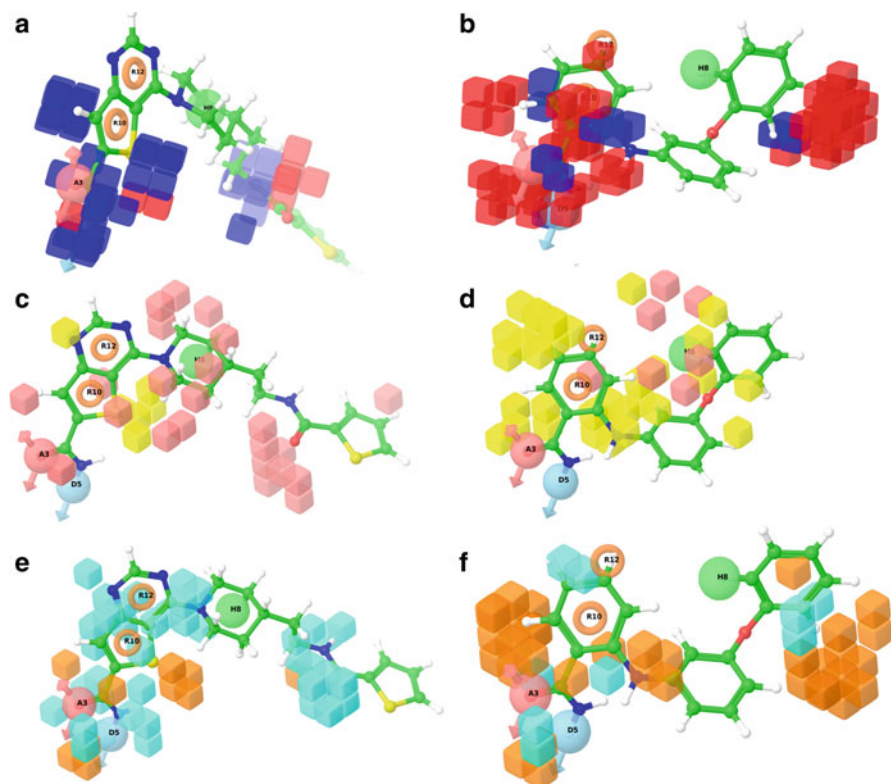


Fig. 5 Contour maps of the most active and inactive molecules: (a) the hydrogen bond donor effect (most active), (b) the least active (*blue* denotes favorable, *red* denotes unfavorable), (c) the hydrophobic effect (most active), (d) the least active (*light red* denotes favorable, *yellow* denotes unfavorable), (e) electron-withdrawing effect (most active), and (f) the least active (*cyan* denotes favorable, *orange* denotes unfavorable) (Reproduced with Permission from ref (Pulla et al. 2016). Copyright, American Chemical Society, 2016)

Zhou et al. in 2007 (Zhou et al. 2007). The method uses atomic relative distance and atomic properties on the bases of three common nonbonded (electrostatic, van der Waals, and hydrophobic) interactions which are directly associated with bioactivities, and then 3D-HoVAIF method derives multidimensional vectors to represent molecular steric structural characteristics for further 3D QSAR analysis. Similarly to previous study, authors conducted a docking study to find the best docking pose and template for alignment. Then authors were able to get good models for a large dataset of 124 compounds and received the following correlation coefficients, $r^2 = 0.789$ and $r^2_{cv} = 0.732$. Authors claim that 3D-HoVAIFA can be applicable to molecular structural characterization and bioactivity prediction. In addition, it was showed that HoVAIFA and docking results are corresponding (Sun et al. 2010), which illustrates that HoVAIFA is an effective methodology for characterization of complex interactions of drug molecules.

One more docking-based 3D QSAR study was published in 2010 by Sakkiah et al. (2010), where authors conducted 3D QSAR pharmacophore-based virtual screening and molecular docking for the identification of potential HSP90 inhibitors. Authors were using HYPO and HYPOGEN (Li et al. 2000) 3D-based pharmacophore models. Based on the training set of 16 compounds, they were able to develop a good model using pharmacophore generation module in Discovery Studio (Accelrys) and then apply it for test set of 30 compounds. For predicting activity, the correlation coefficients of the model for training and test sets were 0.93 and 0.91, respectively. Authors then applied the model to virtual screening of about 160,000 compounds (Maybridge and Scaffold databases) and finally selected 1150 compounds for docking studies. Finally, 36 selected compounds were reported that were showing high activity based on 3D QSAR model and docking analysis. The developed HYPOGEN pharmacophore model that was used for virtual screening of 160,000 compounds from the databases is represented in Fig. 6.

Recently introduced and discussed previously the CMF approach for 3D QSAR analysis was successfully applied for several datasets (Baskin and Zhokhova 2013). Authors applied CMF approach to build 3D QSAR models for eight datasets through the use of five types of molecular fields (the electrostatic, steric, hydrophobic, hydrogen bond acceptor and donor ones). The 3D QSAR models were developed for the following datasets: 114 angiotensin converting enzyme (ACE) inhibitors, 111

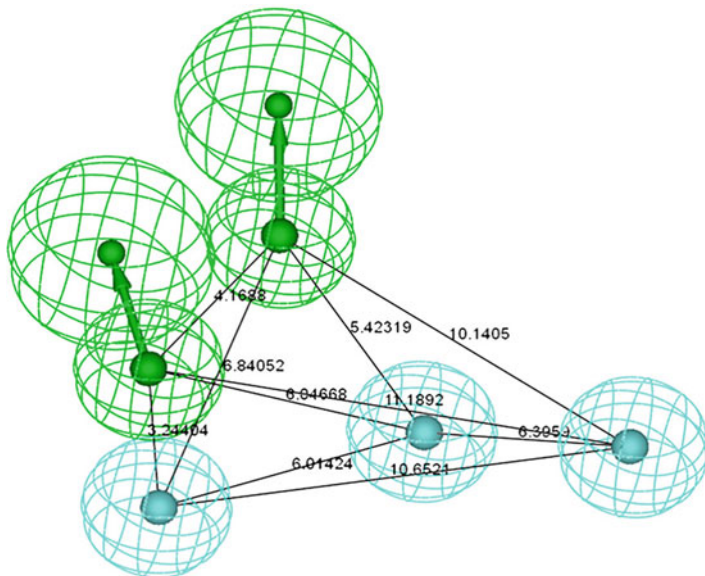


Fig. 6 Catalyst HypoGen pharmacophore model, where H and HBA are illustrated in cyan and green, respectively (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article) (Reproduced with Permission from reference Sakkiah et al. 2010) Copyright Elsevier, 2010)

acetylcholinesterase (AChE) inhibitors, 163 ligands for benzodiazepine receptors (BZR), 322 cyclooxygenase-2 (COX-2) inhibitors, 397 dihydrofolatereductase (DHFR) inhibitors, 66 glycogen phosphorylase b (GPB) inhibitors, 76 thermolysin (THER) inhibitors, and 88 thrombine (THR) inhibitors. Authors were able to get good models and then compare statistical characteristics of the developed models with the same characteristics built earlier for the same datasets using the most popular 3D QSAR methods, CoMFA and CoMSIA, based on the use of molecular fields. Almost for all cases authors received better statistics than in original works. For example, for ACE inhibitors, CMF approach showed $q^2 = 0.72$, while CoMFA and CoMSIA showed 0.68 and 0.65, respectively. 3D QSAR for AChE inhibitors showed for CMF $q^2 = 0.58$, while for CoMFA and CoMSIA, it was 0.52 and 0.48, respectively. 3D QSAR for DHFR inhibitors showed for CMF $q^2 = 0.67$, while for CoMFA and CoMSIA, it was only 0.49 and 0.53, respectively. Other datasets also showed better results. The only one dataset, for BZR receptors, showed not so high values ($q^2 = 0.40$) but comparable with previous data received by CoMFA and CoMSIA – 0.32 and 0.41, respectively. As follows from the results presented in this paper, this particular implementation of the CMF approach provides an appealing alternative to the traditional lattice-based methodology. This method provides either comparable or enhanced predictive performance in comparison with state-of-the-art 3D QSAR methods, such as CoMFA and CoMSIA. Authors also discussed advantages and disadvantages of this approach. The potential advantages of this approach result from the ability to approximate electronic molecular structures with any desirable accuracy level, the ability to leverage the valuable information contained in partial derivatives of molecular fields (otherwise lost upon discretization) to analyze models and enhance their predictive performance, the ability to apply integral transforms to molecular fields and models, etc. The most attractive features of the CMF approach are its versatility and universality. At the same time, of the most serious limitations of the CMF approach, at least in its present form, comes from the mere nature of kernel-based machine learning methods. The amount of computational resources needed to calculate a kernel matrix scales as a square of the number of compounds in the training set. The mean amount of computational resources needed to calculate each element of a kernel matrix also scales as a square of the average number of atoms in molecules. As a result, it becomes impractical to build 3D QSAR models using a training set with more than 300 medium-sized compounds (Baskin and Zhokhova 2013).

Another interesting combination approach, Shih et al. (2011) in 2011 proposed a combination of 3D QSAR methods in order to get better predictivity for the dataset. Authors proposed for the first time a combination approach to integrate the pharmacophore (PhModel) (Taha et al. 2008), CoMFA, and CoMSIA models for B-RAF (RAF family of serine/threonine kinases). The PhModel was implemented by the program Accelrys Discovery Studio 2.1. First, authors established ten PhModels and used them to align diverse inhibitor structures for generating the CoMFA and CoMSIA models. Then the partial least-square (PLS) method was used and known B-RAF inhibitors to validate the prediction ability of CoMFA and CoMSIA models. Finally, the goodness of hit (GH) test score was used as a

benchmark for appraising the prediction ability of CoMFA and CoMSIA models to screen a compound database. Thus, ten PhModels were generated based on the 27 training set inhibitors. Each PhModel included four features: hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic (HY), and ring aromatic (RA). The correlation coefficients for ten PhModels were very good and ranged from 0.964 to 0.902. Authors claim that this approach could be applied to screen inhibitor databases, optimize inhibitor structures, and identify novelty potent or specific inhibitors (Shih et al. 2011).

One more combination study for RAF inhibitors is worth to mention, was performed by Yang et al. (2011) which was applied as a combination of docking, molecular dynamics (MD), molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) calculations, and 3D QSAR analysis to investigate the detailed binding mode between B-RAF kinases with the series of inhibitors and also to find the key structural features affecting the inhibiting activities. Considering the difficulty in the accurate estimation of electrostatic interaction, the QM-polarized ligand docking and GBSA rescoring were applied to predict probable poses of these inhibitors bound into the active site of B-RAF kinase. To obtain the rational conformation for developing 3D QSAR models, authors applied the docking-based conformation selection strategy. Moreover, the detailed interactions were analyzed on the basis of the results from MD simulation and the free energy calculation for two inhibitors with much difference in their activity. Authors investigated 61 B-RAF inhibitors and developed CoMFA and CoMSIA models with $r^2 = 0.917$ and 0.940, respectively. In result, the structure-based 3D QSAR models provided a further structural analysis and modifiable information for understanding the SARs of these inhibitors. The important hydrophobic property of the 3-substitution of B-ring was required to be type 2 inhibitors. The five substitutable positions of the C-ring could be further modified. Authors concluded that the results obtained from the combined computational approach will be helpful for the rational design of novel type 2 RAF kinase inhibitors.

Recently, a group of computational scientists is proposed to apply a protein–protein interaction (PPI) analysis to target small molecules. Since currently in worlds life science, research is going on the booming of interactome studies, a lot of interactions can be measured in a high-throughput way, taking into account that large-scale datasets are already available. Studies show that many different types of interactions can be potential drug targets. This boom of HTS studies greatly broadens the drug target search space, which makes the drug target discovery difficult. In this case, computational methods are highly desired to efficiently provide candidates for further experiments and hold the promise to greatly accelerate the discovery of novel drug targets. Thus, Wang et al. (2016) published a study where they suggested a new method, where inhibition of protein–protein interaction (PPI) analysis offered as a promising source to improve the specificity of drugs with fewer adverse side effects. They proposed a machine learning method to predict PPI targets in a genomic-wide scale. Authors developed a computational method, named as PrePPItar (Wang et al. 2016), to predict PPIs as drug targets by uncovering the potential associations between drugs and PPIs (Fig. 7). Authors investigated the

databases and manually constructed a gold-standard positive dataset for drug and PPI interactions. Their effort leads to a dataset with 227 associations among 63 PPIs and 113 FDA-approved drugs and allowed them to build models and learn the association rules from the data. Also, authors were able to characterize drugs by profiling in chemical structure, drug ATC-code annotation, and side-effect space and represent PPI similarity by a symmetrical S-kernel based on protein amino acid sequence. At the end, a support vector machine (SVM) is used to predict novel associations between drugs and PPIs. The PrePPiTar method was validated on the well-established gold-standard dataset. Authors found that all chemical structures, drug ATC code, and side-effect information are predictive for PPI target. Authors claim that PrePPiTar can serve as a useful tool for PPI target discovery and provide a general heterogeneous data-integrative framework.

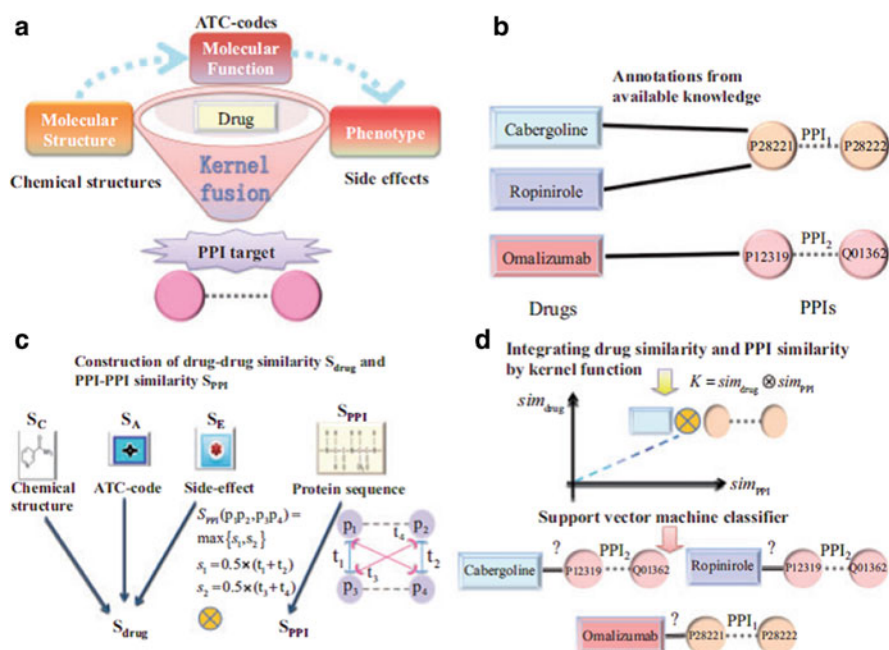


Fig. 7 The proposed flowchart for the PrePPiTar. **(a)** The schematic plot for our PrePPiTar method. PrePPiTar applies the kernel fusion method to integrate multiple information about drug, including chemical structure, ATC code, and drug side effect to detect the interactions between drugs and PPIs. **(b)** Collecting known associations between drugs and PPIs as gold-standard positives in a bipartite graph. **(c)** Calculating drug–drug and PPI–PPI similarity metrics, where t_i ; $i = 1; 2; 3; 4$ are the sequence similarity among proteins. **(d)** Relating the similarity among drugs and similarity among PPIs by Kronecker product kernel and applying SVM-based algorithm to predict the unknown associations between drugs and PPIs (Reproduced with Permission from reference (Wang et al. 2016). Copyright Oxford University Press, 2016)

3D QSARs and Combined Docking Studies of Nanostructured Materials

Nanomaterials are becoming an important component of the modern life and have been the subject of increasing number of investigations involving various areas of natural sciences and technology. However, theoretical modeling of physicochemical and biological activity of these species is still very scarce. The prediction of properties and activities of “classical” substances via correlating with molecular descriptors is a well-known procedure, by application QSAR and 3D QSAR methods. In spite of this, the application of QSAR for the nanomaterials is a very complicated task, because of “nonclassical” structure of these materials. Here, we would like to show first applications of the 3D QSAR and docking methods for nanostructured materials, which are nevertheless possible and can be useful in predicting their various properties and activities (toxicity).

Thus, one of the first 3D QSAR studies for nanostructured materials was provided in 2008. Durdagi et al. 2008a have investigated novel fullerene analogues as potential HIV PR inhibitors. It was the first work where authors analyzed nanostructured compounds for anti-HIV activity using protein–ligand docking and 3D QSAR approaches. Moreover, authors conducted MD simulations of ligand-free and the inhibitor bound HIV-1 PR systems to complement some previous studies and to provide proper input structure of HIV-1 PR in further docking simulations. Then, five different combinations of stereoelectronic fields of 3D QSAR/CoMSIA models were obtained from the set of biologically evaluated and computationally designed fullerene derivatives (where training set = 43 and test set = 6) in order to predict novel compounds with improved inhibition effect. The best 3D QSAR/CoMSIA model yielded a cross-validated r^2 value of 0.739 and a non-cross-validated r^2 value of 0.993. Authors stated that the derived model indicated the importance of steric (42.6%), electrostatic (12.7%), H-bond donor (16.7%), and H-bond acceptor (28.0%) contributions (Fig. 8). In addition, the derived contour plots together with applied de novo drug design were then used as pilot models for proposing the novel analogues with enhanced binding affinities. Interestingly, the investigated by authors the nanostructured compounds have triggered the interest of medicinal chemists to look for novel fullerene-type HIV-1 PR inhibitors possessing higher bioactivity. Later this year, authors published a second study for the same type of fullerene-based nanomaterials (Durdagi et al. 2008b).

The same group published in 2009 another study for fullerene derivatives but functionalized by amino acids (Durdagi et al. 2009). Authors used in silico screening approach in order to propose potent fullerene analogues as anti-HIV drugs. Two of the most promising derivatives showing significant binding scores were subjected to biological studies that confirmed the efficacy of the new compounds. The results showed that new leads can be discovered possessing higher bioactivity. Authors used docking approach together with MD simulations to get the best hits during the virtual screening.

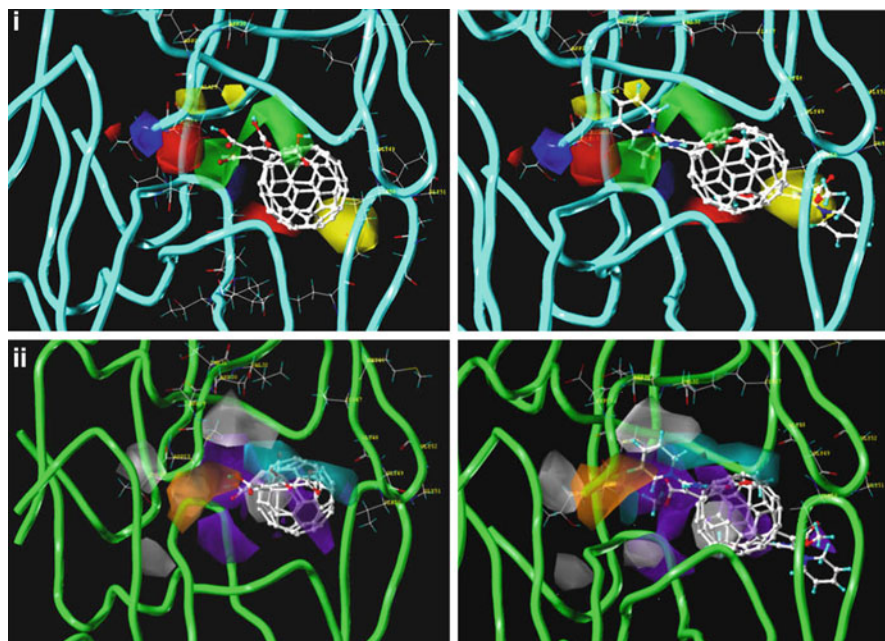


Fig. 8 (i) CoMSIA steric/electrostatic contour maps of template compound 23 (template compound; has best binding affinity in training set, *left* on the figure) and compound 36 (has worst binding affinity in training set, *right* on the figure). Sterically favored areas are shown in *green* color (contribution level of 80 %). Sterically disfavored areas are shown in *yellow* color (contribution level of 20 %). Positive potential favored areas are shown in *blue* color (contribution level of 80 %). Positive potential disfavored areas are shown in *red* color (contribution level of 20 %). (ii) CoMSIA H-bond donor/H-bond acceptor contour maps of compounds 23 and 36 (on the *left* and *right* of the figure, correspondingly). The individual contributions from the H-bond donor and H-bond acceptor favored and disfavored levels are fixed at 80 % and 20 %, respectively. The contours for H-bond donor favored fields have been shown in *cyan* color, while its disfavored fields have been shown in *purple* color. H-bond acceptor favored fields have been shown in *orange* color, while its disfavored fields have been shown in *white* color (Reproduced with Permission from ref (Durdagi et al. 2008a). Copyright Elsevier, 2008)

In 2011, the same group provided further analysis to design better anti-HIV fullerene-based inhibitors (Tzoupis et al. 2011). In this study authors employed a docking technique, two 3D QSAR models, MD simulations and the MM-PBSA method. In particular, authors investigated (1) hydrogen bonding (H-bond) interactions between specific fullerene derivatives and the protease, (2) the regions of HIV-1 PR that play a significant role in binding, (3) protease changes upon binding, and (4) various contributions to the binding free energy, in order to identify the most significant of them. The CoMFA and CoMSIA methods were applied too, to build 3D QSAR models, where good correlation coefficients were received, for both methods, $r^2 = 0.842$ and 0.928 , respectively. Authors claim that the computed binding free energies are in satisfactory agreement with the experimental results.

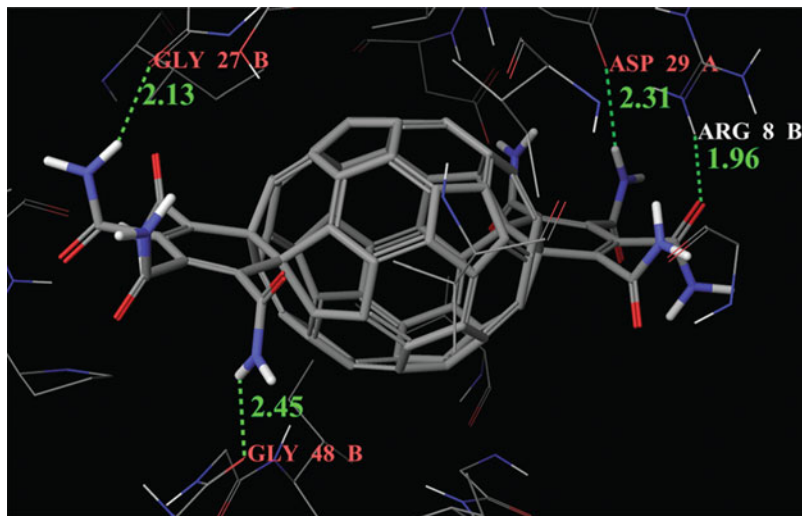


Fig. 9 The binding site interactions. H-bonds formed by the ligand 42 in the binding site (Glide) (Reproduced with Permission from reference (Ahmed et al. 2013). Copyright, Royal Society of Chemistry, 2013)

Another group published in 2013 a study that conducted a comprehensive investigation of fullerene analogues by combined computational approach including quantum chemical, molecular docking, and 3D descriptors-based QSAR (Ahmed et al. 2013). Authors stated that the protein–ligand docking studies and improved structure–activity models have been able both to predict binding affinities for the set of fullerene- C_{60} derivatives and to help in finding mechanisms of fullerene derivative interactions with human immunodeficiency virus type 1 aspartic protease, HIV-1 PR. Protein–ligand docking revealed several important molecular fragments that are responsible for the interaction with HIV-1 PR (Fig. 9). In addition, a density functional theory method has been utilized to identify the optimal geometries and predict physicochemical parameters of 49 studied compounds. The five-variable GA-MLRA-based model showed the best predictive ability ($r^2_{train} = 0.882$ and $r^2_{test} = 0.738$), with high internal and external correlation coefficients.

Calvaresi and Zerbetto (2010) published a study where they investigated a fullerene binding with a set of proteins. Authors investigated about 20 proteins that are known to modify their activity upon interaction with C_{60} . The set was examined using PatchDock (Schneidman-Duhovny et al. 2005) software with an algorithm that appraises quantitatively the interaction of C_{60} and the surface of each protein. The redundancy of the set allowed them to establish the predictive power of the approach that finds explicitly the most probable site where C_{60} docks on each protein. About 80 % of the known fullerene-binding proteins fall in the top 10 % of scorers. The close match between the model and experiments vouches for the accuracy of the model and validates its predictions. Authors identified the sites

of docking and discussed them in view of the existing experimental data available for protein – C60 interaction. In addition, authors identified new proteins that can interact with C60 and discussed for possible future applications as drug targets and fullerene derivative bioconjugate materials.

Later, Calvaresi and Zerbetto (2011) published another study, where they investigated the binding of fullerene C60 with 1099 proteins. They one more time confirmed that hydrophobic pockets of certain proteins can accommodate a carbon cage either in full or in part. Since the identification of proteins that are able to discriminate between different cages is an open issue, they were interested in investigating much larger library than in Calvaresi and Zerbetto (2010). Prediction of candidates is achieved with an inverse docking procedure that accurately accounts for (i) van der Waals interactions between the cage and the protein surface, (ii) desolvation free energy, (iii) shape complementarity, and (iv) minimization of the number of steric clashes through conformational variations. A set of 1099 protein structures is divided into four categories that either select C60 or C70 (p-C60 or p-C70) and either accommodate the cages in the same pocket or in different pockets. Thus, authors also confirmed the agreement with the experiments, where the KcsA potassium channel is predicted to have one of the best performances for both cages.

Recently, in 2015 Xavier et al. (Esposito et al. 2015) published a QSAR study of decorated carbon nanotube investigation for toxicity using 4D fingerprints. In this study, authors proposed detailed mechanisms of action, relating to nanotoxicity, for a series of decorated (functionalized) carbon nanotube complexes based on previously reported QSAR models. Possible mechanisms of nanotoxicity for six endpoints (bovine serum albumin, carbonic anhydrase, chymotrypsin, hemoglobin along with cell viability, and nitrogen oxide production) have been extracted from the corresponding optimized QSAR models. The molecular features relevant to each of the endpoint respective mechanism of action for the decorated nanotubes are also discussed in the study. Based on the molecular information contained within the optimal QSAR models for each nanotoxicity endpoint, either the decorator attached to the nanotube is directly responsible for the expression of a particular activity, irrespective of the decorator's 3D geometry and independent of the nanotube, or those decorators having structures that place the functional groups of the decorators as far as possible from the nanotube surface most strongly influence the biological activity.

A docking study, together with comprehensive DFT analysis was conducted by Saikia et al. (2013). Authors made a simulation to analyze the interaction of nanomaterials with biomolecular systems, where they performed density functional calculations on the interaction of pyrazinamide (PZA) drug with functionalized single-wall CNT (fSWCNT) as a function of nanotube chirality and length, followed by docking simulation of fSWCNT with pncA protein. The functionalization of pristine SWCNT that facilitates in enhancing the reactivity of the nanotubes and formation of such type of nanotube-drug conjugate is thermodynamically feasible. Docking studies predicted the plausible binding mechanism and suggested that PZA loaded fSWCNT facilitates in the target-specific binding of PZA within the protein

following a lock and key mechanism. Authors noticed that no major structural deformation in the protein was observed after binding with CNT, and the interaction between ligand and receptor is mainly hydrophobic in nature. Authors anticipate that these findings may provide new routes toward the drug delivery mechanism by CNTs with long-term practical implications in tuberculosis chemotherapy.

In another study, Turabekova et al. (2014) published a comprehensive study of carbon nanotube and pristine fullerene interactions with Toll-like receptors (TLRs), which are responsible for immune response. Having experimental data on hands and conducting comprehensive protein–ligand investigation, authors were able to show that CNT and fullerenes can bind to certain TLRs. Authors suggested a hypothetical model providing the potential mechanistic explanation for immune and inflammatory responses observed upon exposure to carbon nanoparticles. Specifically, authors performed a theoretical study to analyze CNT and C60 fullerene interactions with the available X-ray structures of TLR homo- and heterodimer extracellular domains. This assumption was based on the fact that similar to the known TLR ligands, both CNTs and fullerenes induce, in cells, the secretion of certain inflammatory protein mediators, such as interleukins and chemokines. These proteins are observed within inflammation downstream processes resulting from the ligand molecule-dependent inhibition or activation of TLR-induced signal transduction. The computational studies have shown that the internal hydrophobic pockets of some TLRs might be capable of binding small-sized carbon nanostructures (5,5 armchair SWCNTs containing 11 carbon atom layers and C60 fullerene). High binding scores and minor structural alterations induced in TLR ectodomains upon binding C60 and CNTs further supported the proposed hypothesis (Fig. 10). Additionally, the proposed hypothesis is strengthened by the indirect experimental

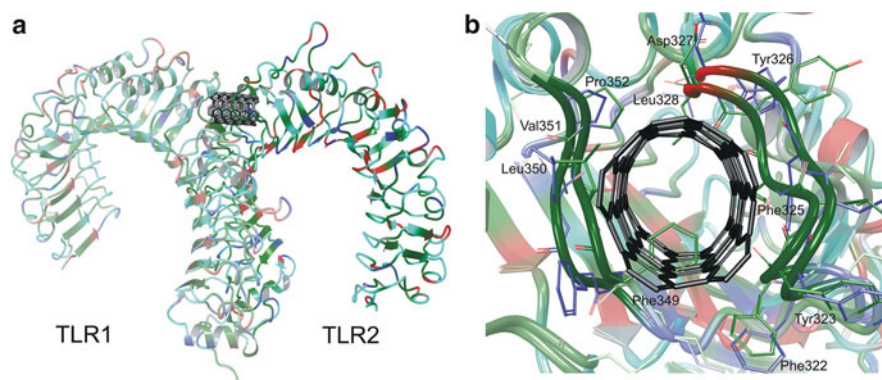


Fig. 10 5,5 CNT-bound TLR1/TLR2 ECDs: (a) 5,5 CNT is bound to the TLR1 and TLR2 ECD interface dimerization area, (b) aligned structures of TLR2 ECDs before (*green* carbon atoms) and after (*blue* carbon atoms) the impact OPLS2005 refinement upon binding 5,5 CNTs. The orientation of two parallel entrance loops and the side chains of hydrophobic Phe349, Phe325, and Leu328 preventing the nanotube from intrusion are shown to be optimized (Reproduced with Permission from reference (Turabekova et al. 2014). Copyright, Royal Society of Chemistry, 2014)

findings indicating that CNTs and fullerenes induce an excessive expression of specific cytokines and chemokines (i.e., IL-8 and MCP1).

Later, this kind of interaction was confirmed by MD simulation provided by Mozolewska et al. (2014). In this study, authors made an attempt to determine if the nanotubes could interfere with the innate immune system by interacting with TLRs. For this purpose, authors used the following TLR structures downloaded from the RCSB Protein Data Bank: TLR2 (3A7C), TLR4/MD (3FXI), TLR5 (3V47), TLR3 (2A0Z), and the complexes of TLR1/TLR2 (2Z7X) and TLR2/TLR6 (3A79). The results of steered molecular dynamics (SMD) simulations have shown that nanotubes interact very strongly with the binding pockets of some receptors (e.g., TLR2), which results in their binding to these sites without substantial use of the external force.

Concluding Remarks

In this chapter, we discussed 3D QSAR and protein-ligand docking methods, recent applications of them for conventional organic compounds design and for nanostructured materials. Despite of all pitfalls, the 3D QSAR approach confirmed the importance and value in drug design and medicinal chemistry. Moreover, the combination of 3D QSAR approach with other techniques, including protein–ligand docking study gives much better improvement in predictions of biologically active compounds and drug candidates. The development of methods for 3D QSAR still continues, giving improved predictions for conventional organic compounds. Thus, we believe that 3D QSAR methods in the near future will be able to encode and model various organic and nanomaterials for important biological and physicochemical property improvement.

Bibliography

- Ahmed, L., Rasulev, B., Turabekova, M., Leszczynska, D., & Leszczynski, J. (2013). Receptor-and ligand-based study of fullerene analogues: comprehensive computational approach including quantum-chemical, QSAR and molecular docking simulations. *Organic and Biomolecular Chemistry*, 11(35), 5798–5808.
- Ajmani, S., Jadhav, K., & Kulkarni, S. A. (2006). Three-dimensional QSAR using the k-nearest neighbor method and its interpretation. *Journal of Chemical Information and Modeling*, 46(1), 24–31.
- Amnerkar, N. D., & Bhusari, K. P. (2010). Synthesis, anticonvulsant activity and 3D-QSAR study of some prop-2-eneamido and 1-acetyl-pyrazolin derivatives of aminobenzothiazole. *European Journal of Medicinal Chemistry*, 45(1), 149–159.
- Araújo, J. Q., de Brito, M. A., Hoelz, L. V. B., de Alencastro, R. B., Castro, H. C., Rodrigues, C. R., & Albuquerque, M. G. (2011). Receptor-dependent (RD) 3D-QSAR approach of a series of benzylpiperidine inhibitors of human acetylcholinesterase (HuAChE). *European Journal of Medicinal Chemistry*, 46(1), 39–51.
- Baskin, I. I., & Zhokhova, N. I. (2013). The continuous molecular fields approach to building 3D-QSAR models. *Journal of Computer-Aided Molecular Design*, 27(5), 427–442.
- Calvaresi, M., & Zerbetto, F. (2010). Baiting proteins with C60. *ACS Nano*, 4(4), 2283–2299.

- Calvaresi, M., & Zerbetto, F. (2011). Fullerene sorting proteins. *Nanoscale*, 3(7), 2873–2881.
- Clark, M., Cramer, R. D., Jones, D. M., Patterson, D. E., & Simeroth, P. E. (1990). Comparative molecular field analysis (CoMFA). 2. Toward its use with 3D-structural databases. *Tetrahedron Computer Methodology*, 3(1), 47–59.
- Cramer, R., & Bunce, J. D. (1987). The Dylomms method: Initial results from a comparative study of approaches to 3d QSAR. In Hadzi D., Jerman-Blazic B. (eds) QSAR in drug design and toxicology (pp. 3–12).
- Cramer, R. D., Patterson, D. E., & Bunce, J. D. (1988). Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society*, 110(18), 5959–5967.
- Cruciani, G., Fontaine, F., & Pastor, M. (2004). *Almond*; 3.3.0. Perugia: Molecular Discovery Ltd.
- Datar, P. A., Khedkar, S. A., Malde, A. K., & Coutinho, E. C. (2006). Comparative residue interaction analysis (CoRIA): A 3D-QSAR approach to explore the binding contributions of active site residues with ligands. *Journal of Computer-Aided Molecular Design*, 20(6), 343–360.
- Dhaked, D. K., Verma, J., Saran, A., & Coutinho, E. C. (2009). Exploring the binding of HIV-1 integrase inhibitors by comparative residue interaction analysis (CoRIA). *Journal of Molecular Modeling*, 15(3), 233–245.
- Dixon, S. L., Smondryev, A. M., Knoll, E. H., Rao, S. N., Shaw, D. E., & Friesner, R. A. (2006). PHASE: A new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening. 1. Methodology and preliminary results. *Journal of Computer-Aided Molecular Design*, 20, 647–671.
- Doweyko, A. M. (1988). The hypothetical active site lattice. An approach to modelling active sites from data on inhibitor molecules. *Journal of Medicinal Chemistry*, 31(7), 1396–1406.
- Durdagi, S., Mavromoustakos, T., Chronakis, N., & Papadopoulos, M. G. (2008a). Computational design of novel fullerene analogues as potential HIV-1 PR inhibitors: Analysis of the binding interactions between fullerene inhibitors and HIV-1 PR residues using 3D QSAR, molecular docking and molecular dynamics simulations. *Bioorganic and Medicinal Chemistry*, 16(23), 9957–9974.
- Durdagi, S., Mavromoustakos, T., & Papadopoulos, M. G. (2008b). 3D QSAR CoMFA/CoMSIA, molecular docking and molecular dynamics studies of fullerene-based HIV-1 PR inhibitors. *Bioorganic and Medicinal Chemistry Letters*, 18(23), 6283–6289.
- Durdagi, S., Supuran, C. T., Strom, T. A., Doostdar, N., Kumar, M. K., Barron, A. R., Mavromoustakos, T., & Papadopoulos, M. G. (2009). In silico drug screening approach for the design of magic bullets: A successful example with anti-HIV fullerene derivatized amino acids. *Journal of Chemical Information and Modeling*, 49(5), 1139–1143.
- Eriksson, L., Jaworska, J., Worth, A. P., Cronin, M. T., McDowell, R. M., & Gramatica, P. (2003). Methods for reliability and uncertainty assessment and for applicability evaluations of classification-and regression-based QSARs. *Environmental Health Perspectives*, 111(10), 1361.
- Espósito, E. X., Hopfinger, A. J., Shao, C.-Y., Su, B.-H., Chen, S.-Z., & Tseng, Y. J. (2015). Exploring possible mechanisms of action for the nanotoxicity and protein binding of decorated nanotubes: Interpretation of physicochemical properties from optimal QSAR models. *Toxicology and Applied Pharmacology*, 288(1), 52–62.
- Evans, D. A., Doman, T. N., Thorner, D. A., & Bodkin, M. J. (2007). 3D QSAR methods: Phase and catalyst compared. *Journal of Chemical Information and Modeling*, 47(3), 1248–1257.
- Fontaine, F., Pastor, M., & Sanz, F. (2004). Incorporating molecular shape into the alignment-free GRid-IN dependent descriptors. *Journal of Medicinal Chemistry*, 47(11), 2805–2815.
- Fontaine, F., Pastor, M., Zamora, I., & Sanz, F. (2005). Anchor-GRIND: Filling the gap between standard 3D QSAR and the GRid-IN dependent descriptors. *Journal of Medicinal Chemistry*, 48(7), 2687–2694.
- Gohlke, H., & Klebe, G. (2002). DrugScore meets CoMFA: Adaptation of fields for molecular comparison (AFMoC) or how to tailor knowledge-based pair-potentials to a particular protein. *Journal of Medicinal Chemistry*, 45(19), 4153–4170.

- Goodford, P. J. (1985). A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *Journal of Medicinal Chemistry*, 28(7), 849–857.
- Gopalakrishnan, B., Aparna, V., Jeevan, J., Ravi, M., & Desiraju, G. (2005). A virtual screening approach for thymidine monophosphate kinase inhibitors as antitubercular agents based on docking and pharmacophore models. *Journal of Chemical Information and Modeling*, 45(4), 1101–1108.
- Gupta, P., Roy, N., & Garg, P. (2009). Docking-based 3D-QSAR study of HIV-1 integrase inhibitors. *European Journal of Medicinal Chemistry*, 44(11), 4276–4287.
- Halgren, T. A. (1996). Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17(5-6), 490–519.
- Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., & Banks, J. L. (2004). Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of Medicinal Chemistry*, 47, 1750–1759.
- Hansch, C., Leo, A., Hoekman, D., & Leo, A. (1995). *Exploring QSAR*. Washington, DC: American Chemical Society.
- Hopfinger, A. J. (1980). A QSAR investigation of dihydrofolate reductase inhibition by Baker triazines based upon molecular shape analysis. *Journal of the American Chemical Society*, 102(24), 7196–7206.
- Hu, R., Barbault, F., Delamar, M., & Zhang, R. (2009). Receptor- and ligand-based 3D-QSAR study for a series of non-nucleoside HIV-1 reverse transcriptase inhibitors. *Bioorganic and Medicinal Chemistry*, 17(6), 2400–2409.
- Isayev, O., Rasulev, B., Gorb, L., & Leszczynski, J. (2006). Structure-toxicity relationships of nitroaromatic compounds. *Molecular Diversity*, 10(2), 233–245.
- Jha, K. K., Samad, A., Kumar, Y., Shaharyar, M., Khosa, R. L., Jain, J., Kumar, V., & Singh, P. (2010). Design, synthesis and biological evaluation of 1, 3, 4-oxadiazole derivatives. *European Journal of Medicinal Chemistry*, 45(11), 4963–4967.
- Karelson, M., Lobanov, V. S., & Katritzky, A. R. (1996). Quantum-chemical descriptors in QSAR/QSPR studies. *Chemical Reviews*, 96(3), 1027–1044.
- Kim, K. H. (2001). Thermodynamic aspects of hydrophobicity and biological QSAR. *Journal of Computer-Aided Molecular Design*, 15(4), 367–380.
- Kim, K. H., Greco, G., & Novellino, E. (1998). A critical review of recent CoMFA applications. In Kubinyi H., Folkers G., Martin Y.C. (eds) In *3D QSAR in drug design* (pp. 257–315). Springer.
- Klebe, G. (2006). Virtual ligand screening: Strategies, perspectives and limitations. *Drug Discovery Today*, 11(13), 580–594.
- Klebe, G., Abraham, U., & Mietzner, T. (1994). Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *Journal of Medicinal Chemistry*, 37(24), 4130–4146.
- Kubinyi, H. (1997a). QSAR and 3D QSAR in drug design Part 1: Methodology. *Drug Discovery Today*, 2(11), 457–467.
- Kubinyi, H. (1997b). QSAR and 3D QSAR in drug design part 2: Applications and problems. *Drug Discovery Today*, 2(12), 538–546.
- Kumar, B. P., Sankar, G., Baig, R. N., & Chandrashekar, S. (2009). Novel Biginelli dihydropyrimidines with potential anticancer activity: A parallel synthesis and CoMSIA study. *European Journal of Medicinal Chemistry*, 44(10), 4192–4198.
- Li, H., Sutter, J., & Hoffmann, R. (2000). HypoGen: An automated system for generating 3D predictive pharmacophore models. In *Pharmacophore perception, development, and use in drug design* (Vol. 2, p. 171).
- Lill, M. A. (2007). Multi-dimensional QSAR in drug discovery. *Drug Discovery Today*, 12(23), 1013–1017.
- Lushington, G. H., Guo, J.-X., & Wang, J. L. (2007). Whither combine? New opportunities for receptor-based QSAR. *Current Medicinal Chemistry*, 14(17), 1863–1877.
- Martin, Y. C. (2009). Let's not forget tautomers. *Journal of Computer-Aided Molecular Design*, 23(10), 693–704.

- Molecular Operating Environment (MOE). (2016). Chemical Computing Group Inc. 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7.
- Moro, S., Braiuca, P., Deflorian, F., Ferrari, C., Pastorin, G., Cacciari, B., Baraldi, P. G., Varani, K., Borea, P. A., & Spalluto, G. (2005). Combined target-based and ligand-based drug design approach as a tool to define a novel 3D-pharmacophore model of human A3 adenosine receptor antagonists: Pyrazolo [4, 3-e] 1, 2, 4-triazolo [1, 5-c] pyrimidine derivatives as a key study. *Journal of Medicinal Chemistry*, *48*(1), 152–162.
- Moro, S., Bacilieri, M., Cacciari, B., Bolcato, C., Cusan, C., Pastorin, G., Klotz, K.-N., & Spalluto, G. (2006). The application of a 3D-QSAR (autoMEP/PLS) approach as an efficient pharmacodynamic-driven filtering method for small-sized virtual library: Application to a lead optimization of a human A 3 adenosine receptor antagonist. *Bioorganic and Medicinal Chemistry*, *14*(14), 4923–4932.
- Mozolewska, M. A., Krupa, P., Rasulev, B., Liwo, A., & Leszczynski, J. (2014). Preliminary studies of interaction between nanotubes and toll-like receptors. *Task Quarterly*, *18*(4), 351–355.
- Oprea, T. I., & Matter, H. (2004). Integrating virtual screening in lead discovery. *Current Opinion in Chemical Biology*, *8*(4), 349–358.
- Ortiz, A. R., Pisabarro, M. T., Gago, F., & Wade, R. C. (1995). Prediction of drug binding affinities by comparative binding energy analysis. *Journal of Medicinal Chemistry*, *38*(14), 2681–2691.
- Pajeva, I. K., Globisch, C., & Wiese, M. (2009). Combined pharmacophore modeling, docking, and 3D QSAR studies of ABCB1 and ABCC1 transporter inhibitors. *ChemMedChem*, *4*(11), 1883–1896.
- Pastor, M., Cruciani, G., McLay, I., Pickett, S., & Clementi, S. (2000). GRIND-INdependent descriptors (GRIND): A novel class of alignment-independent three-dimensional molecular descriptors. *Journal of Medicinal Chemistry*, *43*, 3233–3243.
- Patel, P. D., Patel, M. R., Kaushik-Basu, N., & Talele, T. T. (2008). 3D QSAR and molecular docking studies of benzimidazole derivatives as hepatitis C virus NS5B polymerase inhibitors. *Journal of Chemical Information and Modeling*, *48*(1), 42–55.
- Perkins, R., Fang, H., Tong, W., & Welsh, W. J. (2003). Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology. *Environmental Toxicology and Chemistry*, *22*(8), 1666–1679.
- Perola, E. (2006). Minimizing false positives in kinase virtual screens. *Proteins: Structure, Function, and Bioinformatics*, *64*(2), 422–435.
- Podlogar, B. L., & Ferguson, D. M. (2000). QSAR and CoMFA: A perspective on the practical application to drug discovery. *Drug Design and Discovery*, *17*(1), 4.
- Polanski, J., Gieleciak, R., & Bak, A. (2002). The comparative molecular surface analysis (CoMSA)-a nongrid 3D QSAR method by a coupled neural network and PLS system: Predicting p K a values of benzoic and alkanolic acids. *Journal of Chemical Information and Computer Science*, *42*(2), 184–191.
- Polanski, J., Bak, A., Gieleciak, R., & Magdziarz, T. (2006). Modeling robust QSAR. *Journal of Chemical Information and Modeling*, *46*(6), 2310–2318.
- Pulla, V. K., Sriram, D. S., Viswanatha, S., Sriram, D., & Yogeewari, P. (2016). Energy-based pharmacophore and three-dimensional quantitative structure-activity relationship (3D-QSAR) modeling combined with virtual screening to identify novel small-molecule inhibitors of silent mating-type information regulation 2 homologue 1 (SIRT1). *Journal of Chemical Information and Modeling*, *56*(1), 173–187.
- Puzyn, T., Rasulev, B., Gajewicz, A., Hu, X., Dasari, T. P., Michalkova, A., Hwang, H.-M., Toropov, A., Leszczynska, D., & Leszczynski, J. (2011). Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. *Nature Nanotechnology*, *6*(3), 175–178.
- Ragno, R., Artico, M., De Martino, G., La Regina, G., Coluccia, A., Di Pasquali, A., & Silvestri, R. (2005). Docking and 3-D QSAR studies on indolyl aryl sulfones. Binding mode exploration at the HIV-1 reverse transcriptase non-nucleoside binding site and design of highly active N-(2-hydroxyethyl) carboxamide and N-(2-hydroxyethyl) carbohydrazide derivatives. *Journal of Medicinal Chemistry*, *48*(1), 213–223.

- Raparti, V., Chitre, T., Bothara, K., Kumar, V., Dangre, S., Khachane, C., Gore, S., & Deshmane, B. (2009). Novel 4-(morpholin-4-yl)-N'-(arylidene) benzohydrazides: Synthesis, antimycobacterial activity and QSAR investigations. *European Journal of Medicinal Chemistry*, *44*(10), 3954–3960.
- Rasulev, B., Kušić, H., Leszczynska, D., Leszczynski, J., & Koprivanac, N. (2010). QSAR modeling of acute toxicity on mammals caused by aromatic compounds: The case study using oral LD 50 for rats. *Journal of Environmental Monitoring*, *12*(5), 1037–1044.
- Ravichandran, V., Kumar, B. P., Sankar, S., & Agrawal, R. (2009). Predicting anti-HIV activity of 1, 3, 4-thiazolidinone derivatives: 3D-QSAR approach. *European Journal of Medicinal Chemistry*, *44*(3), 1180–1187.
- Robinson, D. D., Winn, P. J., Lyne, P. D., & Richards, W. G. (1999). Self-organizing molecular field analysis: A tool for structure-activity studies. *Journal of Medicinal Chemistry*, *42*(4), 573–583.
- Saikia, N., Rajkhowa, S., & Deka, R. C. (2013). Density functional and molecular docking studies towards investigating the role of single-wall carbon nanotubes as nanocarrier for loading and delivery of pyrazinamide antitubercular drug onto pncA protein. *Journal of Computer-Aided Molecular Design*, *27*(3), 257–276.
- Sakkiah, S., Thangapandian, S., John, S., Kwon, Y. J., & Lee, K. W. (2010). 3D QSAR pharmacophore based virtual screening and molecular docking for identification of potential HSP90 inhibitors. *European Journal of Medicinal Chemistry*, *45*(6), 2132–2140.
- Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., & Wolfson, H. J. (2005). PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Research*, *33*(Suppl 2), W363–W367.
- Shih, K.-C., Lin, C.-Y., Zhou, J., Chi, H.-C., Chen, T.-S., Wang, C.-C., Tseng, H.-W., & Tang, C.-Y. (2011). Development of novel 3D-QSAR combination approach for screening and optimizing B-Raf inhibitors in silico. *Journal of Chemical Information and Modeling*, *51*(2), 398–407.
- Silverman, B., & Platt, D. E. (1996). Comparative molecular moment analysis (CoMMA): 3D-QSAR without molecular superposition. *Journal of Medicinal Chemistry*, *39*(11), 2129–2140.
- Sun, J., Cai, S., Yan, N., & Mei, H. (2010). Docking and 3D-QSAR studies of influenza neuraminidase inhibitors using three-dimensional holographic vector of atomic interaction field analysis. *European Journal of Medicinal Chemistry*, *45*(3), 1008–1014.
- Taha, M. O., Dahabiyeh, L. A., Bustanji, Y., Zalloum, H., & Saleh, S. (2008). Combining ligand-based pharmacophore modeling, quantitative structure-activity relationship analysis and in silico screening for the discovery of new potent hormone sensitive lipase inhibitors. *Journal of Medicinal Chemistry*, *51*, 6478–6494.
- Todeschini, R., & Gramatica, P. (1998). *3D QSAR in drug design* (Vol. 2, pp. 355–360). Dordrecht: Kluwer/ESCOM.
- Totrov, M. (2008). Atomic property fields: Generalized 3D pharmacophoric potential for automated ligand superposition, pharmacophore elucidation and 3D QSAR. *Chemical Biology and Drug Design*, *71*(1), 15–27.
- Tripos. (2006). SYBYL, version 7.3, 2006, St. Louis.
- Turabekova, M., Rasulev, B., Theodore, M., Jackman, J., Leszczynska, D., & Leszczynski, J. (2014). Immunotoxicity of nanoparticles: A computational study suggests that CNTs and C 60 fullerenes might be recognized as pathogens by Toll-like receptors. *Nanoscale*, *6*(7), 3488–3495.
- Tzoupis, H., Leonis, G., Durdagi, S., Mouchlis, V., Mavromoustakos, T., & Papadopoulos, M. G. (2011). Binding of novel fullerene inhibitors to HIV-1 protease: Insight through molecular dynamics and molecular mechanics Poisson–Boltzmann surface area calculations. *Journal of Computer-Aided Molecular Design*, *25*(10), 959–976.
- Veber, D. F., Johnson, S. R., Cheng, H.-Y., Smith, B. R., Ward, K. W., & Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *Journal of Medicinal Chemistry*, *45*(12), 2615–2623.
- Virasodia, V., Pissurlenkar, R. R., Manvar, D., Dholakia, C., Adlakhia, P., Shah, A., & Coutinho, E. C. (2008). Synthesis, screening for antitubercular activity and 3D-QSAR studies of substituted

- N-phenyl-6-methyl-2-oxo-4-phenyl-1, 2, 3, 4-tetrahydro-pyrimidine-5-carboxamides. *European Journal of Medicinal Chemistry*, 43(10), 2103–2115.
- Walters, D. E., & Hinds, R. M. (1994). Genetically evolved receptor models: A computational approach to construction of receptor models. *Journal of Medicinal Chemistry*, 37(16), 2527–2536.
- Wang, Y.-C., Chen, S.-L., Deng, N.-Y., & Wang, Y. (2016). Computational probing protein–protein interactions targeting small molecules. *Bioinformatics*, 32(2), 226–234.
- Wise, M., Cramer, R. D., Smith, D., & Exman, I. (1983). *Progress in three-dimensional drug design: The use of real-time colour graphics and computer postulation of bioactive molecules in DYLOMMS*. Amsterdam: Elsevier.
- Yang, S.-Y. (2010). Pharmacophore modeling and applications in drug discovery: Challenges and recent advances. *Drug Discovery Today*, 15(11), 444–450.
- Yang, Y., Qin, J., Liu, H., & Yao, X. (2011). Molecular dynamics simulation, free energy calculation and structure-based 3D-QSAR studies of B-RAF kinase inhibitors. *Journal of Chemical Information and Modeling*, 51(3), 680–692.
- Zhokhova, N. I., Baskin, I. I., Bakhronov, D. K., Palyulin, V. A., & Zefirov, N. S. (2009). Method of continuous molecular fields in the search for quantitative structure-activity relationships. *Doklady Chemistry*, 429(1), 273–276.
- Zhou, P., Tian, F., & Li, Z. (2007). Three dimensional holographic vector of atomic interaction field (3D-HoVAIF). *Chemometrics and Intelligent Laboratory Systems*, 87(1), 88–94.

Janna Hastings and Christoph Steinbeck

Contents

Introduction	2164
What Are Ontologies?	2164
Ontology Languages and Logical Expressivity	2165
Ontology Content and Metadata	2167
Chemical Ontologies and Their Uses	2168
Chemical Text Mining	2170
Semantic Similarity	2170
Enrichment Analysis	2172
Cheminformatics and the Semantic Web	2173
Limitations and Research Areas	2175
Conclusions	2177
Bibliography	2178

Abstract

Ontologies are structured controlled vocabularies which encode domain knowledge, backed by sophisticated logic-based computational tools. They enable knowledge-based applications which harness automated reasoning for inference and knowledge discovery. They also enable the semantic and standard annotation of large-scale data, which is ever relevant in the modern age of increased high-throughput data generation and sharing in scientific research. Established chemical ontologies include ChEBI, which encodes the structural classification of chemical entities of biological interest together with their roles. More recently, the chemical information ontology was created to standardize the annotation of cheminformatics software and descriptors. In this chapter, the technology, structure and applications of ontologies within cheminformatics will be described.

J. Hastings (✉) • C. Steinbeck
European Bioinformatics Institute, Hinxton, UK
e-mail: hastings@ebi.ac.uk; steinbeck@ebi.ac.uk

Introduction

Computational methods have become essential to filter, annotate, classify, and discover patterns within large-scale scientific data. *Ontologies* are frequently a core component of such large-scale data management platforms: computable logic-based representations of human domain knowledge in the form of structured controlled vocabularies with metadata, relationships, and other logical axioms (Courtot et al. 2011; Harland et al. 2011), which are used to annotate the data around those entities which are key to the domain of investigation.

Ontologies serve several distinct purposes in research data management. Firstly, ontologies provide standardized terminology and identifiers for entities in a domain, enabling different sources of data to be aggregated, e.g., across distinct databases. This is crucially important as the number of databases covering a domain of interest is proliferating. Secondly, they organize their entities hierarchically, which enables flexible aggregation around scientifically meaningful groupings. Such aggregation is an essential component in several approaches to data-driven scientific discovery. Thirdly, they facilitate browsing and searching, driving the interfaces of several bioinformatics databases. Further, their underlying representation in logical languages allows intelligent applications to perform complex reasoning tasks such as checking for errors and inconsistencies in the represented knowledge.

In the next section, we will introduce ontologies and describe the range of technologies that support and enable their representation. The following section will focus on several examples of ontologies within cheminformatics and chemistry and the application scenarios they are being used in. Finally, the last section will discuss limitations and currently open technological research areas.

What Are Ontologies?

Historically, ontology was a branch of philosophy concerned with studying the nature of existence or *what is*. More recently, the term “ontology” was conscripted as a name for the computational artifacts that were created to represent human domain knowledge for artificial intelligence within computer science, where the objective was to enable logic-based reasoning processes to derive inferences that simulate human reasoning processes (Gruber 2009). In parallel to these developments in computer science, within biology, the genomics revolution that followed the Human Genome Project saw the creation of ontologies as terminological artifacts to solve challenges in the integration of database content relating to genetic discoveries (The Gene Ontology Consortium 2000; Hunter 2002). These different intellectual paradigms all feed into the modern discipline of biomedical ontology or *bio-ontology* (Smith 2003).

Where large-scale data are labeled and populated using unconstrained text, different terminology is often used for similar or identical things. This is a feature of natural human language, and usually humans are easily able to resolve ambiguous

usages of terms and discrepant labels. However, due to the sheer volumes of research data being generated, it is necessary to develop computational methods of aggregating and aligning like with like. One approach to addressing this issue is to adopt shared standards for reference such that the same thing is always referred to by the same identifier. Harnessing such a standard across multiple different databases increases the value of the standard, allowing for easier cross-database integration and querying. Increasingly, computational research in the life sciences requires data integration across multiple disparate fields and methodological approaches in order to gain insight into underlying mechanisms. This is the case, for example, in the development of predictive models to enable personalized and translational medicine. Research results from diverse disciplines such as genetics, molecular biology, physiology, chemistry, psychology, and medicine have to be integrated in order to build a coherent picture of what is known in order to address key research gaps, and ontologies are meeting this need.

Ontologies are the structure in which the meaning of terminology used in a domain can be formalized around identifiers that can be embedded into databases. Bio-ontologies represent biological, medical, or chemical entities of scientific interest together with their properties and the relationships between them in a computationally processable language, enhanced with metadata such as definitions, synonyms, and references to databases. The longest-standing example of a successful ontology of this type is the Gene Ontology (GO) (The Gene Ontology Consortium 2000), which is used *inter alia* to unify annotations between disparate biological databases and for the statistical analysis of large-scale genetic data to identify genes that are significantly enriched for specific functions. The Gene Ontology describes the functions, processes, and cellular components of genes and gene products such as proteins.

Ontology Languages and Logical Expressivity

Bio-ontologies are developed and exchanged in a shared ontology language such as the Web Ontology Language (OWL), version 2 (Grau et al. 2008) or the OBO format (The Gene Ontology Consortium 2012). The OBO language was originally developed within the Gene Ontology project (The Gene Ontology Consortium 2000) in order to provide a more human-readable language than the then available technical alternatives that were the precursors of the OWL language. Many tools still rely on the OBO language with its relatively simple graph-based ontology structure, but the OBO and OWL languages are now largely interconvertible (Hoehndorf et al. 2010), and increasingly the community has moved toward OWL for ontology maintenance in order to harness better performance and expressivity.

OWL ontologies consist of several distinct components. These include classes, which are entities in the domain, known as “terms” in the OBO language, and properties, which are called “relations” in OBO. Metadata such as names and synonyms are included as annotations in OWL and have no impact on the logical

statements of the language. In OWL, properties represent relationships between individuals that belong to classes rather than classes themselves. As such, it is necessary to specify a quantification of each relationship for the members of the class represented by a property axiom. There are two options: a property axiom can express an *existential quantification*, which means that all members of the first class are related to *some* member of the other class, or a *value restriction*, which means that all members of the first class are *only* related to members of that other class (for the specified type of property). For example, you would capture the knowledge that organic molecules have as atomic parts some carbon atom with an existential quantification, since the molecules should have as atomic part some carbon atoms but may have parts of other types, while on the other hand you would capture the knowledge that hydrocarbons have only carbon and hydrogen as atomic parts with a value restriction, since such molecules may only have atomic parts of these types and not other types.

OWL is based on description logics (Baader et al. 2003), a family of decidable logical languages which has been developed specifically for the representation of large-scale terminological knowledge such as is found within large biomedical vocabularies. Logical languages are defined by the types of logical axioms that they support. OWL supports many different types of logical axiom, and the combination of axiom types that is used in a given ontology defines the expressivity of that ontology. There is a trade-off between expressive power and the tractability of a logic-based language. Increasing the expressivity of the language usually means that it takes longer to perform reasoning tasks.

The axiom types available in the OWL language include:

- Declaring and naming atomic classes
- OWL: Thing, the “top” class that is a superclass of every other class in an ontology
- OWL: Nothing, the “bottom” class that is a subclass of every other class in an ontology and is used to highlight inconsistent classes when there are errors in the ontology
- Intersection (AND)
- Union (OR)
- Negation (complement, NOT)
- One of (enumeration)
- Restriction on object property: some values
- Restriction on object property: all values
- Restriction on object property: minimum cardinality
- Restriction on object property: maximum cardinality
- Restriction on data property: value or value range

A powerful feature of OWL is the ability to perform *automatic classification* using fast-performing OWL reasoners. Reasoners are software toolkits that are able to compute inferences on the underlying logic used by the ontology – i.e.,

to logically deduce the consequences of the knowledge that is expressed, including inconsistencies. Examples of reasoners that work with OWL ontologies are Fact++ (Tsarkov and Horrocks 2006), Pellet (Sirin et al. 2007), and HermiT (Shearer et al. 2008), each of which supports different ranges of operators and may have different performance profiles on different reasoning tasks. For an example of OWL reasoning, given the following axioms:

IronAtom subclassOf MetalAtom (1)

MetallicCompound equivalentTo Compound and hasAtom
some MetalAtom (2)

IronOxide subclassOf Compound and hasAtom some IronAtom (3)

An OWL reasoner can automatically infer by 1–3 that IronOxide is a subclass of MetallicCompound.

OWL has *open-world* semantics, which means that inferences can only be drawn based on what is explicitly captured in the knowledge base and that absence of additional information has no implication that the information does not exist. For example, given a statement that *Square* and *Circle* are subclasses of *Shape* in a particular knowledge base, together with the knowledge that some entity *A* has_shape *C*, we nevertheless cannot infer that *C* is either a *Square* or a *Circle*, as we do not know anything about which other shapes may exist that have not been specified (i.e., in the open world). If this type of inference is desired, it is necessary to add an axiom specifying that all shapes are either squares or circles. This is called a *closure axiom*.

The primary editor that is used to maintain OWL ontologies is Protégé (2013). Protégé provides an extensive framework for ontology development, allowing editing of classes and properties, reasoner integration, and ontology visualization, as well as being supported by additional custom plug-ins.

Ontology Content and Metadata

Typically, modern bio-ontologies harness a relatively simple logical expressivity for a relatively large number of axioms (classes and relationships). In this fashion, they are able to encode a broad range of scientific knowledge in these domains – including chemistry – while remaining tractable for the reasoners and other applications, such as editors and visualizations. Alongside the classification hierarchy and the interrelationships between classes, each class typically contains some or all of the following metadata.

- A unique, usually alphanumeric, identifier. Each class is assigned an identifier which is persistent and unique. The identifier is usually stable with regard to changes in the remainder of the metadata associated with the class.

- A label which is also unique across the ontology and unambiguous with regard to what the class means. For an example of disambiguation that might be useful in assigning the label for a class, “nucleus” in an ontology of cellular components differs from “nucleus” in an ontology of atomic parts; thus, they might better be labeled “cell nucleus” and “atomic nucleus,” respectively.
- A definition, which should clearly indicate the intended meaning of the class, the way in which it is to be used, and how it differs from its immediate parent class.
- One or more synonyms or alternative labels. Synonyms may be used to capture short-hand descriptions which are in common use for reference to that class even if the synonym is not unambiguous. They may also capture acronyms or other ways in which the class may be referred to in text. Having a rich set of such synonyms for each class enhances the use of ontologies for automated text and data mining algorithms.
- One or more cross-references to databases and other ontologies. Cross-references enable broader data integration options than are supported by labels and synonym sets alone.

Chemical Ontologies and Their Uses

With an increase in availability of open, large-scale chemical data through public domain projects such as PubChem (Bolton et al. 2008), making sense of and effectively organizing the knowledge in databases for search and retrieval have become a challenge. Traditional large-scale data management methods in chemistry include chemical structure-based algorithmic and statistical methods for the construction of hierarchies and similarity landscapes. These techniques are essential not only for human interaction for effective browsing and searching but also in computational methods for interpreting raw data, such as for evaluating bioactivity patterns associated with chemical structural features (Wegner et al. 2012). Ontology-based classifications provide a complementary method for organizing data which is more text and knowledge oriented than algorithmic clustering methods (Hastings et al. 2012).

The Chemical Entities of Biological Interest (ChEBI) ontology (Hastings et al. 2013) was initially developed as a dictionary for chemical knowledge within biology, to enable standardization of chemical references within bioinformatics databases. This was facilitated by the use of semantics-free, stable identifiers for chemical entities, by the inclusion of extensive synonyms and cross-references, and through striving to represent community agreement about the entities in a given domain such that the ontology could be widely used. ChEBI is the primary chemical annotation and identification standard in many databases including MetaboLights (Haug et al. 2012), UniProt (The UniProt Consortium 2015), BioModels (Li et al. 2010), and Reactome (Matthews et al. 2009) and in the Gene Ontology (Hill et al. 2013). ChEBI is also listed as a secondary identifier for chemical information

in many additional databases such as KEGG (Kanehisa et al. 2006), DrugBank (Wishart et al. 2006), and HMDB (Wishart et al. 2009), enabling ontology-based data integration. As of August 2015, ChEBI contained just over 46,000 fully annotated entities.

ChEBI provides both a structure-based and a role-based sub-ontology, with roots “chemical entity” and “role,” respectively. The “chemical entity” sub-ontology of ChEBI classifies molecular entities into structure-based classes by virtue of shared structural features. Included in the chemical entity ontology are fully specified molecules such as transplatin and erythromycin A and structure-based classes such as steroid and macrolide. These molecules may be furthermore interrelated with structural relationships such as “is enantiomer of” and “is tautomer of” which are appropriate between entities with closely related chemical structures. Included in the “role” sub-ontology are drug usage classes such as antidepressant and antifungal; chemical reactivity classes such as solvent, acid, and base; and biological activity classes such as hormone (Batchelor et al. 2010). Where specific targets are known, the role ontology includes classes which refer to the target specifically, such as “cyclooxygenase inhibitor” for paracetamol. Chemical entities are linked to roles using the “has role” relationship.

An overview of the content of ChEBI showing selected classes from the structure-based and role-based classifications is illustrated in Fig. 1.

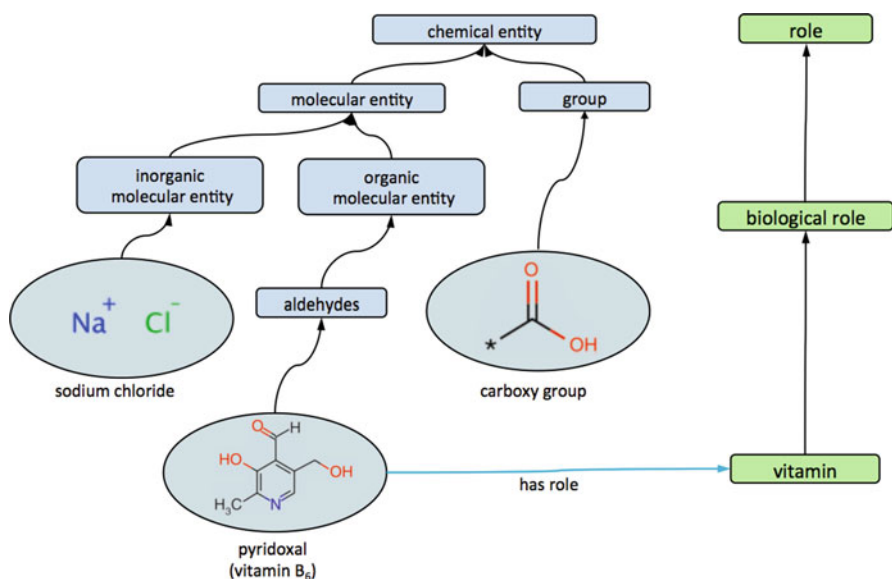


Fig. 1 Core divisions in content in the ChEBI ontology between chemical entities and roles. Fully specified chemical entities such as the vitamin pyridoxal are assigned role categories with the “has role” relationship and structure-based classes with the “is a” relationship. The “has part” relationship may be used to specify important parts of structure-based classes

Chemical Text Mining

ChEBI has been applied to many different use cases outside of its original remit as a dictionary for chemical references in biological databases. The extensive annotation of synonyms and other metadata has made it an important resource for chemical text mining – the automated identification of chemical mentions in text – as implemented in tools such as OSCAR (Jessop et al. 2011).

Results across many different research areas are reported in text in the primary scientific literature. However, the body of literature is growing at such a rate that it is not possible to read all the relevant articles in any field as the volumes are too. Computational processing to identify articles of interest or to amass all the literature in which specific entities are mentioned is increasingly necessary. In the literature, classes of chemical entities are often referenced rather than individual fully specified molecules. Biological knowledge such as the actions of enzymes in biological pathways is also often described in terms of classes or patterns of behavior rather than individual molecules. For this reason, ChEBI's extensive ontology of chemical classes has proved very useful in chemical text mining (Corbett and Murray-Rust 2006).

A closely related task is that of calculating systematic names for chemical structures and calculating structures from systematic chemical names. IUPAC naming rules for chemicals, as described in the “Gold Book” (McNaught and Wilkinson 1997) and implemented in tools including the open source Opsin (Lowe et al. 2011), provide a method for obtaining a systematic name from a given chemical structure and for interpreting a name to yield an appropriate chemical structure. However, it should be noted that systematic names can be unwieldy and lengthy and that chemists in many cases prefer to use shorter trivial names such as “caffeine.” Such trivial names cannot typically be automatically computed and need to be stored in a knowledge base such as ChEBI.

Text processing and knowledge discovery has been enhanced by implementations of ontology-based similarity for improved classification and text mining (Ferreira and Couto 2010). Similarity has many important applications in life science research. Most importantly, similar entities can be expected to behave similarly in similar contexts (Wegner et al. 2012). We can learn something about unknown entities and make predictions about their behavior, by examining knowledge about similar entities. But for this endeavor to yield the best results, the measures used for similarity must have real domain relevance.

Semantic Similarity

Semantic similarity is named as such because it encodes a measure of similarity that uses knowledge over and above the purely structural features of an entity. For example, the two words “cell” and “cell” are structurally identical, but can have very different semantics – the one means “cell” as in organism parts and the other

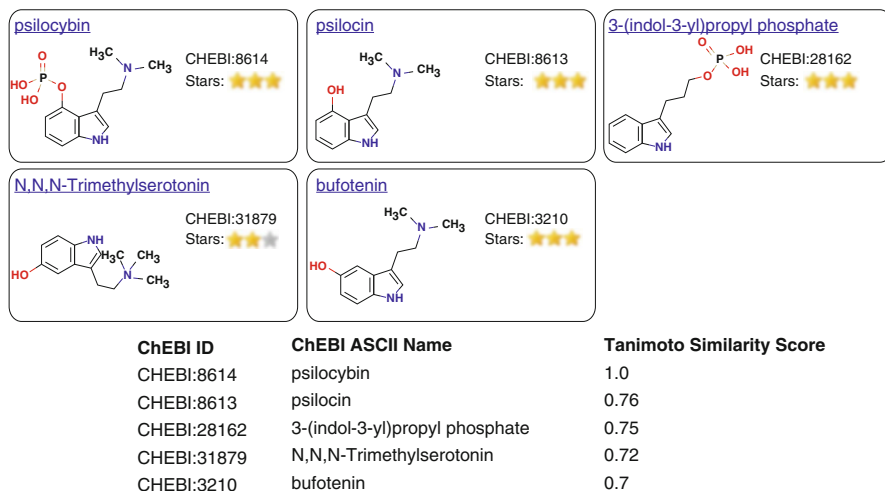


Fig. 2 The chemical structures for psilocybin (CHEBI:8614) and those structures in ChEBI which are similar including psilocin (CHEBI:8613) and 3-(indol-3-yl)propyl phosphate (CHEBI:28162)

“cell” as in for prisoners. Ontologies aim to encode relevant information about the meaning of the entities they represent, through the relationships which are captured for a given class and, in the case of more expressive OWL ontologies, additional features of the logical axioms. In chemistry, the most commonly used measure for calculating similarity between chemical entities uses features of the chemical structure to compute a similarity metric. Representations of chemical structures include the atoms and bonds which constitute the entity, with information about their types and how they are locally arranged. Similarity algorithms pick out features from such representations and binarize them into a fixed-length bit array (a string of 1’s and 0’s) or “fingerprint,” which is then able to be rapidly compared for similarity with the fingerprint from another structure using, for example, a Tanimoto score.

Psilocybin (CHEBI:8614), psilocin (CHEBI:8613), and 3-(indol-3-yl)propyl phosphate (CHEBI:28162) are structurally very similar, and their Tanimoto score along with other similar structures is given in Fig. 2.

ChEBI’s similarity search function gives the similarity score between psilocybin and psilocin at 76 % and between psilocybin and 3-(indol-3-yl)propyl phosphate at 75 %.

Different databases might give different pairwise scorings depending on the underlying features used in the comparison. A human chemist may also give different judgments about the relative similarity of two pairs of compounds depending on the application they have in mind. However, it is clearly not possible to algorithmically make a significant distinction based on a 1 % difference in a similarity metric. While 3D shape-based similarity measures may yield better results in some cases, many organic molecules are flexible and thus may adopt multiple conformers, hindering computation of shape-based similarities. In terms of a semantic similarity

using the ontology relationships captured in ChEBI, however, we can immediately observe rather strong differences between these two pairs. While psilocybin and psilocin are both classified as tryptamine alkaloids (CHEBI:48274) and as having the role hallucinogen (CHEBI:34599), 3-(indol-3-yl)propyl phosphate has immediate structural parent monoalkyl phosphate (CHEBI:25381) and no role annotated. In terms that have strong biological relevance, then, the first pair is much closer in similarity than the second pair. This is not reflected in the structural similarity measure, but would be reflected in a semantic similarity measure using the relationships asserted in the ChEBI ontology. A hybrid similarity metric for chemical entities, called Chym, has been developed, which combines a structure-based similarity measure with a semantic similarity measure based on ChEBI (Ferreira and Couto 2010). In application of Chym to several classification problems, the authors have shown that their hybrid measure yields better (more biologically meaningful) results than a purely structural similarity measure. Recently, the enhanced logical expressivity (disjointness axioms) within ChEBI has been harnessed to implement an improved semantic similarity metric for chemicals (Ferreira et al. 2013).

CMPSim (Grego et al. 2010) is another web tool that also uses the information contained in the ChEBI ontology in order to calculate similarity. In this case, the tool is used to quantify similarity between metabolic pathways. This tool can be used to quickly find the semantic similarity between KEGG pathways. To do so, a pathway is mapped into the ChEBI compounds that participate in it, and then the two pathways are compared by comparing the ChEBI similarities. Within systems biology in general, ChEBI is extensively used for model and pathway annotations (Courtot et al. 2011), and the ontology relationships have furthermore been used to help automate gap filling in the process of the creation of whole-genome metabolic models (Swainston et al. 2011).

Enrichment Analysis

Data that are annotated with ontologies are organized by the structure of the ontologies into groupings which can serve as the categories around which several statistical data mining techniques can operate. One technique that makes use of the structure of ontologies in that way is *ontology-based enrichment analysis*. In an ontology enrichment analysis, an annotated dataset is compared to a background set to find whether some ontology categories are statistically overrepresented in the annotated data as compared to the background. This technique has found widespread use in functional genomics, in which sets of genes annotated in the Gene Ontology are compared to the full set of GO annotations to determine which functional ontology categories are over- or underrepresented in the experimental sample. There are many variations on this theme in different implementations. For example, the genes of interest may be selected as those overexpressed in a particular microarray experiment, and the background reference set may be the full set of genes present in the microarray experiment or may be the full set of genes annotated for a given species (Huang et al. 2009).

It is possible to perform ChEBI role enrichment for metabolic data with MBRole (Chagoyen and Pazos 2011). MBRole performs enrichment analysis from ChEBI role annotations which are grouped into biological roles, chemical roles, and applications. MBRole is also able to perform pathway enrichment analysis using KEGG pathways. Data mining approaches have also been used to establish links between drugs and pathways (Hoehndorf et al. 2012). In this utility, a large-scale mapping between disease pathways and chemicals that may be used to perturb them is provided through the integration of information about drugs, genes, diseases, and pathways. The approach uses a multi-ontology enrichment analysis with a human disease ontology and a chemical ontology in combination.

To aid in the analysis of chemicals in large-scale metabolomics or other -omics measurements, ChEBI has been used to implement a web-based enrichment analysis tool, BiNChE (Moreno et al. 2015). The tool provides the option of doing enrichment analysis with the ChEBI role branch, structure branch, or both. It allows three main types of enrichment analysis: weighted, plain, and fragment. The plain analysis takes as input a list of ChEBI identifiers and uses a binomial test to determine ontology class enrichment. For the weighted analysis, the tool takes as input a list of ChEBI identifiers together with an accompanying set of numbers from 0 to 1, which can represent intensities or any measurement that reflects differential relevance of the small molecules in an experiment. This type of enrichment uses an implementation of the SaddleSum algorithm to compute differential enrichment based on not only the input ChEBI entries but also their weights. Fragment analysis is a special case of weighted analysis, in which only the chemical structure branch of the ontology is used; the input data for this type of analysis would typically come from fragmentation mass spectrometry experiments. In all cases, results are corrected for multiple hypothesis testing. The results of the analysis are displayed in an interactive graph that can be exported as a high-resolution image or in network formats. An example is illustrated in Fig. 3.

Cheminformatics and the Semantic Web

The Semantic Web is a worldwide technological movement which aims to bring data online in readily interpretable, semantically annotated, integrated, and interlinked form. Underlying this effort are a series of technologies including a standard format for representing data as triples (subject-predicate-object, RDF), a standard web-enabled identifier formats (URIs), and the OWL ontology technology for embedded semantics. In this context, chemical and cheminformatics data are becoming available in large interlinked repositories. For example, the PubChem database has recently been made available as linked data in the RDF format (Fu et al. 2015), as was the ChEMBL database (Willighagen et al. 2013) and various other EBI databases (Jupp et al. 2013). Aggregate stores of linked data on the semantic web include Bio2RDF (Belleau et al. 2008) and Chem2Bio2RDF (Chen et al. 2010). Publishing data in this fashion allows for remote usage, without the additional burden of heterogeneous data integration across different database models. Semantic

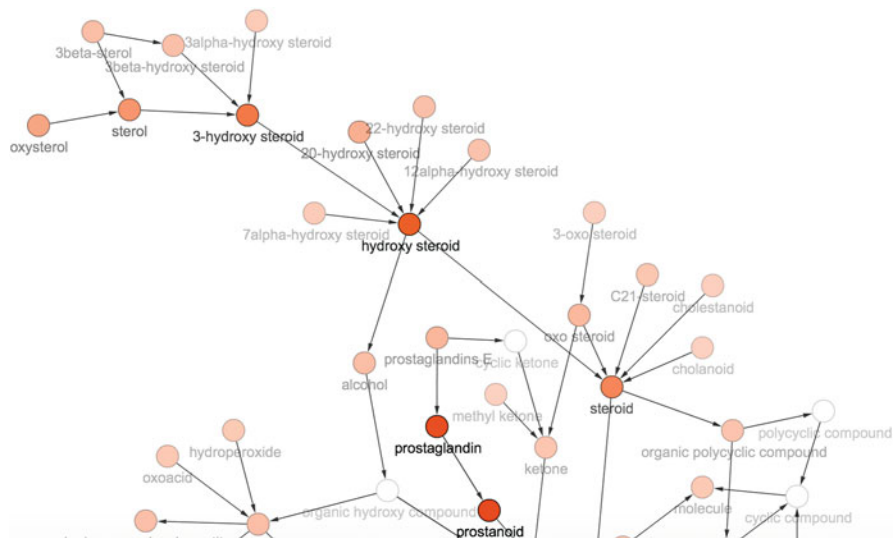


Fig. 3 A portion of a sample output graph of a BiNChE plain enrichment analysis in the structure classification branch of the ontology

web-enabled software is able to fetch the desired data from distributed repositories that support cross-resource query answering over multiple sources.

Annotation of the data using standard vocabularies is key to the success of semantic integration. This is because different vocabularies or data models may refer to the same data types with different labels or identifiers. For example, in cheminformatics, molecular weight as a property of a molecule may be referred to in one database as “MWEIGHT,” while in another it is referred to as simply “WEIGHT” or even “MASS” and in another as “MOLWEIGHT” and so on. These different labels obscure the fact that the data is comparable and should be integrated, thus “hiding” portions of such data from algorithmic processes or forcing a tedious manual disambiguation step. On the other hand, multiple implementations of an algorithm may use the same terminology while nevertheless producing different outputs due to differing heuristics, optimizations, errors, or differences in the interpretation of relevant concepts. This can lead to incorrect deductions when the results of calculations are made available under the same terminological label without further provenance as to which implementation was used to generate the data. A key challenge that arises in this context is thus managing all aspects of the provenance of information published online.

To standardize the annotation of calculated or measured chemical data in the context of the semantic web, the chemical information (CHEMINF) ontology was developed (Hastings et al. 2011). Both measurement and prediction of property values are ways to derive information about chemical or biological properties and represent them so that they can be accessible for further research. Openly published on the semantic web, such values can be reused for diverse subsequent

analyses. It is important, however, to fully detail the algorithms which were used to generate calculated property values and/or the experimental conditions under which the data were generated. To address this need, the CHEMINF ontology includes entities representing chemical graphs and the various formats for encoding them; descriptors, with definitions and axioms describing what they are specifically about; algorithms and their software implementations, with axioms describing their inputs and outputs; and chemical qualities, which are needed to specify exactly which quality a chemical descriptor is describing.

CHEMINF was one of the ontologies used in the semantic annotation of the large-scale PubChem database in RDF (Fu et al. 2015). As a deposition-based aggregation archive database, PubChem calculates and represents many different types of chemical property, identifier, and other attributes, organized around a standardized chemical structure. For example, PubChem calculates 3D coordinates, physical properties (e.g., molecular weight, XLogP3), and descriptors (e.g., InChI, SMILES, IUPAC names). Furthermore, all such information that is provided by depositors from their own data collections are faithfully represented in PubChem. There is thus a pertinent need for the standard representation of, firstly, the type for each data item, taken from the CHEMINF descriptor classification; secondly, the algorithm and toolkit used (where applicable) to calculate the descriptor in question, taken from the CHEMINF software and algorithm classification; and, finally, the source of the data (where not internally calculated), taken from a CHEMINF vocabulary of database names. All entries in CHEMINF have dereferenceable URIs, that is, their identifiers are themselves linked data ready.

Other ontologies used in the PubChem RDF implementation include ChEBI for chemical classification, the Units Ontology (UO) (Gkoutos et al. 2012) for the accurate representation of standard units of measure, and the Citation Typing Ontology (CiTO) (Shotton 2010) for the representation of associated references from the literature.

Limitations and Research Areas

There is a fundamental challenge for the comprehensive integration of chemical structure-based algorithms with logic-based OWL reasoning, in that the OWL language is not able to represent the full detail of cyclic structures, including therefore the actual structures of molecular entities containing rings (Magka et al. 2011). Given this limitation, progress has been made to partially represent aspects of chemical structure and expose those to OWL-based reasoning. One of the first applications of OWL for chemical classification was Villanueva-Rosales and Dumontier (2007), in which functional groups were encoded within an OWL ontology used for group-based automated ontology classification. Care, however, had to be taken to only encode the detail of the functional group structures insofar as they were not cyclic. The same team extended this work to include axiomatic definitions of lipid classes in the Lipid Ontology (Chepelev et al. 2011). In both of these cases, the detection of specific functional groups in molecular structures

is done using normal algorithmic approaches (Wegner et al. 2012). A prototype of a self-classifying ontology for chemicals in ChEBI was presented in Chepelev et al. (2012), which includes an algorithm for the discovery of shared features among groups of chemical structures, the representation of such features in an OWL ontology, and the automatic classification of that ontology using OWL reasoning. The features which are detected include common functional groups and additionally the presence of charges and cycles.

Different types of structural features used in chemical class definitions were identified in Hastings et al. (2012). These include:

- Presence or absence of specific parts. In the general case for chemical structures which may include cycles, in order to classify molecules based on the detail of the parts of their structures, a different formalism to OWL is needed for the chemical ontology. Description graphs have been proposed as one such formalism (Magka et al. 2011).
- The number or count of specific parts. While ChEBI does not capture this information at present, it is within the scope of the OWL formalism through capturing *cardinality* constraints on the “has part” relationships. For example, a cardinality constraint could capture the fact that a tricarboxylic acid is a class of chemical entities that has three carboxy groups as part. The algorithms reported in Chepelev et al. (2012) were capable of discovering such constraints on classes. However, within that approach, it is not possible to exclude the inference that a tricarboxylic acid – with three carboxy groups – is also a dicarboxylic acid, with *at least two* carboxy groups. OWL is capable of representing this constraint accurately, but the performance of the reasoning over a knowledge base including many such constraints is still an open challenge.
- Calculated properties of the chemical such as overall charge or molecular weight. OWL is capable of handling data properties such as integers or strings. These allow the definition of classes referring to particular values or value ranges. For example, it is possible to define *small* molecules as the molecules whose molecular weight is less than 800 daltons. These data restrictions are supported by OWL reasoners just as are object property restrictions.
- Topological features, such as polycyclic cages or molecular knots. Topological features such as overall regularity or cage structure falls outside the expressivity of OWL, although some previous work has investigated the use of higher-order logics for that purpose (Kutz et al. 2012).
- Structural formulae, e.g., hydrocarbons, strictly defined, excluding heteroatoms. Due to the open-world semantics of OWL, everything that is not explicitly stated in the ontology is assumed to be *not known to hold* rather than *known not to hold*. For example, ChEBI contains “organic molecule” and “inorganic molecule” as two classes. However, if it is not explicitly stated in the ontology that *all* molecules are either organic or inorganic and none are both, the ontology cannot infer that a molecule is inorganic simply from a statement to the effect that it is not organic. This open-world property of the semantics is a challenge for strict and exclusive class definitions. This is an area where the closed-world semantics

of alternative formalisms such as the description graphs formalism described in Magka et al. (2011) may be more practical. A particularly challenging class to define within a logic-based formalism is one captured by a parameterized molecular formula, such as alkenes which are described by the formula C_nH_{2n} . Constraints on number of atoms of particular sorts can be expressed using OWL cardinality restrictions, but this facility does not allow the relationship between the number of carbons and the number of hydrogens to be expressed.

Another open research area in the context of text and data mining is the prediction of novel ontology annotations for as yet unannotated data. The Gene Ontology annotation team use a set of rules for automatically predicting ontology annotations for as yet unannotated data, but these are mainly based on knowledge about the orthology of genes rather than data mining existing annotations or examining the structure of the ontology (Camon et al. 2004). Ontologies provide categories for data to be aggregated into which categories can then form the input into machine learning algorithms and statistical models of various sorts that can in turn provide predictions for novel data. Supervised methods, such as Bayesian classifiers, decision trees, and support vector machines, can be used to classify compounds for a particular functional activity class. However, these approaches result in binary output for a particular class membership rather than allocation of compounds to the ontology dataset. Supervised machine learning for prediction of chemical class membership based on an existing hierarchy would require large training sets of chemicals that are already classified into such a hierarchy. Although ChEBI could in principle act as such a training set, the size of the classified data is still a tiny fraction of the enormous chemical space, and the problem is further exacerbated by the fact that many of the ChEBI leaf nodes thus far contain a fairly low number of representative classified structures. Manually constructed classifications may furthermore be far from complete in the sense that an arbitrary compound belongs to a vast number of classes yet will only have been classified under those deemed to be the most *relevant* – a judgment which remains, for the foreseeable future, highly subjective.

Conclusions

We have given some background to ontologies as they are being used in cheminformatics to support standardization, knowledge discovery, and data annotation. In general, the use of ontologies for standardized annotations reduces the overhead of data integration, sharing, and reuse.

An exciting direction that life sciences ontology development is currently expanding into is that of increasing interrelationships between the ontologies themselves. For example, relationships link from the Gene Ontology to the ChEBI ontology are being created to explicitly represent chemical participation in biological processes (Hill et al. 2013). This shows potential for enabling whole-systems modeling that is needed to enable truly translational research in which structural chemistry informs biological research and discovery in real-world applications.

Acknowledgements ChEBI is supported by the BBSRC under grant agreement number BB/K019783/1 within the “Bioinformatics and biological resources” fund.

Bibliography

- Baader, F., Calvanese, D., McGuinness, D., Nardi, D., & Patel-Schneider, P. (2003). *Description logic handbook* (2nd ed.). Cambridge: Cambridge University Press.
- Batchelor, C., Hastings, J., Steinbeck, C. (2010). Ontological dependence, dispositions and institutional reality in chemistry. In A. Galton & R. Mizoguchi (Eds.), *Proceedings of the 6th Formal Ontology in Information Systems Conference*, Toronto.
- Belleau, F., Nolin, M., Tourigny, N., Rigault, P., & Morissette, J. (2008). Bio2RDF: Towards a mashup to build bioinformatics knowledge systems. *Journal of Biomedical Informatics*, *41*, 706–716. doi:10.1016/j.jbi.2008.03.004.
- Bolton, E. E., Wang, Y., Thiessen, P. A., & Bryant, S. H. (2008). *PubChem: Integrated platform of small molecules and biological activities* (pp. 217–241). American Chemical Society, Washington, DC.
- Camon, E., Magrane, M., Barrell, D., Lee, V., Dimmer, E., Maslen, J., Binns, D., Harte, N., Lopez, R., & Apweiler, R. (2004). The Gene Ontology Annotation (GOA) database: Sharing knowledge in uniprot with gene ontology. *Nucleic Acids Research*, *32*(suppl 1), D262–D266. doi:10.1093/nar/gkh021. http://nar.oxfordjournals.org/content/32/suppl_1/D262.abstract.
- Chagoyen, M., & Pazos, F. (2011). MBRole: Enrichment analysis of metabolomic data. *Bioinformatics*, *27*(5), 730–731. doi:10.1093/bioinformatics/btr001. <http://bioinformatics.oxfordjournals.org/content/27/5/730.abstract>.
- Chen, B., Dong, X., Jiao, D., Wang, H., Zhu, Q., Ding, Y., & Wild, D. (2010). Chem2Bio2RDF: A semantic framework for linking and data mining chemogenomic and systems chemical biology data. *BMC Bioinformatics*, *11*(1), 255. doi:10.1186/1471-2105-11-255. <http://www.biomedcentral.com/1471-2105/11/255>.
- Chepelev, L., Riazanov, A., Kouznetsov, A., Low, H. S., Dumontier, M., & Baker, C. (2011). Prototype semantic infrastructure for automated small molecule classification and annotation in lipidomics. *BMC Bioinformatics*, *12*(1), 303. <http://dx.doi.org/10.1186/1471-2105-12-303>.
- Chepelev, L. L., Hastings, J., Ennis, M., Steinbeck, C., & Dumontier, M. (2012). Self-organizing ontology of biochemically relevant small molecules. *BMC Bioinformatics*, *13*, 3.
- Corbett, P., & Murray-Rust, P. (2006). High-throughput identification of chemistry in life science texts. In M. Berthold, R. Glen, & I. Fischer (Eds.), *Computational life sciences II* (pp. 107–118). Springer, Berlin/Heidelberg.
- Courtot, M., Juty, N., Knüpfer, C., Waltemath, D., Zhukova, A., Dräger, A., Dumontier, M., Finney, A., Golebiewski, M., Hastings, J., Hoops, S., Keating, S., Kell, D. B., Kerrien, S., Lawson, J., Lister, A., Lu, J., Machne, R., Mendes, P., Pocock, M., Rodriguez, N., Villegier, A., Wilkinson, D. J., Wimalaratne, S., Laibe, C., Hucka, M., & Novère, N. L. (2011). Controlled vocabularies and semantics in systems biology. *Molecular Systems Biology*, *7*, 543.
- Ferreira, J. D., & Couto, F. M. (2010). Semantic similarity for automatic classification of chemical compounds. *PLoS Computational Biology*, *6*(9), e1000937. doi:10.1371/journal.pcbi.1000937.
- Ferreira, J. D., Hastings, J., & Couto, F. M. (2013). Exploiting disjointness axioms to improve semantic similarity measures. *Bioinformatics*, *29*, 2781–2787.
- Fu, G., Batchelor, C., Dumontier, M., Hastings, J., Willighagen, E., & Bolton, E. (2015). PubChemRDF: Towards the semantic annotation of pubchem compound and substance databases. *Journal of Cheminformatics*, *7*, 34.
- Gkoutos, G. V., Schofield, P. N., & Hoehndorf, R. (2012). The units ontology: A tool for integrating units of measurement in science. *Database*, *2012*. doi:10.1093/database/bas033. <http://database.oxfordjournals.org/content/2012/bas033.abstract>.

- Grau, B. C., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P., & Sattler, U. (2008). OWL 2: The next step for OWL. *Web Semantics*, 6, 309–322. doi:10.1016/j.websem.2008.05.001. <http://portal.acm.org/citation.cfm?id=1464505.1464604>.
- Grego, T., Ferreira, J. D., Pesquita, C., Bastos, H., Vicoso, D. V., Freire, J., & Couto, F. M. (2010). Chemical and metabolic pathway semantic similarity. Technical report, LASIGE, Faculty of Sciences, University of Lisbon.
- Gruber, T. R. (2009). Ontology. In L. Liu & M. T. Özsu (Eds.), *Encyclopedia of database systems*. Springer. <http://tomgruber.org/writing/ontology-definition-2007.htm>.
- Harland, L., Larminie, C., Sansone, S. A., Popa, S., Marshall, M. S., Braxenthaler, M., Cantor, M., Filsell, W., Forster, M. J., Huang, E., Matern, A., Musen, M., Saric, J., Slater, T., Wilson, J., Lynch, N., Wise, J., & Dix, I. (2011). Empowering industrial research with shared biomedical vocabularies. *Drug Discovery Today*, 16(21–22), 940–947. doi:10.1016/j.drudis.2011.09.013. <http://www.sciencedirect.com/science/article/pii/S1359644611003035>.
- Hastings, J., Chepelev, L., Willighagen, E., Adams, N., Steinbeck, C., Dumontier, M. (2011). The chemical information ontology: Provenance and disambiguation for chemical data on the biological semantic web. *PLoS One*, 6(10), e25513. doi:10.1371/journal.pone.0025513.
- Hastings, J., de Matos, P., Dekker, A., Ennis, M., Harsha, B., Kale, N., Muthukrishnan, V., Owen, G., Turner, S., Williams, M., & Steinbeck, C. (2013). The ChEBI reference database and ontology for biologically relevant chemistry: Enhancements for 2013. *Nucleic Acids Research*, 41(Database issue), D456–D463.
- Hastings, J., Magka, D., Batchelor, C., Duan, L., Stevens, R., Ennis, M., & Steinbeck, C. (2012). Structure-based classification and ontology in chemistry. *Journal of Cheminformatics*, 4(1), 8. doi:10.1186/1758-2946-4-8. <http://www.jcheminf.com/content/4/1/8>.
- Haug, K., Salek, R. M., Conesa, P., Hastings, J., de Matos, P., Rijnbeek, M., Mahendrakar, T., Williams, M., Neumann, S., Rocca-Serra, P., Maguire, E., Gonzalez-Beltran, A., Sansone, S. A., Griffin, J. L., & Steinbeck, C. (2012). Metabolights—an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Research*. doi:10.1093/nar/gks1004. <http://nar.oxfordjournals.org/content/early/2012/10/28/nar.gks1004.abstract>.
- Hill, D. P., Adams, N., Bada, M., Batchelor, C., Berardini, T. Z., Dietze, H., Drabkin, H. J., Ennis, M., Foulger, R. E., Harris, M. A., Hastings, J., Kale, N. S., de Matos, P., Mungall, C. J., Owen, G., Roncaglia, P., Steinbeck, C., Turner, S., & Lomax, J. (2013). Dovetailing biology and chemistry: Integrating the Gene Ontology with the ChEBI chemical ontology. *BMC Genomics*, 14, 513.
- Hoehndorf, R., Dumontier, M., & Gkoutos, G. V. (2012). Identifying aberrant pathways through integrated analysis of knowledge in pharmacogenomics. *Bioinformatics*, 28(16), 2169–2175. doi:10.1093/bioinformatics/bts350. <http://bioinformatics.oxfordjournals.org/content/28/16/2169.abstract>.
- Hoehndorf, R., Oellrich, A., Dumontier, M., Kelso, J., Rebholz-Schuhmann, D., & Herre, H. (2010). Relations as patterns: Bridging the gap between obo and owl. *BMC Bioinformatics*, 11(1), 441. doi:10.1186/1471-2105-11-441. <http://www.biomedcentral.com/1471-2105/11/441>.
- Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, 37(1), 1–13.
- Hunter, L. (2002). Ontologies for programs, not people. *Genome Biology*, 3, 1002.1–1002.2.
- Jessop, D. M., Adams, S. E., Willighagen, E. L., Hawizy, L., & Murray-Rust, P. (2011). Oscar4: A flexible architecture for chemical text-mining. *Journal of Cheminformatics*, 3, 41.
- Jupp, S., Malone, J., Bolleman, J., Brandizi, M., Davies, M., Garcia, L., Gaulton, A., Gehant, S., Laibe, C., Redaschi, N., Wimalaratne, S. M., Martin, M., Novère, N. L., Parkinson, H., Birney, E., & Jenkinson, A. M. (2013). The EBI RDF platform: Linked open data for the life sciences. *Bioinformatics*, 30, 1338–1339.

- Kanehisa, M., Goto, S., Hattori, M., Aoki-Kinoshita, K., Itoh, M., Kawashima, S., Katayama, T., Araki, M., & Hirakawa, M. (2006). From genomics to chemical genomics: New developments in KEGG. *Nucleic Acids Research*, *34*, D354–D357. doi:10.1093/nar/gkj102.
- Kutz, O., Hastings, J., & Mossakowski, T. (2012). Modelling highly symmetrical molecules: Linking ontologies and graphs artificial intelligence: Methodology, systems, and applications. In A. Ramsay & G. Agre (Eds.), *Artificial intelligence: Methodology, systems, and applications* (Lecture notes in computer science, Vol. 7557, chap. 11, pp. 103–111). Springer, Berlin/Heidelberg. doi:10.1007/978-3-642-33185-5_11. http://dx.doi.org/10.1007/978-3-642-33185-5_11.
- Li, C., Donizelli, M., Rodriguez, N., Dharuri, H., Endler, L., Chelliah, V., Li, L., He, E., Henry, A., Stefan, M., Snoep, J., Hucka, M., Le Novère, N., & Laibe, C. (2010). BioModels database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, *4*, 92.
- Lowe, D. M., Corbett, P. T., Murray-Rust, P., & Glen, R. C. (2011). Chemical name to structure: Opsin, an open source solution. *Journal of Chemical Information and Modeling*, *51*(3), 739–753. doi:10.1021/ci100384d. <http://pubs.acs.org/doi/abs/10.1021/ci100384d>.
- Magka, D., Motik, B., & Horrocks, I. (2011). Modelling structured domains using description graphs and logic programming. Technical report, Department of Computer Science, University of Oxford.
- Matthews, L., Gopinath, G., Gillespie, M., Caudy, M., Croft, D., de Bono, B., Garapati, P., Hemish, J., Hermjakob, H., Jassal, B., Kanapin, A., Lewis, S., Mahajan, S., May, B., Schmidt, E., Vastrik, I., Wu, G., Birney, E., Stein, L., & D'Eustachio, E. (2009). Reactome knowledgebase of human biological pathways and processes. *Nucleic Acids Research*, *37*, D619–D622.
- McNaught, A. D., & Wilkinson, A. (1997). IUPAC compendium of chemical terminology (2nd ed., the “Gold Book”). Oxford: Blackwell Scientific Publications. doi:doi:10.1351/goldbook. [XMLon-linecorrectedversion:http://goldbook.iupac.org](http://goldbook.iupac.org). (2006-) created by M. Nic, J. Jirat, B. Kosata; updates compiled by A. Jenkins.
- Moreno, P., Beiskens, S., Harsha, B., Muthukrishnan, V., Tudose, I., Dekker, A., Dornfeldt, S., Taruttis, F., Grosse, I., Hastings, J., Neumann, S., & Steinbeck, C. (2015). BiNChE: A web tool and library for chemical enrichment analysis based on the chEBI ontology. *BMC Bioinformatics*, *16*, 56.
- Shearer, R., Motik, B., & Horrocks, I. (2008). HermiT: A highly-efficient OWL reasoner. In C. Dolbear, A. Ruttenberg, & U. Sattler (Eds.), *Proceedings of the 5th Workshop on OWL: Experiences and Directions*, Karlsruhe.
- Shotton, D. (2010). CiTO, the citation typing ontology, and its use for annotation of reference lists and visualization of citation networks. *Journal of Biomedical Semantics*, *1*(Suppl 1), S6.
- Sirin, E., Parsia, B., Cuenca Grau, B., Kalyanpur, A., & Katz, Y. (2007). Pellet: A practical OWL-DL reasoner. *Journal of Web Semantics*, *5*, 51–53.
- Smith, B. (2003). Ontology. In L. Floridi (Ed.), *Blackwell guide to the philosophy of computing and information* (pp. 155–166). Oxford: Blackwell.
- Swainston, N., Smallbone, K., Mendes, P., Kell, D. B., & Paton, N. W. (2011). The SuBliMinaL Toolbox: Automating steps in the reconstruction of metabolic networks. *Journal of Integrative Bioinformatics*, *8*, 186.
- Protégé Team, T. (2013). The Protégé ontology editing tool. <http://protege.stanford.edu/>. Last accessed Mar 2013.
- The Gene Ontology Consortium. (2000). Gene ontology: Tool for the unification of biology. *Nature Genetics*, *25*, 25–29.
- The Gene Ontology Consortium. (2012). The OBO language, version 1.2. http://www.geneontology.org/GO.format.obo-1_2.shtml. Last accessed Oct 2012.
- The UniProt Consortium. (2015). UniProt: A hub for protein information. *Nucleic Acids Research*, *43*, D204–D212.
- Tsarkov, D., & Horrocks, I. (2006). FaCT++ description logic reasoner: System description. In *Proceedings of the International Joint Conference on Automated Reasoning (IJCAR 2006)*, Seattle (pp. 292–297). Springer.

- Villanueva-Rosales, N., & Dumontier, M. (2007). Describing chemical functional groups in OWL-DL for the classification of chemical compounds. In *Proceedings of the OWL: Experiences and Directions (OWLED 2007)*, Innsbruck.
- Wegner, J. K., Sterling, A., Guha, R., Bender, A., Faulon, J. L., Hastings, J., O'Boyle, N., Overington, J., Van Vlijmen, H., & Willighagen, E. (2012). Cheminformatics. *Communications of the ACM*, 55(11), 65–75.
- Willighagen, E. L., Waagmeester, A., Spjuth, O., Ansell, P., Williams, A. J., Tkachenko, V., Hastings, J., Chen, B., & Wild, D. J. (2013). The ChEMBL database as linked open data. *Journal of Cheminformatics*, 5, 23.
- Wishart, D., Knox, C., Guo, A., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z., & Woolsey, J. (2006). DrugBank: A comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Research*, 34, D668–D672. doi:10.1093/nar/gkj067.
- Wishart, D. S., Knox, C., Guo, A. C. C., Eisner, R., Young, N., Gautam, B., Hau, D. D., Psychogios, N., Dong, E., Bouatra, S., Mandal, R., Sinelnikov, I., Xia, J., Jia, L., Cruz, J. A., Lim, E., Sobsey, C. A., Shrivastava, S., Huang, P., Liu, P., Fang, L., Peng, J., Fradette, R., Cheng, D., Tzur, D., Clements, M., Lewis, A., De Souza, A., Zuniga, A., Dawe, M., Xiong, Y., Clive, D., Greiner, R., Nazyrova, A., Shaykhtudinov, R., Li, L., Vogel, H. J., Forsythe, I. (2009). HMDB: A knowledgebase for the human metabolome. *Nucleic Acids Research*, 37(Database issue), D603–D610. doi:10.1093/nar/gkn810. <http://dx.doi.org/10.1093/nar/gkn810>.

Adam Liwo, Cezary Czaplewski, Stanisław Ołdziej,
Bartłomiej Zaborowski, Dawid Jagieła, and Jooyoung Lee

Contents

Introduction	2184
Force-Field Calibration for Fold Recognition	2186
Energy Gap and Z-Score Optimization for Ab Initio Protein Structure Prediction and Folding Simulations	2187
Funnel Sculpting	2189
Maximum Likelihood Optimization	2193
Conclusions and Outlook	2195
Bibliography	2196

A. Liwo (✉)

Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

Center for In Silico Protein Structure and School of Computational Sciences, Korea Institute for
Advanced Study, Seoul, Republic of Korea

e-mail: adam@sun1.chem.univ.gda.pl; adam.liwo@ug.edu.pl

C. Czaplewski • B. Zaborowski • D. Jagieła

Faculty of Chemistry, University of Gdańsk, Gdańsk, Poland

e-mail: cezary.czaplewski@ug.edu.pl; lightnir@gmail.com

J. Lee

Center for In Silico Protein Structure and School of Computational Sciences, Korea Institute for
Advanced Study, Seoul, Republic of Korea

e-mail: jlee@kias.re.kr

S. Ołdziej

Laboratory of Biopolymer Structure, Intercollegiate Faculty of Biotechnology, University of
Gdańsk and Medical University of Gdańsk, Gdańsk, Poland

Abstract

Force fields are empirical energy functions that are designed to compute the properties of molecular systems at a relatively low cost, subject to the condition that there are no chemical changes in a system. A force field is a sum of various energy contributions which were developed and parameterized using small model systems. A long-sought goal in designing a force field is that the simulation of a biomolecular system (mainly globular proteins) reproduces the native structure(s) and the folding/assembly process of the system. A systematic way to achieve this goal is to train a force field using a number of macromolecules (e.g., proteins) to determine the balance among constituting energy terms by adjusting force-field parameters so that the force field properly reproduces the native structures of the training molecules. In this chapter, we discuss the principles and methods used for the calibration of a force field suitable for various applications.

Introduction

Empirical force fields to treat biological macromolecules, including proteins, have a long history of development (Mackerell 2004). Initially, they were defined as sums of neoclassical expressions that include bond stretch, bond angle, torsional angle, point-charge electrostatics, and van der Waals terms. Parameters for bond lengths, bond angles, torsional angles, and van der Waals terms were determined based on various spectroscopic experimental data (crystal, infrared, NMR, and microwave spectroscopy data) of small molecules. Point charges were determined by fitting to the molecular electrostatic potentials of model molecules calculated by using molecular quantum mechanics methods. These parameters were subsequently adjusted to reproduce the geometry, infrared frequencies, heats of associations, etc., of small molecular systems (Némethy et al. 1992; Pearlman et al. 1995; Brooks et al. 2009). Thus-generated force fields were reasonably good to study the dynamics of proteins near their experimental structures, but the computational resources and simulation techniques available at that time did not allow one to use them to study protein folding from unfolded structures. Therefore, the capacity of these early force fields to fold proteins could not be properly assessed. Nevertheless, some of them, e.g., the ECEPP force field developed by Scheraga and coworkers, were successfully used to predict structures of rigid cyclic peptides such as gramicidin S (Dybert et al. 1975) ahead of experiment.

In late 1970s and early 1980s, coarse-grained force fields for proteins began to emerge. Levitt and Warshel (Levitt and Warshell 1975; Levitt 1976) developed a physics-based force field for proteins, representing the protein backbone by three pseudo-atoms (identified with C^α , N, and O) and the side chain by one extended atom. The interaction potentials were defined and determined as Boltzmann-averaged energies over approximate energy surfaces of a model system calculated with all-atom empirical potentials and partially by taking empirical hydration

energies of amino acid residues. A few years later, an alternative approach to develop coarse-grained force fields was proposed by Crippen and coworkers, which was based on the statistical potentials determined from a database of known protein structures (Obatake and Crippen 1981; Crippen and Wiswanadhan 1984). When these researchers applied then-available global optimization methods to explore the energy surface of protein sequences with known 3D structures, their force fields produced protein-like structures, but not native-like structures.

In the early 1980s, a protein structure prediction approach known as threading or fold recognition was proposed. In this approach, the native structure of a protein is sought among decoys generated from known protein structures (Bryant and Lawrence 1993; Miller et al. 1996). This approach was developed for both on-lattice and off-lattice representation of protein conformations. The decoy which gives the lowest energy (or equivalently the best score) is considered as the prediction of the native fold. For the success of this approach, the corresponding energy/score function was trained on a large number of known protein structures. The training was achieved usually by optimizing the parameters so that the most native-like structure corresponds to the lowest Z-score one. The Z-score is a quantity taken from statistics and is defined as the ratio of the difference between the energy of a structure and the average energy of sampled structures to the standard deviation of the sample.

The first method to determine a force field with folding properties (i.e., operating in the conformational search space not restricted to a fixed set of decoys) was developed by Crippen and Snow (Crippen and Snow 1990). The method was based on the maximization of the energy gap between the lowest-energy simulated native-like structure and the lowest-energy simulated nonnative structure. The physics-based justification of such an approach was provided by Shakhnovich and coworkers (1991) and Sali et al. (1994) who, based on their lattice model studies of proteins, concluded that the larger the energy gap is, the quicker the folding happens. Even though the original Crippen's results did not survive long because it turned out very quickly that their energy-embedding method used to generate the decoys was insufficient to find the global energy minimum (Snow 1992) and it was found later that a large energy gap is neither a necessary nor a sufficient condition for foldability (Camacho and Thirumalai 1996; Klimov and Thirumalai 1996), the energy-gap maximization still remains as one of the principles of force-field training. It was later demonstrated based on on-lattice model studies (Camacho and Thirumalai 1996; Klimov and Thirumalai 1996) that a large energy gap is insufficient to assure folding and that folding transition should precede collapse of a polymer chain into compact but unorganized structures. A few years later (Liwo et al. 2004), this result was generalized to the requirement that for a foldable energy landscape, the formation of native-structure elements should occur in an ordered manner, which observation provided the basis for the hierarchical method of force-field optimization (Liwo et al. 2002, 2007; Ołdziej et al. 2004a, b). Wolynes and coworkers, on the other hand, provided a physical justification of Z-score optimization by relating this quantity to the difference between the folding and glassy transitions of proteins (Brungelson and Wolynes 1987; Goldstein et al. 1992).

For a protein to fold, its folding-transition temperature should be well above its glassy-transition temperature. The ratio of the difference of these two temperatures to the glassy-transition temperature is related to Z-score. Z-score optimization also is a principle that underpins many successful force-field optimization approaches.

This chapter summarizes the force-field calibration methods, starting from the two initial concepts outlined above.

Force-Field Calibration for Fold Recognition

As outlined in Introduction, the fold-recognition problem is defined as finding the most native-like structure among a given set of decoys. The scoring function is usually based on energy or pseudo-energy. Therefore, depending on whether Z-score or energy gap is the scoring function, the force-field optimization problem can be defined as the minimization of the target function given by Eqs. 1 or 2, respectively:

$$\Delta E = E_{nat} - \min_{i \neq nat} E_i \quad (1)$$

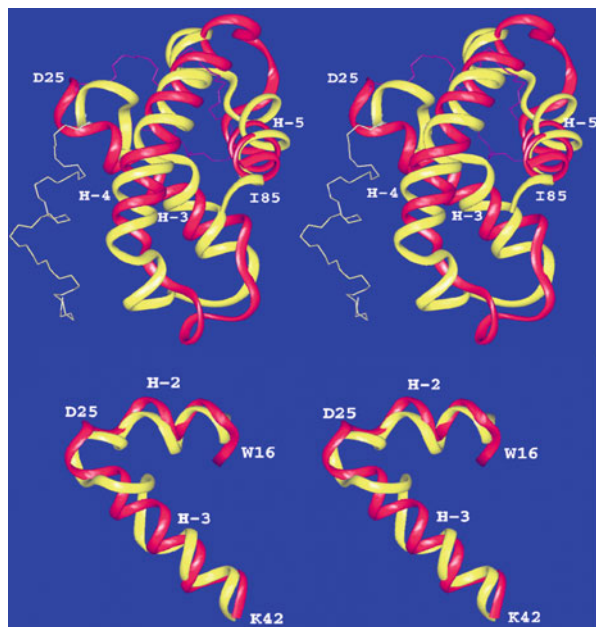
$$Z = \frac{E_{nat} - \frac{1}{N} \sum_{i \neq nat} E_i}{\sqrt{\frac{1}{N} \sum_{i \neq nat} E_i^2 - \left(\frac{1}{N} \sum_{i \neq nat} E_i \right)^2}} \quad (2)$$

where *nat* is the index of the native structure of the training protein and *N* is the number of decoys, excluding the native structure.

The energy function, *E*, is usually a linear function of parameters which makes the optimization problem a linear programming problem if the score function is the energy gap (Eq. 1) or nonlinear programming problem when the Z-score function is used (Eq. 2). Examples of successful force fields developed using this principle are the force field used in the LOPP server (Meller and Elber 2001) and that developed by Zhang and Skolnick (2004).

Optimization of a force field for fold recognition is relatively easy because the set of decoys is fixed. However, the accuracy of fold recognition depends on the completeness of the decoy set (threading database). Therefore, the method inevitably fails if the database does not include the appropriate fold. Often, it is difficult to identify the correct fold in the database if it is of less common type. A variant of threading, in which the energy of decoys is minimized with the target sequence before the score function is calculated, was developed in our laboratory and applied to the optimization of the early version of the UNRES force field (Liwo et al. 1997). This approach was subsequently further studied by Schug and Wenzel (2006). In this case, the force-field optimization requires iterations because the minimized geometry of decoys depends on force-field parameters. On the other hand, the resulting force fields are applicable not only to the folds used in training but also to new-fold proteins for de novo folding studies; a prediction made with

Fig. 1 *Top panel:* the predicted structure of the 25–85 fragment of HDEA (red ribbon) superposed on the experimental structure (PDB code 1BG8) Yang et al. (1998) (yellow ribbon); the C α rmsd is 4.2 Å. *Bottom panel:* same for the 16–42 fragment; the C α rmsd is 2.9 Å (Reproduced with permission from Figs. 1 and 2 of Liwo et al. (1999))



the UNRES force field optimized in this way coupled with the global optimization conformational space annealing (CSA) method (Lee et al. 1997, 1999) successfully generated the best protein model for target T0063 of CASP3 (Fig. 1) (Liwo et al. 1999).

Energy Gap and Z-Score Optimization for Ab Initio Protein Structure Prediction and Folding Simulations

The target functions of this optimization are the same as for force-field optimization for fold recognition; however, instead of using a fixed set of decoys, we need to generate many diverse low-energy decoys using the energy function under consideration. Because of high cost of such computations, the number of training proteins is much less than that used in optimization for fold recognition (where the set of decoys is fixed). In the very first approach developed by Crippen and Snow, the energy-embedding method was used. They also used the experimental structure to compute the native reference (Crippen and Snow 1990; Seetharamulu and Crippen 1991). However, this method failed to include many low-energy structures, and for this reason, the resulting force field worked only with the energy-embedding method. When a more powerful Monte Carlo minimization (MCM) search method (Li and Scheraga 1987) was applied with the identical force field, nonnative structures with lower energies than the native reference structure were found (Snow 1992).

In our laboratory, we developed a variant of energy-gap and Z-score optimization method (Lee et al. 2001; Pillardy et al. 2001) in which decoys are generated with the powerful CSA conformational search technique (Lee et al. 1997; Lee and Scheraga 1999). The experimental structure was not used as the native reference; instead, an approach called local CSA was developed to perturb the experimental structure in order to generate low-energy near-native structures. Therefore, the target functions given by Eqs. 1 and 2 had to be redefined by Eqs. 3 and 4, respectively.

$$\Delta E = \min_{i \in \{nat\}} E_i - \min_{i \notin \{nat\}} E_i \quad (3)$$

$$Z = \frac{\frac{1}{N_{nat}} E_i - \frac{1}{N_{nonnat}} \sum_{i \notin \{nat\}} E_i}{\sqrt{\frac{1}{N_{nonnat}} \sum_{i \notin \{nat\}} E_i^2 - \left(\frac{1}{N_{nonnat}} \sum_{i \notin \{nat\}} E_i \right)^2}} \quad (4)$$

where $\{nat\}$ denotes the set of native-like decoys (generated by restrained or unrestrained simulations with the current energy function parameters), N_{nat} is the number of native-like decoys, and N_{nonnat} is the number of nonnative-like decoys. The optimization involves iterating cycles, each of which consists of decoy generation and target function optimization, as shown in Fig. 2.

Using this approach, we managed to produce a version of the UNRES force field that folded proteins of all structural classes (Pillardy et al. 2001). However, the transferability of this force field was not well established. Later, the method was formulated as a sequence of linear programming problems and applied to the optimization of the united-residue UNRES and the all-atom CHARMM force fields (Lee et al. 2002, 2004a, b, 2008).

Wolyne and coworkers used the Z-score optimization method with the umbrella sampling Brownian dynamics as a search engine to optimize their force field based on associative-memory Hamiltonian (Eastwood et al. 2002, 2003; Fujitsuka et al. 2004). The resulting force field had some success in protein structure prediction (Prentiss et al. 2006).

When using a flexible decoy set, regardless of the target function to optimize (e.g., Eq. 3 or 4), two problems are encountered: (i) instability of the iterative process and (ii) glassiness of the resulting force field. The reason for the glassiness is that although local CSA or umbrella sampling Brownian dynamics generate reasonable low-energy near-native structures, there is no guarantee, especially with complex training proteins, that the optimization will provide accessible pathways for the proteins to fold starting from the unfolded state to the folded structures. The numerical instability problem results from the fact that the target function optimized in a given cycle can produce a force field with which qualitatively distinct decoys are generated; these new decoys often have very low energies and are usually nonnative. The reason for this is that the largest contribution to the target function comes from the native-like structure or the lowest-energy nonnative structures, while the other

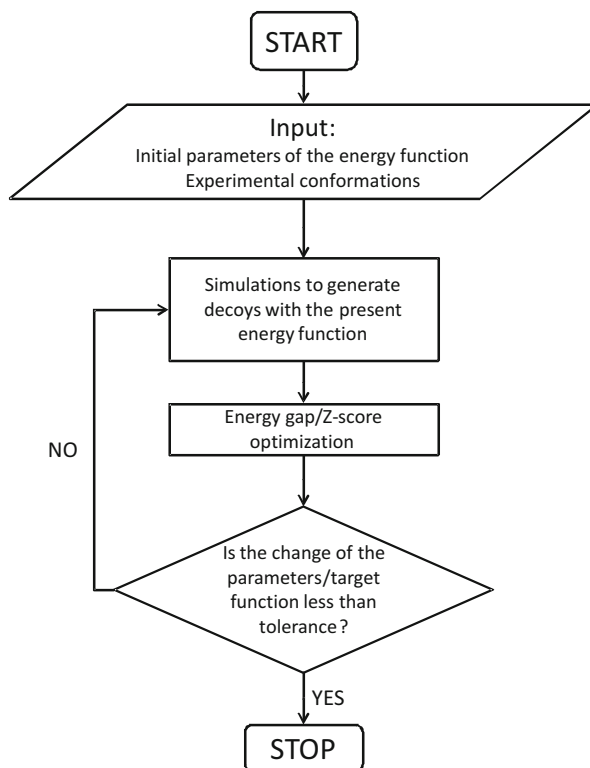


Fig. 2 A general iterative scheme of force-field optimization

structures are largely ignored. Therefore, the optimization of a force field is possible only for training proteins with simple folds (e.g., the three- α -helix bundle (Lee et al. 2001) or a three-stranded β -sheet fold (Pillardiy et al. 2001)). The development of a transferable force field that can fold proteins with more complex α -helical folds or $\alpha + \beta$ proteins still remains a challenge. Methods to consider in order to overcome the drawbacks in the energy-gap and/or Z-score optimization methods are described in the next sections.

Funnel Sculpting

Funnel sculpting is an approach developed by Micheletti and coworkers (2001) and Fain and Levitt (2003), which is based on the Wolynes' concept that the energy landscape of a foldable protein should look like a funnel; the lower the energy is, the fewer accessible conformational states should appear. As pointed out in the previous section, neither the energy-gap nor the Z-score optimization method guarantees a funnel-like energy landscape; instead, the *Schott-funnel* landscape with a flat and

rugged bottom is often obtained, resulting in the *glassiness* of the force field. The method of funnel sculpting is based on the maximization of the correlation between energy and native likeness. Although it never reached beyond the testing stage, this idea was a big step forward in the field of energy function optimization.

Hierarchical Optimization

The idea that the (free) energy should decrease with increasing native likeness was implemented in the hierarchical optimization method (Liwo et al. 2002). In this approach, the conformational space of a training protein is divided into levels, each of which consists of conformations with common native-structure element(s) (e.g., a native α -helix, a native β -hairpin, a native supersecondary structure motif, etc., or combinations thereof). The more native-like elements a level has, the higher rank it is assigned; the native structure has all native-like elements and, consequently, has the highest rank. The idea is to optimize the energy function so that the lower free energy a level has, the higher its native likeness is. The free energy of each level can be calculated from the results of simulations with a given force field. The levels and their hierarchy should follow the experimental thermal folding/unfolding pathway, as outlined in Fig. 3.

In the initial version of the hierarchical optimization approach (Liwo et al. 2002; Ołdziej et al. 2004a, b), the CSA global optimization method was implemented to generate decoys. To assign a particular conformation to a level, a binary number was assigned to each conformation, depending on the conformity of its segments (including the whole conformation) with the native structure; this number increases with increasing native likeness (Ołdziej et al. 2004b). To account for gradual folding, decoys were collected from various stages of CSA, i.e., after 8,000, 16,000, and 64,000 energy-minimized conformations, for stages 1, 2, and 3, respectively. The conformations of stage 1 usually contained only one native-structure element, and the conformations of stage 3 were all native-like. The target function consisted of quartic penalty terms to maintain the free energy gaps between levels above the preassigned values, Δ (see Eq. 5). As in the energy-gap and Z-score optimization approaches, hierarchical optimization is an iterative process that consists of decoy generation and target function minimization. The target function is defined as

$$\Phi = \sum_{\substack{\text{training} \\ \text{proteins}}} \sum_{\beta} \sum_{i < j} w_i g[F_j(\beta) - F_i(\beta); \Delta_{ij}^{(1)}, \Delta_{ij}^{(2)}(\beta)] \quad (5)$$

where

$$F_i = -\frac{1}{\beta} \ln \sum_{i \in \{i\}} \exp(\omega_i - \beta E_i) \quad (6)$$

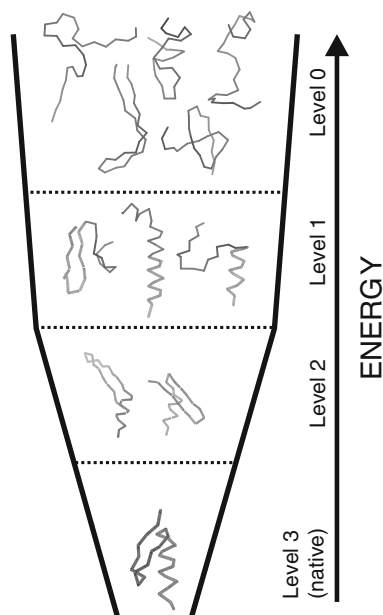


Fig. 3 Schematic illustration of the energy-ordering of the conformational-hierarchy levels, which is the goal of the algorithm for optimizing the potential function by hierarchical approach, using the 1FSD peptide (Dahiyat and Mayo 1997) as an example. The energies of the conformations should decrease with their increasing “native-likeness”. The highest energy level (*Level 0*) is occupied by structures with either no or nonnative secondary structure. Next (*Level 1*) is occupied by the structures with one native secondary structure element (the N-terminal β -hairpin or the C-terminal α -helix; the native-like structure fragments are indicated by thicker lines). Yet lower energy (*Level 2*) have structures with both α -helix and β -hairpin, but no or incorrect packing of these two substructures and/or shifted turn in the β -hairpin. Finally, the native-like structures, with α -helix and β -hairpin packed correctly occupy the lowest energy level (*Level 3*). Because the number of structures with more and more defined native-like elements decreases, such ordering of structures leads to diminishing conformational entropy following the energy decrease, which is highly desirable in order to find the native structure quickly in a spontaneous energy-driven search of the conformational space. Reproduced with permission from Figure 1 of Liwo et al. (2002)

$$g(x; x_{\min}, x_{\max}) = \begin{cases} \frac{1}{4}(x - x_{\min})^4 & x < x_{\min}, \\ \frac{1}{4}(x - x_{\max})^4 & x > x_{\max}, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

The indices i and j indicate level numbers ($\{i\}$ denoting the conformations of level i). The absolute temperature is denoted by T and the inverse temperature $\beta = 1/k_B T$, where k_B is the Boltzmann constant. T runs from below to above the folding temperature. $\Delta_{ij}^{(1)}$ and $\Delta_{ij}^{(2)}$ are the lower and the upper bounds, respectively,

of the free energy gap between levels i and j , The first sum is taken over all training proteins. For each conformation, ω_i is the logarithm of the frequency with which a conformation is to be counted; it is 0 for canonical simulations, and for multicanonical simulations, it is calculated by the weighted histogram analysis method (WHAM) (Kumar et al. 1992; Liwo et al. 2007).

With this version of hierarchical optimization method, more complex training proteins could be treated that contained α , β , and $\alpha + \beta$ folds (Ołdziej et al. 2004a, b). A transferable force field was produced with the use of four training proteins, which are identified by their PDB codes: 1GAB (a three- α -helix bundle), 1E0L (a three-stranded antiparallel β -sheet), 1E0G, and 1IGD ($\alpha + \beta$ proteins) (Ołdziej et al. 2004a). This force field produced reasonably good results in CASP5 and CASP6 blind-test protein-structure prediction experiments and performed well even in coarse-grained molecular dynamics simulations. However, while the glassiness problem was effectively eliminated by the introduction of a hierarchy of gradually folded structures, the instability problem mentioned in the previous section still remained, and the force field with the parameters obtained by minimizing the target function in a given iteration generated usually nonnative decoys in the subsequent CSA simulations. This problem was somewhat alleviated (but not completely resolved) by including the representatives of the decoys from all iterations in computing the free energies.

While the force field optimized as described above performed well in finding native-like structures and pathways of protein folding, it was not good in reproducing folding thermodynamics. Therefore, the hierarchical optimization method was further modified. The first modification was to require that the free energy of a level decreases with increasing native likeness below the folding-transition temperature, it increases above the folding-transition temperature, and the free energies of all levels become equivalent at the folding-transition temperature. The folding-transition temperatures of the training proteins were taken from the experiment, and the free energy gaps at different temperatures were estimated from the experimental temperature dependence of the free energy of the unfolding of the training proteins. The second modification was made to use replica-exchange molecular dynamics (Hansmann and Okamoto 1994; Pande et al. 2003) to generate decoys and the weighted histogram analysis method (Kumar et al. 1992) to compute free energy gaps; at early optimization stages, umbrella sampling replica exchange was used in order to generate native-like structures in the conformational ensemble (Liwo et al. 2007).

The revised hierarchical optimization method resulted in versions of the UNRES force field which reproduced reasonably good thermodynamics of the protein folding (Liwo et al. 2007). To alleviate the iteration instability problem, a Shannon entropy term was added to the target function to keep the shape of the probability distribution of the conformations of a given level close to the initial one and, therefore, to avoid the appearance of a handful of *lead* conformations on which the target function is focused (Liwo et al. 2007). With this approach, the UNRES force field was parameterized with the 1GAB (three- α -helix-bundle) protein, which proved very reliable in predicting the structures and folding pathways of α -helical

proteins; when distance restraints were implemented, it was successfully used to study functionally important *open-to-closed-conformation* transitions of a bacterial Hsp70 chaperone (Golas et al. 2012) and to study the association of the iron-sulfur-binding (Isu1) protein with the Jac1 Hsp40 co-chaperone (Mozolewska et al. 2015). Another version of UNRES was obtained by using 1E0L, and it was successfully used to study amyloid formation from the A β 42 peptide (Rojas et al. 2010, 2011).

Even with the modifications mentioned above, the optimization procedure was not sufficiently stable. Therefore, an alternative method of scanning the parameter space with the use of two small proteins, tryptophan cage and tryptophan zipper, was implemented (He et al. 2009), which bypassed the local minimization stage. The resulting force field turned out to be quite successful, as proved in the CASP10 exercise (He et al. 2013), although the predicted secondary-structure information needs to be provided for more reliable predictions. This force field was also used to study biological processes such as Pick1 binding to the Bar1 domain (He et al. 2011).

Maximum Likelihood Optimization

Although the hierarchical method performs much better in the optimization of the energy functions for free simulations, its disadvantage is a large degree of arbitrariness of the assignment of conformations to levels. Cutoffs on the number of native contacts, the root-mean-square deviation (RMSD) of the respective fragment from the experimental structure, etc., must be applied to assign conformations to particular hierarchy levels, and cases when two very similar conformations are assigned to different structural levels occur frequently. Moreover, experimental NMR structures often contain ensembles of conformations with quite flexible parts, which must be deleted from comparison.

In addition to this, the idea of hierarchical optimization conforms strictly only with the nucleation-condensation mechanism of protein folding. However, experimental and theoretical studies of protein folding demonstrated that the folding process occurs through a molten globule stage (Ptitsyn 1995). In particular, in our simulations of the N-terminal part of the B-domain of staphylococcal protein A (PDB code: 1BDD) (Maisuradze et al. 2010), we observed that the protein preserves its overall shape even at temperatures higher than the folding-transition temperature, with the native contacts in place or shifted by a few residue, and only the secondary structure is melted. This observation was confirmed by a detailed NMR study of tryptophan cage (Hałabis et al. 2012) (Fig. 4).

On the basis of the facts pointed out above, we recently developed a new method of energy function optimization, which is based on achieving maximum similarity between calculated and experimental ensembles at the temperatures at which the experimental ensembles were determined (Zaborowski et al. 2015). The maximum likelihood method is a natural way to solve this problem: each conformation determined experimentally can be considered as a point in the conformational space, at which the Boltzmann probability calculated from the energy function could be

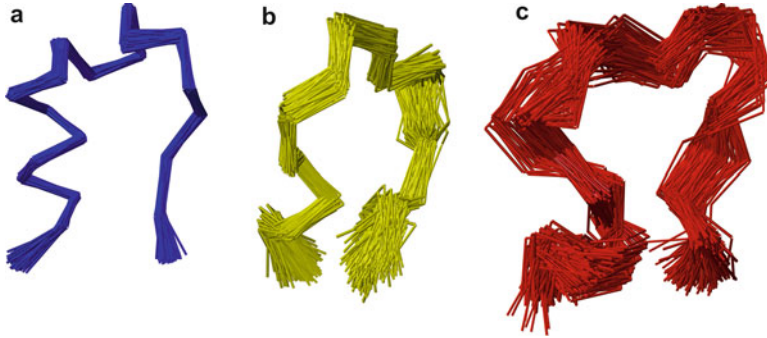


Fig. 4 NMR-determined ensembles of the tryptophan cage miniprotein at $T = 280$ K (a), $T = 305$ K (b), and $T = 313$ K (c). The pictures were drawn based on the data of Hałabis et al. (2012)

evaluated and the sum of logarithms of probabilities maximized. However, to do so, the exact method to calculate the partition function would be needed. Simulations provide only the relative probabilities of the conformations from the simulated set, and these calculations cannot be extended beyond this set. To overcome this difficulty, an approach has been recently designed in our laboratory, in which each experimental conformation is assigned with contributions from all of the simulated conformations with Gaussian weights depending on the distance from the experimental conformation. The target function is defined as

$$l_N^* = \ln Q(T) - \sum_{j=1}^N a_j \left(\omega_j - \frac{U_j}{RT} \right) \quad (8)$$

where U_j denotes the energy of the j th simulated conformation, ω_j is defined below Eq. 6, $Q(T)$ is the partition function computed from simulations, and the coefficients $a_i, i = 1, 2, \dots, n$ are defined by Eq. 9; it should be noted that these quantities comprise the similarity of the simulated conformations to the experimental conformations. It should also be noted that we made use of the fact that by Eq. 9, the coefficients a_i are normalized to 1.

$$a_j = \frac{1}{A} \sum_{i=1}^n \frac{\exp\left(-\frac{\|\mathbf{X}_i - \mathbf{Y}_j\|^2}{2s^2}\right)}{\sum_{k=1}^N \exp\left(-\frac{\|\mathbf{Y}_j - \mathbf{Y}_k\|^2}{2s^2}\right)} w_i \quad (9)$$

$$A = \sum_{j=1}^N \sum_{i=1}^n \frac{\exp\left(-\frac{\|\mathbf{X}_i - \mathbf{Y}_j\|^2}{2s^2}\right)}{\sum_{k=1}^N \exp\left(-\frac{\|\mathbf{Y}_j - \mathbf{Y}_k\|^2}{2s^2}\right)} w_i \quad (10)$$

Here, \mathbf{X}_i is the i th experimental conformation, \mathbf{Y}_j is the j th simulated conformation, and s is the overlap radius, which is the parameter of the method.

In the zero temperature limit, using a single reference experimental conformation and replacing the soft Gaussian similarity function with a hard cutoff on the distance of the calculated from the experimental structure, the maximum likelihood method becomes equivalent to the energy-gap maximization method (Zaborowski et al. 2015).

The method was tested with optimization of the UNRES force field based on the experimental ensembles of conformations of the tryptophan cage determined at three temperatures, including the folding-transition temperature. Hamiltonian replica exchange molecular dynamics with restraints derived from the mean experimental conformation corresponding to each temperature was used to generate decoys. As opposed to the energy-gap, Z-score, or hierarchical optimization methods discussed in the previous sections, the iteration process was very stable, and a fast convergence was observed. The reason for this is that the target function expressed by Eq. 8 contains energies, not the exponents of energies as in the case of the hierarchical optimization method or the minimum energies as in the case of the energy-gap/Z-score optimization. Therefore, provided that the set of simulated conformations is dense enough near the experimental conformations (which can always be achieved when umbrella sampling Hamiltonian replica exchange is used), no lead conformations appear to bias the results of local minimization of the target function. The UNRES force field obtained by maximum likelihood optimization described above was transferable to other α -helical proteins, although it was trained with only one small protein. It could even reproduce some $\alpha + \beta$ folds (Zaborowski et al. 2015).

Very recently, maximum likelihood optimization was applied to calibrate the NARES-2P force field for nucleic acids. In this application of the maximum likelihood optimization method, the target function was extended by the addition of the sums of squares of the differences between the calculated and experimental heat capacity and the fraction of paired nucleic acid base profiles (He et al. 2015). With the experimental data of only two small DNA molecules, the resulting force field was able to reproduce the experimental melting temperatures, enthalpies, and entropies of 18 small DNA molecules with good accuracy.

Conclusions and Outlook

Calibration of macromolecular energy functions is a new frontier in force-field development. The methods used for the energy function calibration presented in this chapter (Bryant and Lawrence 1993; Miller et al. 1996; Brungelson and Wolynes 1987; Goldstein et al. 1992; Meller and Elber 2001; Schug and Wenzel 2006; Lee et al. 2001, 2002, 2004a,b, 2008; Pillardy et al. 2001; Liwo et al. 2002, 2007; Ołdziej et al. 2004a,b; Zaborowski et al. 2015) demonstrate the importance of a close interaction between theory and experiment. Apparently, experimental structures alone are not sufficient to obtain force fields that can

reproduce structures, dynamics, and thermodynamics of the folding of proteins and other biological macromolecules, protein-ligand and protein-protein association, DNA hybridization, etc.

The modified maximum likelihood function defined by Eq. 8 (Zaborowski et al. 2015) appears to be the best one for the optimization of force fields for free simulations. With this approach, both the instability of a single interaction of the optimization process and the arbitrariness of splitting the set of decoys into the subsets with various native-like elements (or just native-nonnative if energy-gap optimization is pursued) are avoided. On the other hand, the procedure itself is computationally very demanding because only convergent decoy sets can be accepted to start a new iteration. Therefore, the method seems to be the most appropriate to optimize coarse-grained force fields in which the number of degrees of freedom is much reduced (e.g., only two per residue in the UNRES and NARES-2P force fields) compared to all-atom force fields.

Acknowledgements This work was supported from the Polish National Science Centre by grant UMO-2013/10/M/ST4/00640 and from the Foundation for Polish Science by grant Mistrz 7./2013. Jooyoung Lee was supported by a National Research Foundation of Korea (NRF) grant funded by the Government of Korea (MEST) (2008-0061987). This research was conducted by using the resources of (a) the Polish Grid Infrastructure (PL-GRID); (b) the 846-processor Beowulf cluster at Baker Laboratory of Chemistry and Chemical Biology, Cornell University; (c) our 45-processor Beowulf cluster at the Faculty of Chemistry, University of Gdańsk; (d) the Informatics Center of the Metropolitan Academic Network (IC MAN) in Gdańsk; (e) the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputer Center; and (f) the Interdisciplinary Centre for Mathematical and Computer Modeling (ICM) at the University of Warsaw.

Bibliography

- Brooks, B. R., Brooks III, C. L., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Cafisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoseck, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., & Karplus, M. (2009). Charmm: The biomolecular simulation program. *Journal of Computational Chemistry*, *30*, 1545–1615.
- Brungelson, J. D., & Wolynes, P. G. (1987). Spin glasses and the statistical mechanics of protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, *84*, 7524–7528.
- Bryant, S. H., & Lawrence, C. E. (1993). An empirical energy function for threading protein sequence through the folding motif. *Proteins: Structure, Function, and Genetics*, *16*, 92–112.
- Camacho, C. J., & Thirumalai, D. (1996). A criterion that determines fast folding of proteins: A model study. *Europhysics Letters*, *35*, 627–632.
- Crippen, G. M., & Snow, M. E. (1990). A 1.8 Å resolution potential function for protein folding. *Biopolymers*, *29*, 1479–1489.
- Crippen, G. M., & Wiswanadhan, V. N. (1984). A potential function for conformational analysis of peptides. *International Journal of Peptide and Protein Research*, *24*, 279–296.
- Dahiyat, B. I., & Mayo, S. L. (1997). De novo protein design: Fully automated sequence selection. *Science*, *278*, 82–87.
- Dygert, M., Gö, N., & Scheraga, H. A. (1975). Use of a symmetry condition to compute the conformation of gramicidin S. *Macromolecules*, *8*, 750–761.

- Eastwood, M. P., Hardin, C., Luthey-Schulten, Z., & Wolynes, P. G. (2002). Statistical mechanical refinement of protein structure prediction schemes: Cumulant expansion approach. *Journal of Chemical Physics*, *117*, 4602–4615.
- Eastwood, M. P., Hardin, C., Luthey-Schulten, Z., & Wolynes, P. G. (2003). Statistical mechanical refinement of protein structure prediction schemes. II. Mayer cluster expansion approach. *Journal of Chemical Physics*, *118*, 8500–8512.
- Fain, B., & Levitt, M. (2003). Funnel sculpting for in silico assembly of secondary structure elements of proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 10700–10705.
- Fujitsuka, Y., Takada, S., Luthey-Schulten, Z. A., & Wolynes, P. G. (2004). Optimizing physical energy functions for protein folding. *Proteins: Structure, Function, and Genetics*, *54*, 88–103.
- Golas, E. I., Maisuradze, G. G., Senet, P., Ołdziej, S., Czaplewski, C., Scheraga, H. A., & Liwo, A. (2012). Simulation of the opening and closing of Hsp70 chaperones by coarse-grained molecular dynamics. *Journal of Chemical Theory and Computation*, *8*, 1334–1343.
- Goldstein, R. A., Luthey-Schulten, Z. A., & Wolynes, P. G. (1992). Protein tertiary structure recognition using optimized hamiltonians with local interactions. *Proceedings of the National Academy of Sciences of the United States of America*, *89*, 9029–9033.
- Halabis, A., Żmudzińska, W., Liwo, A., & Ołdziej, S. (2012). Conformational dynamics of the trp-cage miniprotein at its folding temperature. *Journal of Physical Chemistry B*, *116*, 6898–6907.
- Hansmann, U. H. E., & Okamoto, Y. (1994). Comparative study of multicanonical and simulated annealing algorithms in the protein folding problem. *Physica A*, *212*, 415–437.
- He, Y., Liwo, A., & Scheraga, H. A. (2015). Optimization of a nucleic acids united-residue 2-point model (NARES-2P) force field with a maximum-likelihood approach. *Journal of Chemical Physics*, *143*, 243111.
- He, Y., Liwo, A., Weinstein, H., & Scheraga, H. A. (2011). PDZ binding to the BAR domain of PICK1 is elucidated by coarse-grained molecular dynamics. *Journal of Molecular Biology*, *405*, 298–314.
- He, Y., Mozolewska, M. A., Krupa, P., Sieradzan, A. K., Wirecki, T. K., Liwo, A., Kachlishvili, K., Rackovsky, S., Jagieła, D., Ślusarz, R., Czaplewski, C. R., Ołdziej, S., & Scheraga, H. A. (2013). Lessons from application of the UNRES force field to predictions of structures of CASP10 targets. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 14936–14941.
- He, Y., Xiao, Y., Liwo, A., & Scheraga, H. A. (2009). Exploring the parameter space of the coarse-grained UNRES force field by random search: Selecting a transferable medium-resolution force field. *Journal of Computational Chemistry*, *30*, 2127–2135.
- Klimov, D. K., & Thirumalai, D. (1996). Criterion that determines the foldability of proteins. *Physical Review Letters*, *76*, 4070–4073.
- Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A., & Rosenberg, J. M. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, *13*, 1011–1021.
- Lee, J., Kim, S., & Lee, J. (2004a). Design of a protein potential energy landscape by parameter optimization. *Journal of Physical Chemistry B*, *108*, 4525–4534.
- Lee, J., Kim, S.-Y., & Lee, J. (2004b). Optimization of potential energy parameters for folding of several proteins. *Journal of the Korean Physical Society*, *44*, 594–598.
- Lee, J., Lee, J., Sasaki, T. N., Sasai, M., Seok, C., & Lee, J. (2008). De novo protein structure prediction by dynamic fragment assembly and conformational space annealing. *Proteins: Structure, Function, and Bioinformatics*, *79*, 2403–2417.
- Lee, J., Liwo, A., Ripoll, D. R., Pillardy, J., & Scheraga, H. A. (1999). Calculation of protein conformation by global optimization of a potential energy function. *Proteins: Structure, Function, and Genetics*, *37*(Suppl. 3), 204–208.
- Lee, J., Park, K., & Lee, J. (2002). Full optimization of linear parameters of a united residue protein potential. *Journal of Physical Chemistry B*, *106*, 11647–11657.
- Lee, J., Ripoll, D. R., Czaplewski, C., Pillardy, J., Wedemeyer, W. J., & Scheraga, H. A. (2001). Optimization of parameters in macromolecular potential energy functions by conformational

- space annealing. *Journal of Physical Chemistry B*, *105*, 7291–7298.
- Lee, J., & Scheraga, H. A. (1999). Conformational space annealing by parallel computations: Extensive conformational search of Met-enkephalin and of the 20-residue membrane-bound portion of melittin. *International Journal Of Quantum Chemistry*, *75*, 255–265.
- Lee, J., Scheraga, H. A., & Rackovsky, S. (1997). New optimization method for conformational energy calculations on polypeptides: Conformational space annealing. *Journal of Computational Chemistry*, *18*, 1222–1232.
- Levitt, M. (1976). A simplified representation of protein conformations for rapid simulation of protein folding. *Journal of Molecular Biology*, *104*, 59–107.
- Levitt, M., & Warshell, A. (1975). Computer simulation of protein folding. *Nature*, *253*, 694–698.
- Li, Z., & Scheraga, H. A. (1987). Monte Carlo-minimization approach to the multiple-minima problem in protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, *84*, 6611–6615.
- Liwo, A., Arlukowicz, P., Czaplowski, C., Ołdziej, S., Pillardy, J., & Scheraga, H.A. (2002). A method for optimizing potential-energy functions by a hierarchical design of the potential-energy landscape: Application to the UNRES force field. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 1937–1942.
- Liwo, A., Khalili, M., Czaplowski, C., Kalinowski, S., Ołdziej, S., Wachucik, K., & Scheraga, H.A. (2007). Modification and optimization of the united-residue (UNRES) potential energy function for canonical simulations. I. Temperature dependence of the effective energy function and tests of the optimization method with single training proteins. *Journal of Physical Chemistry B*, *111*, 260–285.
- Liwo, A., Lee, J., Ripoll, D. R., Pillardy, J., & Scheraga, H. A. (1999). Protein structure prediction by global optimization of a potential energy function. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 5482–5485.
- Liwo, A., Ołdziej, S., Czaplowski, C., Kozłowska, U., & Scheraga, H. A. (2004). Parameterization of backbone-electrostatic and multibody contributions to the UNRES force field for protein-structure prediction from *ab initio* energy surfaces of model systems. *Journal of Physical Chemistry B*, *108*, 9421–9438.
- Liwo, A., Pincus, M. R., Wawak, R. J., Rackovsky, S., Ołdziej, S., & Scheraga, H. A. (1997). A united-residue force field for off-lattice protein-structure simulations. II: Parameterization of local interactions and determination of the weights of energy terms by Z-score optimization. *Journal of Computational Chemistry*, *18*, 874–887.
- Mackerell, A. D., Jr. (2004). Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, *25*, 1584–1604.
- Maisuradze, G. G., Senet, P., Czaplowski, C., Liwo, A., & Scheraga, H. A. (2010). Investigation of protein folding by coarse-grained molecular dynamics with the UNRES force field. *Journal of Physical Chemistry A*, *114*, 4471–4485.
- Meller, J., & Elber, R. (2001). Linear programming optimization and a double statistical filter for protein threading protocols. *Proteins: Structure, Function, and Genetics*, *45*, 241–261.
- Micheletti, C., Seno, F., Maritan, A., & Banavar, J. R. (2001). An optimal procedure to extract interaction potentials for protein folding. *Computational Materials Science*, *20*, 305–310.
- Miller, R. T., Jones, D. T., & Thornton, J. M. (1996). Protein fold recognition by sequence threading: Tools and assessment techniques. *The FASEB Journal*, *10*, 171–178.
- Mozolewska, M., Krupa, P., Scheraga, H. A., & Liwo, A. (2015). Molecular modeling of the binding modes of the iron-sulfur protein to the jac1 co-chaperone from *Saccharomyces cerevisiae* by all-atom and coarse-grained approaches. *Proteins: Structure, Function, and Bioinformatics*, *83*, 1414–1426.
- Némethy, G., Gibson, K. D., Palmer, K. A., Yoon, C. N., Paterlini, G., Zagari, A., Rumsey, S., & Scheraga, H. A. (1992). Energy parameters in polypeptides. 10. Improved geometrical parameters and nonbonded interactions for use in the ECEPP/3 algorithm with application to proline-containing peptides. *The Journal of Physical Chemistry*, *96*, 6472–6484.
- Obatake, M., & Crippen, G. M. (1981). Residue-residue potential function for conformational analysis of proteins. *The Journal of Physical Chemistry*, *85*, 1187–1197.

- Oldziej, S., Łągiewka, J., Liwo, A., Czaplewski, C., Chinchio, M., Nancias, M., & Scheraga, H. A. (2004a). Optimization of the UNRES force field by hierarchical design of the potential-energy landscape. 3. Use of many proteins in optimization. *Journal of Physical Chemistry B*, *108*, 16950–16959.
- Oldziej, S., Liwo, A., Czaplewski, C., Pillardy, J., & Scheraga, H. A. (2004b). Optimization of the UNRES force field by hierarchical design of the potential-energy landscape: 2. Off-lattice tests of the method with single proteins. *Journal of Physical Chemistry B*, *108*, 16934–16949.
- Pande, V. S., Baker, I., Chapman, J., Elmer, S., Kaliq, S., Larson, S. M., Rhee, Y. M., Shirts, M. R., Snow, C. D., Sorin, E. J., & Zagrovic, B. (2003). Atomistic protein folding simulations on the submillisecond timescale using worldwide distributed computing. *Biopolymers*, *68*, 91–109.
- Pearlman, D., Case, D., Caldwell, J., Ross, W., Cheatham III, T., DeBolt, S., Ferguson, D., Seibel, G., & Kollman, P. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Computer Physics Communications*, *91*, 1–41.
- Pillardy, J., Czaplewski, C., Liwo, A., Wedemeyer, W. J., Lee, J., Ripoll, D. R., Arlukowicz, P., Oldziej, S., Arnautova, Y. A., & Scheraga, H. A. (2001). Development of physics-based energy functions that predict medium-resolution structure for proteins of the α , β , and α/β structural classes. *Journal of Physical Chemistry B*, *105*, 7299–7311.
- Prentiss, M. C., Hardin, C., Eastwood, M. P., Zong, C. H., & Wolynes, P. G. (2006). Protein structure prediction: The next generation. *Journal of Chemical Theory and Computation*, *2*, 705–716.
- Pittsytyn, O. B. (1995). Molten globule and protein folding. *Advances in Protein Chemistry*, *47*, 83–229.
- Rojas, A., Liwo, A., Browne, D., & Scheraga, H. A. (2010). Mechanism of fiber assembly; treatment of $A\beta$ -peptide aggregation with a coarse-grained united-residue force field. *Journal of Molecular Biology*, *404*, 537–552.
- Rojas, A., Liwo, A., & Scheraga, H. A. (2011). A study of the α -helical intermediate preceding the aggregation of the amino-terminal fragment of the $A\beta$ -amyloid peptide (1–28). *Journal of Physical Chemistry B*, *115*, 12978–12983.
- Sali, A., Shakhnovich, E., & Karplus, M. (1994). How does a protein fold? *Nature*, *369*, 248–251.
- Schug, A., & Wenzel, W. (2006). An evolutionary strategy for all-atom folding of the 60-amino-acid bacterial ribosomal protein L20. *Biophysical Journal*, *90*, 4273–4280.
- Seetharamulu, P., & Crippen, G. M. (1991). A potential function for protein folding. *Journal of Mathematical Chemistry*, *6*, 91–110.
- Shakhnovich, E. I., Farzdtinov, G., Gutin, A. M., & Karplus, M. (1991). Protein folding bottlenecks – A lattice Monte-Carlo simulation. *Physical Review Letters*, *67*, 1665–1668.
- Snow, M. E. (1992). Powerful simulated-annealing algorithm locates global minimum of protein-folding potentials from multiple starting points. *Journal of Computational Chemistry*, *13*, 579–584.
- Yang, F., Gustafson, K. R., Boyd, M. R., & Wlodawer, A. (1998). Crystal structure of *Escherichia coli* HDEA. *Nature Structural Biology*, *5*, 763–764.
- Zaborowski, B., Jagieła, D., Czaplewski, C., Hałabis, A., Lewandowska, A., Żmudzińska, W., Oldziej, S., Karczyńska, A., Omieczynski, C., Wirecki, T., & Liwo, A. (2015). A maximum-likelihood approach to force-field calibration. *Journal of Chemical Information and Modeling*, *55*, 2050–2070.
- Zhang, Y., & Skolnick, J. (2004). Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics*, *57*, 702–710.

Georgios Leonis, Georgia Melagraki, and Antreas Afantitis

Contents

Introduction	2202
Popular Software Tools for Chemoinformatics	2203
Chemoinformatics Studies Using the Enalos KNIME Nodes and Enalos Cloud Platform ..	2205
Risk Assessment Tool for the Toxicity Prediction of Iron Oxide Nanoparticles Through Enalos InSilicoNano Cloud Platform	2206
Cellular Uptake Prediction and Virtual Screening of Nanoparticles via Enalos InSilicoNano Platform	2211
Identification of Organic Materials as Corrosion Inhibitors Based on Enalos KNIME Nodes	2214
KNIME-Based Design and Prediction of Compound Structures	2221
Prediction of Yellow Fever Inhibitors from ChEMBL Database Through KNIME Classification Analysis	2221
De Novo Design of Synthetically Feasible Compounds Through the Reaction Vectors Approach	2223
Bibliography	2226

Abstract

In this chapter, we present a brief description of compound datasets and programs developed to serve chemoinformatics as well as, more specifically, nanoinformatics purposes. Emphasis has been placed on publicly available tools and particularly on KNIME (Konstanz Information Miner), the most widely used freely available platform for data processing and analysis. Among a multitude

G. Leonis
Novamechanics Ltd, Nicosia, Cyprus

G. Melagraki (✉) • A. Afantitis (✉)
InSilicoLab LP, Athens, Greece
Novamechanics Ltd, Nicosia, Cyprus
e-mail: melagraki@novamechanics.com; melagraki@insilicolab.eu;
afantitis@novamechanics.com; afantitis@insilicolab.eu

of studies that have demonstrated the usefulness of chemoinformatics tools to chemical and medicinal applications, herein we present indicative cases of five successful KNIME-based approaches. The first two studies include the risk assessment of nanoparticles (NPs) through the Enalos InSilicoNano platform, namely, (1) the prediction of the toxicity of iron oxide NPs and (2) the cellular uptake prediction of computationally designed NPs with the aid of reliable quantitative nanostructure–activity relationships (QNAR) models. The third case study deals with the recognition of organic substances as corrosion inhibitors through the construction of predictive quantitative structure–property relationships (QSPR) models with Enalos KNIME nodes. Finally, two more cases are briefly described and involve the accurate prediction of yellow fever inhibitors from the ChEMBL database and the *de novo* design of compounds with the reaction vectors methodology. The aim of this work is to familiarize the interested reader with the freely available *in silico* tools in KNIME analytics platform and to demonstrate their value and effectiveness toward specific computational applications.

Introduction

Chemoinformatics employs computational methods and information technology to deal with chemical problems (Leach and Gillet 2007). Current efforts in the field of drug discovery are particularly concerned with the handling of chemical structural information so that properties of a ligand are optimized to address the multiple demands of a potent drug (Brown 1998). Chemoinformatics mainly emerged due to the enormous amount of data that has been generated by recent drug discovery attempts, including high-throughput screening and combinatorial chemistry methodologies (Russo 2002). Chemoinformatics include several modeling approaches aiming at successful drug design. For instance, the development of quantitative relationships between the observed biological activities and the chemical structures through construction of quantitative structure–activity relationships (QSARs) models and the prediction of ligand–protein structures via docking approaches are among the most widely used techniques. It is important to highlight that the methods employed in chemoinformatics are usually developed to handle large sets of chemical structures and their different properties (usually referred as molecular descriptors) and thus should be appropriate for big data analysis.

Through the course of the years, the development of chemoinformatics approaches has been greatly assisted by data mining tools and open-source software. The recently released curated small-molecule databases are valuable assets for testing and validating chemoinformatics algorithms and tools (Hu and Bajorath 2012). There are many databases (either publicly available or proprietary) containing a number of chemical substances. The size of typical databases ranges between tens of thousands and millions of compound entries (Leach and Gillet 2007). The most well used public databases are PubChem (Wang et al. 2009), ZINC (Irwin et al. 2012), BindingDB (Liu et al. 2007), and

ChEMBL (Gaulton et al. 2012). ZINC contains the three-dimensional structures of commercially available compounds, which were constructed to be used in structure-based virtual screening, while PubChem, BindingDB, and ChEMBL also include (bio)activity information. Additionally, PubChem provides screening information data sets (SIDS) and the three-dimensional structures of the majority of the compounds, and in BindingDB and ChEMBL databases the activities of compounds can be assembled into relevant classes. Activity classes are of particular usefulness toward benchmarking of new computational methodologies (Hu and Bajorath 2012).

Recently, virtual databases have been built. Virtual compounds are substances that have not been observed/synthesized so far, but they could be synthesized. This broader consideration allows for the construction of even larger virtual libraries containing billions of compounds (Leach and Gillet 2007).

The release of these important sources of information that can be systematically explored have boosted the development of many software tools for chemoinformatics that make extensive use of datasets. Among others, software tools have been particularly emerged to tackle research practices, such as data mining, virtual screening and machine learning, molecular selectivity analysis, and visualization of structure-activity relationships (SARs).

In this chapter, we briefly review some of the most popular chemoinformatics tools with particular emphasis on programs that are either publicly available or at least free for academic purposes. Since KNIME is the most widely used, freely available platform for chemoinformatics applications, we will specifically present tools that have been integrated into KNIME and are offered as KNIME nodes to execute several important tasks for chemoinformatics analysis. Such applications include CDK, Indigo, RDKit, Vernalis, CACTVS, Enalos, Lhasa, OpenBabel, OCHEM, Chemical Identifier Resolver, ErlWood, EMBL-EBI Nodes, and CheS- Mapper.

Popular Software Tools for Chemoinformatics

Konstanz Information Miner (KNIME) is an open-source analytics platform, which is the leading tool for wide-ranging data processing, integration, analysis, and exploration (Berthold et al. 2008). It enables the visual creation of data flows (so-called pipelines), the selective execution of specified analysis steps, and the presentation of the results via interactive views on models and data. KNIME offers intuitive use and high level of scalability, which currently render it the most popular platform for chemoinformatics applications. Therefore, as already mentioned, this chapter will be mostly devoted to the description of software tools that employ KNIME to accomplish their functions.

The Chemistry Development Kit (CDK) is an open-source and development chemoinformatics software (Steinbeck et al. 2003). In collaboration with the KNIME group, the CDK nodes for KNIME have been recently developed. These nodes provide features regarding chemical compound handling, such as several file conversion applications for molecules, calculation and drawing of 2D and 3D

structures, symmetry group calculations, fingerprint calculation, proper handling of hydrogen atoms, and molecular property estimations, among others.

Indigo (developed by GGA Software Services LLC) is a software tool for organic chemistry (<http://lifescience.opensource.epam.com/indigo/>). Manipulation and functionality of organic structures with the Indigo nodes for KNIME can be obtained through conversions to Kekulé and aromatic states, handling of hydrogen, molecular properties generation, fingerprint comparison, R-group decomposition, stereochemistry calculation, and component separation. Additional functionalities include file conversions among SDF, SMILES, and CML formats; detection of drawing errors in structures; 2D structure generation; and structure matching.

RDKit (<http://www.rdkit.org/>) also provides chemoinformatics applications through KNIME, for instance, substructure filtering and searching, 2D and 3D structure generation, chemical reactions, molecular fingerprinting, salt separation from compounds, and R-group decomposition.

Vernalis Research (<http://www.vernalis.com/research>) employs its KNIME nodes to assist structure-based and fragment-based drug discovery. It includes several functionalities, such as flow control, PDB and sequence tools, I/O applications, matched molecular pairs, and fingerprint properties.

Lhasa Limited nodes for KNIME offer additional operations on the evaluation of binary classification models and table manipulation (<http://www.lhasalimited.org/>).

The OpenBabel chemoinformatics package (<http://openbabel.org/>) is primarily a file converter toolbox (KNIME chemistry nodes). It can also filter molecular files with SMARTS and has a wide applicability in analyzing molecular modeling and bioinformatics data (O'Boyle et al. 2011).

The Online Chemical Modeling Environment (OCHEM) is a platform, which aims to simplify the procedures for performing QSAR calculations (Sushko et al. 2011). This is achieved through combination of experimental results taken from a database and a modeling procedure. The database contains thousands of entries and a user-friendly application environment. The use of current KNIME nodes for OCHEM is restricted to some running predictions and the data export/import.

KNIME nodes development from ErlWood also offers interpretation and handling of structure-activity relationship data, as well as various compound viewing facilities (<https://tech.knime.org/community/erlwood>).

The European Bioinformatics Institute (www.ebi.ac.uk) provides the EMBL-EBI KNIME nodes, which make use of the Chemical Entities of Biological Interest (ChEBI) database to obtain database files via ChEBI IDs, substructure, or keyword searches. ChEBI is a publicly available depository of small chemical molecules.

The visualization node CheS-Mapper (Gutlein et al. 2012) employs the characteristics of small molecule structures to perform clustering according to feature similarity criteria. 3D structure depiction and embedding of compound datasets are also supported.

The Chemical Identifier Resolver (CIR, developed by the CADD group at the National Cancer Institute) nodes for KNIME enable the recognition of a chemical structure provided that an identifier is known (<http://cactus.nci.nih.gov/chemical/>



Fig. 1 Freely available Enalos KNIME nodes

structure). CIR is a resolver for various structure identifiers and can also convert a particular structure identifier into another one.

Finally, the functionality of the Enalos KNIME nodes (<https://tech.knime.org/community/enalos-nodes>) developed by NovaMechanics Ltd will be briefly described. Enalos nodes for the KNIME platform are associated with several important aspects regarding data analysis and curation for chemoinformatics and nanoinformatics. The Enalos nodes (Fig. 1) among others provide (1) domain-similarity based on (i) euclidean distances or (ii) leverages, (2) fit quality and predictive power of a QSAR model with the Model Acceptability Criteria node, (3) the newly developed fast generation of all possible substitutions of a lead compound, and (4) calculation of important molecular descriptors with the Mold2 node. Mold2 is able to evaluate large and diverse sets of molecular descriptors from two-dimensional chemical structure information (Hong et al. 2008). Comparison of Mold2 descriptors with descriptors calculated from commercial software on several published datasets showed that Mold2 descriptors yield models with higher quality than other packages and also produce sufficient structural information.

In the following section, we present three case studies of *in silico* approaches developed by our group that involve the utilization of Enalos KNIME nodes for building predictive models.

Chemoinformatics Studies Using the Enalos KNIME Nodes and Enalos Cloud Platform

The KNIME analytics platform contains a multitude of processing nodes for data I/O, modeling, analysis, and data mining. It integrates well with Weka programming language for machine learning applications. Over the last 30 years, a vast amount of data has been generated within the areas of chemo-, bio-, and nanoinformatics. KNIME has emerged as one of the most reliable open-source data mining tools for the prediction of chemical properties and applications, such as virtual screening of chemicals and nanoparticles (NPs), chemical library enumeration, virtual library creation, building QSAR/QSPR, ADMET, and pharmacokinetics models, as well as prediction of various biological effects of organic compounds and NPs.

In this section, we present three case studies that involve the use of Enalos KNIME nodes in predicting (i) the toxicity of iron oxide NPs, (ii) the cellular uptake of organic NPs (virtual screening is also demonstrated), and (iii) the inhibitory potency of organic substances against corrosion.

In the first two case studies Enalos InSilico platform is also introduced. Enalos InSilico platform is a Cloud platform built to host a variety of predictive models to address the need for risk assessment and virtual screening. Workflows included in Enalos InSilico platform are constructed based on reliable data information, and each workflow combines advanced in silico tools to yield accurate predictions. Predictive workflows are available in a user-friendly format and include toxicity, biological activity, and property evaluation models.

Risk Assessment Tool for the Toxicity Prediction of Iron Oxide Nanoparticles Through Enalos InSilicoNano Cloud Platform

Nanoparticles (NPs) are known for their unique optical, electronic, and mechanical properties, which have led to the rapid evolution of nanotechnology materials being applied in a wide range of commercial, technological, and therapeutic applications in fields such as environment, industry, defense, electronics, and biomedicine. The latter has enjoyed great scientific, technological, and commercial progress in different NP applications (Gajewicz et al. 2012; Cohen et al. 2013).

Along with the apparent increasing use of NPs, concerns on their effect upon the environment and human health have been raised. Since toxicity assessment of NPs via traditional experimental routes often require expensive and time-consuming procedures, computational approaches such as quantitative nanostructure–activity relationships (QNARs) have been successfully used to predict the toxic effects of NPs (Vrontaki et al. 2015; Kleandrova et al. 2014a, b; Speck-Planche et al. 2015; Winkler et al. 2013, 2014; Shao et al. 2013; Toropov et al. 2013). However, the computational investigation of NP toxicity is seriously hindered by the lack of available NP descriptors, organized datasets, and systematic experimental data for NPs. Only few organized datasets on NP toxicity are available so far. Among these toxicity datasets on nanostructures, metal oxide NP data have been investigated in several computational studies (Fourches et al. 2010; Liu et al. 2011, 2013a; Puzyn et al. 2011; Zhang et al. 2012).

A fully validated QNAR model is presented, which was constructed based on toxicity data of iron oxide NPs with different core, coating, and surface modifications (Melagraki and Afantitis 2015; Shaw et al. 2008; Liu et al. 2013). The initial dataset was constructed with 44 iron oxide NPs that comprised a core with either Fe₂O₃ or Fe₃O₄ coating, including cross-linked dextran, PVA, or other, and various surface modifications (Shaw et al. 2008; Liu et al. 2013). Values for descriptors such as the size, R1 and R2 relaxivities, and zeta potential along with a coating-indicative parameter were considered as independent variables for the model development. The values of the input variables for each NP and the corresponding toxicity class are shown in Table 1.

Table 1 NP properties, bioactivity, and predictions

ID	NP ^b	Size	ZP	R1 ^c	R2	Coating	NHit ^d	Class	Prediction	Domain
1 ^a	NP1	36	-19.9	19	45	Cross-linked dextran	1	Inactive	Inactive	Reliable
2 ^a	NP2	30	-9.22	26	74	Cross-linked dextran	1	Inactive	Inactive	Reliable
3	NP3	32	5.9	21	54	Cross-linked dextran	3	Inactive	Inactive	-
4	NP4	74	-2.72	21	153	Cross-linked dextran	2	Inactive	Inactive	-
5	NP5	27	3.34	17	36	Cross-linked dextran	0	Inactive	Inactive	-
6 ^a	NP6	29	1.95	22	51	Cross-linked dextran	2	Inactive	Inactive	Reliable
7 ^a	NP7	38	-10.1	21	62	Cross-linked dextran	1	Inactive	Inactive	Reliable
8	NP8	33	-19.5	22	49	Cross-linked dextran	0	Inactive	Inactive	-
9	NP9	36	-14	19	45	Cross-linked dextran	3	Inactive	Inactive	-
10	NP10	28	3.24	19	39	Cross-linked dextran	1	Inactive	Inactive	-
11	NP11	31	-9.46	23	59	Cross-linked dextran	4	Inactive	Inactive	-
12	NP12	31	3.64	19	49	Cross-linked dextran	17	Active	Active	-
13 ^a	NP14	28	2.34	19	39	Cross-linked dextran	4	Inactive	Inactive	Reliable
14	NP15	24	-11.7	22	54	Cross-linked dextran	1	Inactive	Inactive	-
15	NP16	37	0.766	21	52	Cross-linked dextran	2	Inactive	Inactive	-
16	NP17	38	-20.7	21	62	Cross-linked dextran	3	Inactive	Inactive	-
17	NP18	38	-9.08	21	62	Cross-linked dextran	0	Inactive	Inactive	-
18	NP19	31	-3.61	19	49	Cross-linked dextran	8	Active	Active	-
19 ^a	NP20	38	-9.34	21	62	Cross-linked dextran	7	Active	Inactive	Reliable
20	NP21	28	-9.23	15	40	Cross-linked dextran	4	Inactive	Inactive	-
21	NP22	36	-21.9	36	122	Cross-linked dextran	2	Inactive	Inactive	-
22	NP23	31	-6.11	20	45	Cross-linked dextran	3	Inactive	Inactive	-
23	NP26	40	-12	15	30	PVA	3	Inactive	Active	-
24 ^a	NP27	40	-3.77	15	30	PVA	0	Inactive	Active	Reliable
25	NP28	40	-7.57	15	30	PVA	5	Active	Active	-
26	NP29	40	0.25	15	30	PVA	7	Active	Active	-
27	NP30	40	-6.05	15	30	PVA	5	Active	Active	-
28	NP31	20	-12.3	0.5	0.5	PVA	4	Inactive	Active	-
29	NP32	20	-4.22	0.5	0.5	PVA	8	Active	Active	-
30	NP33	20	-7.15	0.5	0.5	PVA	0	Inactive	Active	-
31 ^a	NP34	20	-4.3	0.5	0.5	Other	13	Active	Active	Reliable
32	NP35	20	-12.1	0.5	0.5	PVA	8	Active	Active	-
33	NP36	20	-15.6	0.5	0.5	Other	9	Active	Active	-
34	NP37	20	-16.1	0.5	0.5	PVA	5	Active	Active	-
35	NP38	20	-4.7	0.5	0.5	PVA	13	Active	Active	-
36	NP39	20	-6.47	0.5	0.5	PVA	9	Active	Active	-
37 ^a	NP40	20	-6.54	0.5	0.5	PVA	6	Active	Active	Reliable
38	NP41	20	-10.8	0.5	0.5	Other	2	Inactive	Inactive	-
39 ^a	NP42	20	-7.7	0.5	0.5	PVA	6	Active	Active	Reliable
40 ^a	NP43	20	-6.75	0.5	0.5	PVA	6	Active	Active	Reliable

(continued)

Table 1 (continued)

ID	NP ^b	Size	ZP	R1 ^c	R2	Coating	NHit ^d	Class	Prediction	Domain
41	NP45	23	-13.6	29	62	Other	1	Inactive	Inactive	-
42	NP46	33	-14.5	36	106	Other	1	Inactive	Inactive	-
43 ^a	NP47	28	-9.23	32	60	Other	0	Inactive	Inactive	Reliable
44 ^a	NP48	25	-37	29	49	Other	1	Inactive	Inactive	Reliable

^aTest Set

^bThe 44 NPs were obtained from the following studies (Liu et al. 2011; Epa et al. 2012)

^cR1: spin-lattice relaxivity; R2: spin-spin relaxivity

^dNHit: the number of hits identified for each NP across the 64 bioactivity measures (4 cell lines X 4 assays X 4 concentrations)

NPs were evaluated with different assays in various cell types and concentrations that produced a 64-component vectorial metric. Each NP was characterized either active or inactive according to the number of hits obtained across the 64 bioactivity measures. For the QNAR developed KNIME workflow, the following steps have been integrated: (1) data preprocessing, (2) variable selection, (3) model development and (4) validation, and (5) determination of domain of applicability [via the Enalos Domain – Similarity node that defines applicability domain (APD) based on euclidean distances]. The publicly available set of Enalos KNIME nodes can be accessed through either the KNIME Community or NovaMechanics website (www.novamechanics.com/knime.php or www.insilicotox.com/index.php/products/enalos-knime-nodes-community-contributions/) (Melagraki and Afantitis 2013).

Before model running, the available data were separated into training set and test set with the partitioning node in KNIME. According to the data in the training set, the most significant descriptors were selected (Witten et al. 2005; Hall et al. 2009). The NPs used in the validation set were not further employed during model development. Among all available techniques, the J48 modeling method yielded the most predictively powerful model for the available data.

The proposed predictive model was validated internally and externally regarding goodness-of-fit, robustness, and predictivity, thus totally meeting the criteria recommended by the Organization for Economic Cooperation and Development (OECD).

To validate the performance of the model, the following parameters were considered (Afantitis et al. 2011; Mouchlis et al. 2012):

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{Specificity} = \text{TN} / (\text{TN} + \text{FP}) \text{ and}$$

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{FN} + \text{TN}),$$

where TP true positive, FP false positive, TN true negative, and FN false negative.

The confusion matrix is presented below:

	Positive predicted	Negative predicted
Positive observed (Active)	TP	FN
Negative observed (Inactive)	FP	TN

Table 2 Confusion matrix (training set)

	Positive predicted	Negative predicted
Positive observed (Inactive)	16	0
Negative observed (Active)	4	11

Table 3 Confusion matrix (test set)

	Positive predicted	Negative predicted
Positive observed (Inactive)	7	1
Negative observed (Active)	1	4

Additionally, the Y-randomization test demonstrated the robustness and the statistical significance of the predictive model.

The ability to perform virtual screening of NPs that were not originally included in the dataset is particularly important, especially if there is an indication on its reliability. For this purpose, it is crucial to determine the limits of the model's domain of applicability. This will identify NPs that are excluded from the area of reliable predictions of the proposed model.

In this work, the Enalos Domain-Similarity node calculated the domain of applicability using euclidean distances to estimate the similarity between NPs belonging to the training and test sets. More details on the domain of applicability calculation can be found elsewhere (Melagraki and Afantitis 2013; Afantitis et al. 2008; Tropsha 2010).

Based on the reported toxicity data for the 44 iron oxide NPs used in this study, 31 compounds were included in the training set while the other 13 in the test set. As mentioned above, subsequent model development was based on the NP structures of the training set. For the NPs, experimental parameters such as the size, R1 and R2 relaxivities, and zeta potential were combined with a coating-specific parameter and were used as inputs for the model development. The three coating categories which have been considered are: PVA, cross-linked dextran, and other. Among the above descriptors, the subset that best describes the variation of toxicity with NP properties (as demonstrated by the variable selection algorithm) includes the R1 and R2 relaxivities and coating.


After the model development from the training data was achieved, the toxicity prediction for the test set followed. Observed classes and predictions for the 44 NPs of the initial set are shown in Table 1. Confusion matrices for the training and the test sets are shown in Tables 2 and 3, respectively. The evaluation of the

performance of the training set yielded: precision = 80 %, sensitivity = 100 %, specificity = 73.3 %, and accuracy = 87.1 %. Regarding the test set, the respective parameters are: precision = 87.5 %, sensitivity = 87.5 %, specificity = 80 %, and accuracy = 84.6 %.

The applicability domain was defined for NPs within the test set and the cutoff value was estimated to be 0.906. For all structures of the test set, values range from 0 to 0.448, therefore, all predictions are considered reliable. The above set of validation measurements highlights the accuracy, significance, and robustness of the produced model.

The predictive workflow is publicly available via the Enalos Cloud platform. In silico design and screening may be performed through the Enalos Cloud platform by visiting the iron oxide model web page (enalos.insilicotox.com/QNAR_IronOxide_Toxicity/). The user can initiate a prediction by either manually entering the selected NP properties (e.g., zeta potential, size etc.) or by importing a CSV file (.csv) with NP properties for high-throughput virtual screening (HTVS) (Scheme 1).

When properties are uploaded for a set of NPs and the input values have been included, the predictive model is used and a prediction is obtained. The generated

 **Enalos QNAR Iron Oxide Toxicity Platform**

MNP Number	Size (nm)	ZP (mV)	R1 (mM-IS-1)	R2 (mM-IS-1)	Coating
1					Other ▼
2					Other ▼
3					Other ▼
4					Other ▼
5					Other ▼
6					Other ▼
7					Other ▼
8					Other ▼
9					Other ▼
10					Other ▼
11					Other ▼
12					Other ▼
13					Other ▼
14					Other ▼
15					Other ▼
16					Other ▼
17					Other ▼
18					Other ▼
19					Other ▼
20					Other ▼

Import a CSV file for High Throughput Virtual Screening (.csv)

No file selected.

Scheme 1 Screenshot of Enalos iron oxide platform input page

Enalos QNAR Iron Oxide Toxicity Platform

Knime report powered by Birt

"Prediction"	"Domain"
inactive	reliable
inactive	reliable
inactive	reliable
inactive	reliable
inactive	reliable
inactive	reliable
inactive	reliable
active	reliable
active	reliable
active	reliable
active	reliable
active	reliable
active	reliable
inactive	reliable
inactive	reliable

Date: Dec 20, 2014 10:07 AM
www.knime.org

Author: NovaMechanics Ltd

1 of 1

Scheme 2 Screenshot of Enalos QNAR iron oxide toxicity platform results

output provides a summary of the results in a pdf-like format or a CSV file, which contains all relevant information for further analysis (Scheme 2). The results include predicted values for each nanostructure and a notification on the reliability of predictions based on the domain of applicability limits.

As already mentioned, the Enalos Cloud platform allows the performance of a preliminary *in silico* testing by virtually screening a set of NPs, based on the validated model. The predictive QNAR model described here can be accessed at the web page of Enalos QNAR Iron Oxide Toxicity Platform (http://enalos.insilicotox.com/QNAR_IronOxide_Toxicity/).

This study outlines Enalos InSilico platform as a useful tool to facilitate the computer-aided NP design by acting as a source of toxicity prediction for novel NPs.

Cellular Uptake Prediction and Virtual Screening of Nanoparticles via Enalos InSilicoNano Platform

In this section, a validated quantitative nanostructure–activity relationships (QNAR) model that can predict the cellular uptake of organic nanoparticles is presented (Melagraki and Afantitis 2014). The model is publicly available through Enalos InSilico Cloud platform in the QNAR_PaCa2 web page (http://enalos.insilicotox.com/QNAR_PaCa2/) and can be used for new-structure predictions that are designed and/or uploaded to the server. The Enalos InSilico web service functionality was successfully used for the virtual screening of a set of PubChem structures, which were selected to recognize structures similar to that of a known active compound. The model was based on Mold2 descriptors and the k-nearest neighbors (kNN) algorithm.

The selected engineered NPs (ENPs) used for model development, possess the same metal core but different organic coating (Weissleder et al. 2005). The model building was based on a KNIME workflow, especially designed for this purpose. Initially, for the development of the model, all data including organic molecules and cellular uptake values were preprocessed and randomly divided into training set (89 compounds, which were used in the model development) and validation (25 compounds) set. The Enalos Mold2 KNIME node was used for the calculation of 777 descriptors for each compound. During the correlation analysis some descriptors were eliminated, thus leaving 382 of them to be used as inputs for the QNAR model development.

Next, the CfsSubset variable selection with BestFirst evaluator method was applied to identify nine descriptors as the most representatives of the structural features that define the biological profile of the studied NPs.

The nine descriptors are: (D461) Geary topological structure autocorrelation length-7 weighted by atomic van der Waals volume, (D467) Geary topological structure autocorrelation length-5 weighted by atomic Sanderson electronegativities, (D599) number of total quaternary C-sp³, (D649) number of group secondary aliphatic amines, (D712) number of group donor atoms for hydrogen bonds, (D714) number of group CH₃R and CH₄, (D753) number of group phenol or enol or carboxyl OH, (D758) number of group Al₂-NH, and (D775) hydrophilic factor index. The physical meaning of the above descriptors can be found in the original publication (Melagraki and Afantitis 2014). The optimized value of k within kNN application was 2 (Franco-Lopez et al. 2001).

The proposed model was successfully validated with the methodology applied in the previous section. The validation results are shown in Fig. 2.

The R²_{LOO} was calculated to be 0.74. Also, the Y-randomization test verified the model's robustness and statistical significance. The decreased values of the correlation coefficient indicate the low possibility of chance correlation.

After the model was validated, the reliability of a given prediction was suggested through domain of applicability calculations (cutoff value = 2.153) (Mouchlis et al.

Criterion	Assessment	Result
$R^2 > 0.6$	PASS	$R^2 = 0.848$
$R_{cvext}^2 > 0.5$	PASS	$R_{cvext}^2 = 0.82$
$(R^2 - R_0^2) / R^2 < 0.1$	PASS	$(R^2 - R_0^2) / R^2 = 0.038$
$(R^2 - R'0^2) / R^2 < 0.1$	PASS	$(R^2 - R'0^2) / R^2 = 0.0$
$abs(R_0^2 - R'0^2) < 0.1$	PASS	$abs(R_0^2 - R'0^2) = 0.032$
$0.85 < k < 1.15$	PASS	$k = 1.019$
$0.85 < k' < 1.15$	PASS	$k' = 0.979$

Model Predictive

Fig. 2 Model evaluation summary results

2012; Zhang et al. 1995; Papa et al. 2009). It was concluded that the proposed model requires only the structural information from the organic compounds involved and was confirmed to be accurate and reliable within applicability limits. Thus, the model could be considered a useful tool toward cellular uptake determination of NPs.

The model was made available online through Enalos InSilico platform for fast in silico predictions for a set of given compounds. Screenshots of the Enalos InSilico web service and the results page are presented in Schemes 3 and 4.

Name	Description
QNAR_PaCa2	QNAR model correlating chemical descriptors and MNP cellular uptakes (similar nanoparticle core with different surface modifiers)

Scheme 3 Screenshot of Enalos InSilico platform input page for QNAR_PaCa2 model

Prediction of MNPs Uptake in PaCa2 Cancer Cells

Knime report powered by Birt

"PaCa2 cellular uptake (log ₁₀ [nanoparticles]/cell pM)"	"Domain of Applicability Prediction"
4.429	reliable
4.293	reliable
3.494	reliable
4.197	reliable
3.895	reliable
3.895	reliable
4.197	reliable
4.197	reliable
3.895	reliable
4.197	reliable
3.523	reliable
3.495	reliable
4.209	reliable
3.895	reliable
4.283	reliable

Date: Jun 5, 2014 6:07 AM
www.knime.org

Author: NovaMechanics Ltd

1 of 1

Scheme 4 Screenshot of Enalos InSilico platform results for QNAR_PaCa2 model

As shown in Schemes 3 and 4, the user can design or enter a chemical structure and obtain a prediction. The aforementioned workflow will calculate the descriptors and the output will be rapidly generated (within seconds). One may experiment with different scaffolds and structures and observe the structural features that induce a certain effect. Also, the user can exploit the proposed QNAR model and then scan specific structures for a preliminary *in silico* testing. In this way, it is possible to offer QSAR/QNAR results for immediate sharing and implementation. It was recently pointed out (Tetko 2012) that the use of predictive models as software tools will probably increase in the future, and this will activate the reuse of knowledge, which in turn will result in further developments.

As already mentioned, the proposed model and web platform can be used in virtual screening studies for the prioritization of new compounds. To demonstrate the usefulness of the produced model, potent compounds from the PubChem database were identified using similarity calculations based on molecular quantum numbers (MQNs) (Melagraki and Afantitis 2014). The virtual screening procedure was employed for the recognition of the first 1000 neighbors of **compound 36** (isatoic anhydride), in terms of chemical similarity. The 1000 resulting compounds were tested with the online tool through Enalos InSilico platform (via an sdf file, which contains all the structures) regarding their cellular uptake. Compounds were next classified by increasing potency and the most promising ones were selected for screening. The predictions for the first 20 compounds are shown in Table 4.

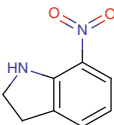
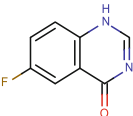
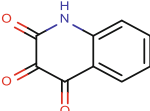
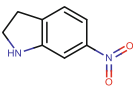
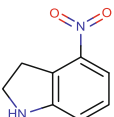
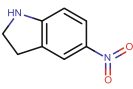
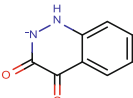
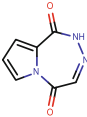
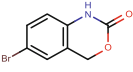
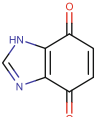
Within the proposed framework, the Enalos InSilico platform emerges as an invaluable application for the evaluation of novel NPs, which have not been experimentally tested or synthesized. An additional, important aspect of the approach is that the above tools can be further expanded and applied to polymer–NP structures that are currently gaining increasing attention.

Identification of Organic Materials as Corrosion Inhibitors Based on Enalos KNIME Nodes

One of the most efficient ways to prevent metal corrosion in acidic media is the development of novel corrosion inhibitors (Ebenso et al. 2012). Organic inhibitors, which contain heteroatoms (e.g., oxygen, nitrogen, sulfur) and possess multiple bonds, have been considered for various corrosion systems, metals, and alloys. Inhibition is obtained with creation of physical and/or chemical absorption film on the surface of the metal (El Ashry et al. 2012). The planarity of heterocycles and the presence of lone electron pairs on the heteroatoms are crucial factors that control the absorption of these compounds on the metallic surface.

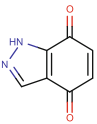
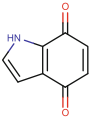
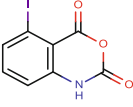
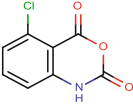
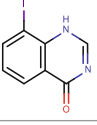
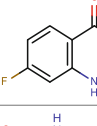
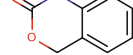
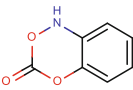
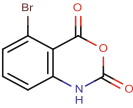
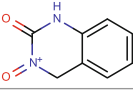
One drawback regarding the design of corrosion inhibitors is the time-consuming and costly nature of the process. On the other hand, computational techniques, such as the quantitative structure–property relationships (QSPR) methodology has greatly advanced the efficient modeling and prediction of new or modified corrosion inhibitors (Lee et al. 2012; Toropov et al. 2012).

Table 4 Virtual screening results for the most promising compounds in PubChem database

ID	Compound	Predicted value PaCa2 cellular uptake (log ₁₀ [NP]/cell)	Domain of applicability (limit: 2.153)
679		4.41	0.03
604		4.41	0.01
958		4.41	0.06
676		4.40	0.01
678		4.40	0.05
677		4.39	0.02
107		4.39	0.02
368		4.38	0.10
293		4.37	0.09
493		4.37	0.10

(continued)

Table 4 (continued)

ID	Compound	Predicted value PaCa2 cellular uptake (log ₁₀ [NP]/cell)	Domain of applicability (limit: 2.153)
494		4.37	0.10
550		4.36	0.11
196		4.35	0.06
200		4.35	0.06
626		4.35	0.06
602		4.35	0.06
981		4.34	0.10
925		4.34	0.10
192		4.34	0.06
65		4.34	0.05

In this study, the modeling and prediction of corrosion inhibition for steel in acidic environment through the development of QSPR with the aid of the Enalos KNIME nodes is described (Berthold et al. 2008; Melagraki and Afantitis 2013). The development of a predictive kNN model was realized by first calculating Mold2 molecular descriptors for the organic inhibitors with the Enalos Mold2 KNIME node. The predictive model was assessed with the Enalos Model Acceptability Criteria KNIME node. The domain of model applicability was determined using the Enalos Domain KNIME nodes.

Corrosion inhibition data for steel in acidic medium from various organic chemicals were collected from the literature (El Ashry et al. 2012) and were assembled in a single database. Inhibitors involve triazole, oxadiazole, and thiadiazole derivatives; aromatic hydrazides and Schiff bases; benzimidazole and 2- substituted derivatives; as well as pyridine derivatives. A set of 55 organic inhibitors in different concentrations yielded a total of 186 inhibition data.

The structural features of the studied corrosion inhibitors were evaluated with the Mold2 software. A total of 777 descriptors were initially calculated for each compound based on topological, geometrical, and structural criteria. From these, only 320 descriptors were used as possible inputs for the construction of the QSPR model, since the remaining descriptors were filtered out due to poor discrimination power (Ojha and Roy 2011).

Among the different modeling methodologies screened in KNIME platform, the k-nearest neighbors (kNN) technique (with an optimized value $k = 3$) was selected as the most appropriate for the specific data (Hall et al. 2009). Details on the kNN algorithm and methodology can be found elsewhere (Franco-Lopez et al. 2001).

The predictive model was validated internally and externally according to the standards of QSAR model acceptance, as imposed by the OECD. The complete dataset was randomly split into 70:30 ratio (training set: validation set) with the partitioning KNIME node. The combinations in the test set did not participate in the training procedure. The statistical criteria that were used to determine the robustness, reliability, and predictive ability of the model are: the coefficient of determination between experimental values and predicted values (R^2), validation via external test set, leave-one-out cross validation procedure and quality of fit and predictive ability of a continuous QSAR model based on Tropsha's tests (Tropsha 2010). The statistics of the validation procedure are shown in Table 5.

Table 5 Statistical parameters of the QSPR model

R^2 training (n = 131)	0.96
RMSE training	4.90
R^2_{LOO}	0.73
R^2_{pred} (n=55)	0.84
RMSE _{pred}	9.83

The predictive scheme included a KNIME workflow, which operated the following actions:

1. Compounds along with corrosion inhibition data were uploaded and preprocessed.
2. The calculation and selection of descriptors was realized.
3. The kNN methodology was carried out.
4. The produced model was validated.
5. The domain of applicability was determined.

The initial dataset of 186 corrosion inhibitors was randomly divided into 131 training set compounds and 55 validation set compounds (ratio 70:30). As mentioned above, only compounds from the training set were used to develop the QSPR models, and 320 descriptors were selected as possible inputs during the development.

The CfsSubset variable selection with the BestFirst evaluator method (Witten et al. 2005) was next applied on the training set to identify the most significant among the 320 descriptors. Thus, the concentration along with seven descriptors was selected as the most important parameter for the model development.

The selected descriptors include number of oxygen (D026), structural information content order-1 index (D282), Geary topological structure autocorrelation length-8 weighted by atomic Sanderson electronegativities (D470), Moran topological structure autocorrelation length-7 weighted by atomic polarizabilities (D509), lowest eigenvalue from Burden matrix weighted by van der Waals order-6 (D545), highest eigenvalue from Burden matrix weighted by van der Waals order-4 (D575), number of group Ar-CH=X (D741), and the concentration (C in mM).

The description of the selected descriptors will be briefly discussed below (Todeschini and Consonni 2009).

A combination of the descriptors offers a unified representation of the compound and high selective power. Descriptors D026 and D741 indicate the number of oxygen atoms and Ar-CH=X groups that may be present in the compound. Descriptor D282 encodes the structural information content order-1 index. This descriptor represents a graph theoretical invariant, which considers the molecular graph as being a source of different probability distributions, to which the information theory is applied (Todeschini and Consonni 2009). D470 encodes information as described by Geary topological structure autocorrelation length-8 weighted by atomic Sanderson electronegativities. The Geary index denotes spatial autocorrelation and is a distance-type function which varies from zero to infinite. D509 encodes information regarding the atomic polarizabilities combined with Moran topological structure autocorrelation length-7. Moran coefficient (value range between -1 and +1) indicates the spatial autocorrelation and is associated with atomic properties, the number of atoms, and the topological distance between them. Descriptors D545 and D575 represent the lowest eigenvalue from Burden matrix (Burden 1989) weighted by van der Waals order-6 and the highest eigenvalue from Burden matrix weighted by van der Waals order-4, respectively (Burden et al. 2009). Burden descriptors

weighted by van der Waals have been shown to be very selective descriptors and in turn useful for similarity searching (Todeschini and Consonni 2009).

The above descriptors are associated with different weights that affect the corrosion inhibition across compounds. Therefore, according to the positive or negative impact of each descriptor, novel compounds with specified properties may be designed.

After comparison with the results from other methodologies (Melagraki and Afantitis 2013), it was concluded that the kNN approach yielded an accurate and powerful model that reliably predicts the efficiency of corrosion inhibition. Then, one may safely conclude that the selected descriptors encode the structural characteristics of the substances related to corrosion inhibition.

The experimental vs. predicted corrosion inhibition values for the training set and test set compounds are shown in Fig. 3. Outliers have been indicated and presented in the original article (Melagraki and Afantitis 2013).

The Enalos Model Acceptability Criteria KNIME node has been applied to the data (Fig. 4) and the model passed Tropsha's requirements for predictive ability.

R^2 is the determination coefficient between experimental and predicted values and model prediction on the test set (R^2_{pred}). The model was particularly stable with respect to the inclusion/exclusion of compounds measured by the leave-one-out (LOO) cross validation procedure ($R^2_{\text{LOO}} = 0.73$).

Another measure of robustness and statistical significance of a QSPR model is the Y-randomization test, which further validated our approach. An additional validation test was conducted to evaluate the predictive power of the method independently of the partitioning of the dataset (Melagraki and Afantitis 2013).

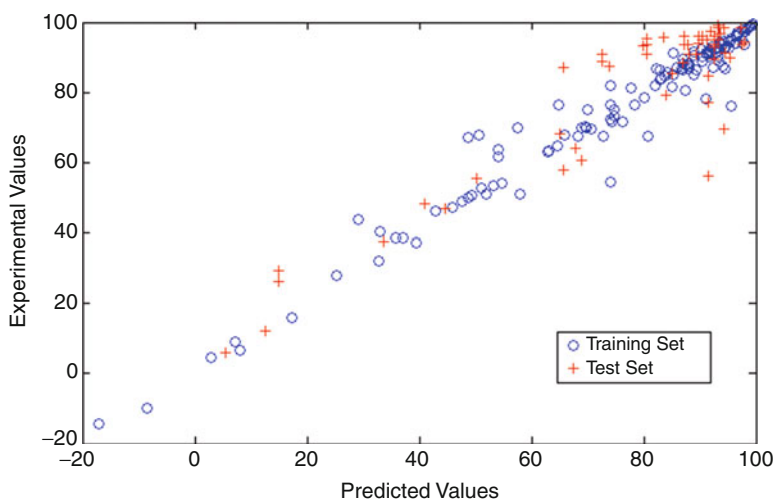


Fig. 3 Experimental vs. predicted values for the training and test set

Criterion	Assessment	Result
$R^2 > 0.6$	PASS	$R^2 = 0.842$
$R_{cv}^2 > 0.5$	PASS	$R_{cv}^2 = 0.826$
$(R^2 - R_0^2) / R^2 < 0.1$	PASS	$(R^2 - R_0^2) / R^2 = 0.0010$
$(R^2 - R'0^2) / R^2 < 0.1$	PASS	$(R^2 - R'0^2) / R^2 = 0.019$
$abs(R_0^2 - R'0^2) < 0.1$	PASS	$abs(R_0^2 - R'0^2) = 0.014$
$0.85 < k < 1.15$	PASS	$k = 1.023$
$0.85 < k' < 1.15$	PASS	$k' = 0.964$

Model Predictive

Fig. 4 Enalos Model Acceptability Criteria KNIME node screenshot

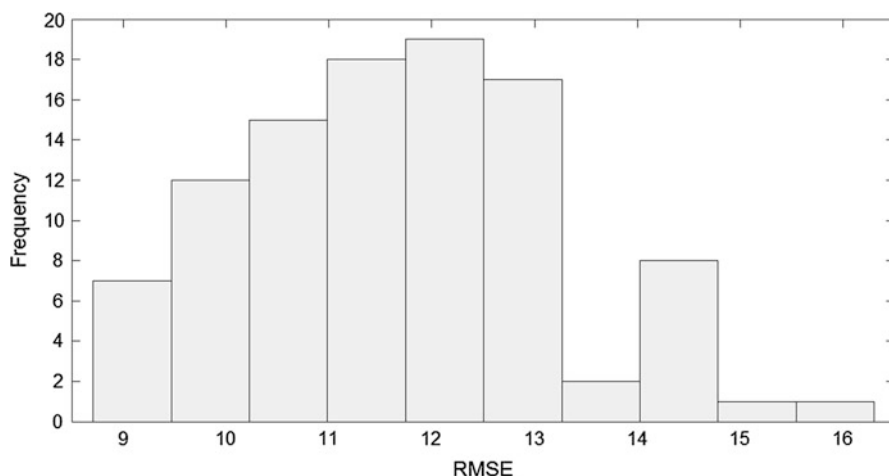


Fig. 5 Distribution of the RMSE values (100 random splits)

The distribution of the root mean squared error (RMSE) values is shown in Fig. 5.

The applicability domain cutoff value was estimated to be 3.774 and 0.183 for similarity (Mouchlis et al. 2012) and leverage (Afantitis et al. 2008) calculations, respectively. Similarity calculations for all compounds in the test set had values which range between 0.015 and 1.23. However, the leverage predicted response of a simple pyridine (0.378) resulted from a significant model extrapolation, and it is the only prediction that may be considered unreliable.

The present approach, due to its high predictive power and the minimal requirement of only 2D structure information of a compound, could be a very useful tool for the determination of the corrosion inhibition. Moreover, this modeling method considerably decreases the time and cost required to experimentally design corrosion inhibitors. Also, the method may be applied to the screening of regular or virtual chemical databases, thus seeking new organic compounds with specific

properties. For this purpose, the applicability domain will be an invaluable tool for discarding “divergent” chemical structures.

KNIME-Based Design and Prediction of Compound Structures

The broad applicability of KNIME in the area of medicinal and pharmaceutical chemistry has also rendered it a valuable tool for structure prediction and analysis. Thus, KNIME workflows have been implemented over the years to perform drug design, virtual screening, molecular modeling, QSAR, QSPR, structure classification, and clustering applications among others.

In this section, we present two KNIME applications regarding the prediction of yellow fever inhibitors and a knowledge-based method for the de novo design of novel compounds.

Prediction of Yellow Fever Inhibitors from ChEMBL Database Through KNIME Classification Analysis

Yellow fever (YF) is an acute infection, which is transmitted by arthropods and mosquitoes to humans (Agnihotri et al. 2012), and is caused by the mosquito-borne yellow fever flavivirus (YFV). The family of flaviviruses also includes other RNA viruses, such as the hepatitis C virus (HCV), the dengue virus (DENV), the West Nile virus (WNV), the Japanese encephalitis virus (JEV), and the bovine viral diarrhea virus (BVDV) among others (Agnihotri et al. 2012; Julander 2013).

YFV is a major health risk in particular regions of South America and Africa since almost 200,000 new infections and 30,000 deaths are observed every year (Chatelain et al. 2013); importantly, the fatality rate may reach 60 % in severe cases.

Currently, several regular antiviral drugs have been tested against YF disease, but no chemotherapeutic medication has been developed to specifically target YFV. An anti-YFV vaccine (17D) is used to prevent the infection; however, it has been observed to cause systemic infections and side effects in some patients (Julander 2013).

Therefore, computational approaches, such as virtual screening and modeling are particularly suitable to assist the discovery of new anti-YFV compounds. In this context, Moorthy and Poongavanam collected a number of compounds from the literature that were found (experimentally or computationally) to inhibit the YFV and employed them for the development of KNIME classification models using the Naive Bayes approach (Narayana Moorthy and Poongavanam 2015).

For this purpose, a set of 379 YFV inhibitors were collected from the ChEMBL database (<https://www.ebi.ac.uk/chembl/>). After the initial treatment of the data (namely, salt removal, 3D structure generation, and energy minimization of the structures), 30 two-dimensional descriptors of the compounds were calculated with the CDK tool as implemented in KNIME (Berthold et al. 2008; Beisken et al. 2013). In total, 16 classification models were developed using the Weka

data mining software (Hall et al. 2009). The first 12 models refer to individual datasets, while models 13–16 were developed from a combined dataset, which contained 309 compounds. Before the development of the models, the dataset was partitioned into training (65 %) and test (35 %) set based on a sampling consideration that distributes the inhibitors and noninhibitors evenly between sets. The inhibitor definition was restricted to specific activity criteria (i.e., $IC_{50} \leq 10 \mu\text{M}$ for inhibitors; $IC_{50} > 10 \mu\text{M}$ for noninhibitors). The quality of models was investigated through various activity thresholds (10, 30, 50, and 100 μM) and statistical parameters, such as the sensitivity, specificity, G-mean, Matthew's correlation coefficient (MCC), and overall accuracy.

Principal component analysis (PCA) revealed that the dataset does not contain distinct clusters and its diversity is adequately represented by the training set. Despite the presence of some outliers, it was observed that they were not structurally similar whatsoever. Additionally, it was shown that the majority of inhibitors is highly affected by the topological polar surface area (TPSA) and polar bonds. This indicated the more hydrophobic nature of noninhibitors compared to inhibitors.

The BestFirst attribute selection module of Weka was employed to select a set of 24 physicochemical descriptors in order to construct the Naive Bayes classification models for YFV inhibition. It was shown that the six datasets (activity cutoff 30 μM) perform equally well with an overall accuracy for the test set being >75 %.

The model quality (based on MCC) was higher for all datasets except for one, which performed poorly. Models derived from all datasets were statistically significant with MCC and G-mean values >0.7. The significance of the dataset was also verified by the high F-score values (>0.8), while sensitivity and specificity parameters were >0.75 for most models. The F-score is calculated through the following equation:

$$\text{F-score} = \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (1)$$

Also, the predictive ability of the scheme was investigated by combining all the datasets into one. Thus, 70 compounds (Krečmerová et al. 2007) were used in the test set, while the remaining 309 compounds were combined into one dataset to train the model. It was observed that an activity cutoff of 50 μM yields a relatively good accuracy. Despite that models developed by different activity thresholds (10 or 30 μM) were successful in distinguishing inhibitors (>90 %) from noninhibitors (65 %), they lacked a balanced class distribution (predicting adequately both classes instead of predicting accurately only one), which clearly affects the quality of the model. Therefore, new models were constructed based on the 309 compounds set (200 compounds in the training set, 109 compounds in the test set). From the models for each activity threshold (10, 30, 50, and 100 μM), one model (at 50 μM) displayed superior performance over the others. That is, it predicted 92 % of inhibitors and 78 % of noninhibitors with a MCC coefficient of 0.71, and a G-mean score (0.84), which reflects the high quality. Finally, the best YFV inhibition model

(at 50 μM) was implemented into KNIME workflow to become freely available for use in medicinal chemistry applications.

De Novo Design of Synthetically Feasible Compounds Through the Reaction Vectors Approach

Early attempts to develop successful de novo design programs for compounds have been seriously hampered by the unfeasible synthesis of the proposed structures (Boda and Johnson 2006; Lewis and Leach 1994; Schneider and Fechner 2005; Gillet and Johnson 1998). More recent approaches that may facilitate synthetic routes generate a fixed set of transformations, which are in turn applied to starting structures to generate new molecules (Fechner and Schneider 2006, 2007; Lameijer et al. 2006; Schürer et al. 2005; Vinkers et al. 2003). However, these methods are usually restricted by the limited number of “reactions” that can be performed.

Reaction-like methodologies have been used in QSAR and data mining studies. For instance, the “matched molecular pairs” scheme (Leach et al. 2006) evaluates the change of a property with respect to a single structural change. Similar approaches perform clustering of molecule pairs according to descriptor difference vectors, which are constructed by subtracting the vector representation of the “product” molecule from the vector representation of the “reactant” molecule (Sheridan et al. 2006). Therefore, the pairs of molecules which belong to the same cluster represent similar transformations. Thus, changes in activities associated with the pairs may be used to estimate the effect of a particular transformation on activity.

A knowledge-based approach regarding the de novo design of compounds based on reaction vectors has been proposed by Patel et al. (Patel et al. 2009). The method of reaction vectors is related to the descriptor difference vectors approach (Sheridan et al. 2006) and characterizes the structural changes, which occur at the reaction center.

The authors’ flexible procedure enabled the automatic collection of ever-increasing data on reactions, which are available in several databases. Reaction vectors could be applied to novel starting molecules through a structure generation algorithm to predict new structures. Each applied transformation is derived from a reactions database, and since it is associated with a reaction from the literature, one may be confident toward the synthetic feasibility of the proposed molecules. The vectors are automatically collected from a set of reactions, which is not restricted by size or reaction type, therefore without involving any complex reaction strategy.

The principle of the method is the utilization of the information taken from a set of reactions (parent reactions) in deriving reaction vectors. Reaction vectors are used to describe the reaction environment as well as changes at the reaction center. Then, the reaction vectors are extracted from the knowledge base and may be applied to input starting materials in order to predict new product compounds for synthesis. In some cases, successful predictions may require the operation of a reaction vector to a second starting molecule.

In the past years, reaction vectors have been used to search and classify chemical reactions (Broughton et al. 2003). As already mentioned, a reaction vector represents the changes that occur during a reaction. This is described in a descriptor scheme that is enhanced in the products (positive descriptors, they represent bonds gained in products) and is diminished from the reactants (negative descriptors, they represent bonds lost from reactants). In de novo design, the practical role of reaction vectors is to propose synthetic ways to new compounds by achieving a balance between specificity and generality. Namely, the reaction vectors should be specific enough in order to avoid their application in environments which disfavor the course of a reaction, but encoding part of the environment will hamper their ability to generate novel structures. It was shown that such a balance may be obtained by combining atom pairs separated by one and two bonds (Patel et al. 2007). The reaction vectors employed by Patel et al. are modifications of the original descriptors developed by Carhart et al. (1985). More detailed information on reaction vectors can be found in Patel et al. and references therein (Patel et al. 2009).

The authors revealed that many reactions in the Lilly database are incomplete, and they implemented a reaction cleaning algorithm to overcome this drawback. The reaction cleaning algorithm is applied to the reactions before calculating the atom pair descriptors for the reactants and products. The reaction vector is next calculated as follows:

$$D = P - R \quad (2)$$

where D is the reaction vector, P is the product vector (the sum of the vectors of individual products), and R is the reactant vector (the sum of the vectors of individual reactants). The algorithms have been implemented with the aid of the JoeLib toolkit (joelib.sourceforge.net/) with the de novo design tool available in KNIME.

The de novo design algorithm was validated internally by reproducing reactions in the Lilly database and externally by reproducing the already known synthetic pathways of two drugs. Internal validation was achieved by generating 90 % of the known parent products for each reaction in a dataset containing 5695 reactions and 2866 reaction vectors. The external validation involved the syntheses of the intermediate of the antithrombotic drug (S)-(+)-clopidogrel bisulfate (Wang et al. 2007) and the antidepressant venlafaxine (Kavitha and Rangappa 2004). In each case, the products were successfully generated using a 5839 reagents set and a dataset, which comprised 24,418 reactions and 16,859 reaction vectors. The reaction vectors procedure was successful in reproducing the known synthetic routes for both drugs; however, it was suggested that this approach also yields quite a few alternative products.

The applicability of the algorithm to de novo design was demonstrated in three cases: (1) the ability of a reaction vector to generate novel and diverse products as regards the parent reaction, (2) prediction of analogs of a lead compound, and (3) application to the enumeration of a compound library.

In the first case, 10 reactants were randomly selected from the smaller (2866 vectors/5695 reactions) dataset and were input to the de novo design tool. Products were generated employing the 16,859 reaction vectors from the larger dataset. Structures containing more functional groups than others usually yielded more solutions, therefore, the number of products that were created lies within 0 and 44. Moreover, the average similarities between predicted compounds and products of the parent reactions vary significantly (0.15–0.96). It is therefore suggested that the diversity of the products depends on both the starting molecule and the entries in the reaction database.

Next, the predictive power of the algorithm in lead optimization cases was explored by considering drugs as starting compounds and generating products after single-step transformations. The products were assumed to be synthetically feasible if they were identified upon search in SciFinder Scholar. The procedure was based on penicillin G, Prozac, and aspirin (starting structures) and was performed with the larger set of reaction vectors (16,859) and the 5839 reagent set as described above. Penicillin G was associated with the most potential products (24), while 20 and 12 products were generated for Prozac and aspirin, respectively. The reaction vector methodology was shown to realize the generation of synthetically accessible molecules that share close structural similarities with a parent compound. Thus, this approach could be particularly valuable in structure-activity relationship (SAR) studies by proposing structures for synthesis that explore sufficiently the conformational space around a known lead compound.

Finally, the proposed methodology can be used to construct a virtual library of products by listing all possible products from a single reaction employing a specific set of reagents. For instance, the authors used 6-bromoquinoxalin-2-one as starting material, 628 boronic acids as the reagents (extracted from the ACD), and a Suzuki coupling reaction to construct an enumerated library with 292 compounds. It was shown that this scheme may be used to direct a parallel synthesis route according to a single reaction with varying reagents.

Moreover, the way the reaction vector methodology recognizes the environment of a reaction is clearly demonstrated. For the transformation that is applied, there is a literature precedence, and thus a high level of confidence is achieved regarding the completion of the reaction. However, this requires that complete coverage of a generic reaction has been obtained through adequate examples in the reaction database.

The approach has been implemented in KNIME and this allows the facile customization of the procedure for a variety of applications. Overall, the usefulness of the proposed algorithm is to provide sets of molecules, which are associated with multiple purposes, and at the same time being synthetically feasible.

Acknowledgments This work is supported by funding under the Seven Research Framework Programme of the European Union: Project THALAMOSS (HEALTH.2012.1.2-1 grant agreement no. 306201) and Project nanoMILE (grant agreement no. 310451).

Bibliography

- Afantitis, A., Melagraki, G., Sarimveis, H., Koutentis, P. A., Markopoulos, J., & Igglessi-Markopoulou, O. (2008). Development and evaluation of a QSPR model for the prediction of diamagnetic susceptibility. *QSAR & Combinatorial Science*, 27(4), 432–436.
- Afantitis, A., Melagraki, G., Koutentis, P. A., Sarimveis, H., & Kollias, G. (2011). Ligand – based virtual screening procedure for the prediction and the identification of novel β -amyloid aggregation inhibitors using Kohonen Maps and Counterpropagation Artificial Neural Networks. *European Journal of Medicinal Chemistry*, 46, 497–508.
- Agnihotri, S., Narula, R., Joshi, K., Rana, S., & Singh, M. (2012). In silico modeling of ligand molecule for non structural 3 (NS3) protein target of flaviviruses. *Bioinformatics*, 8(3), 123–127.
- Beisken, S., Meinel, T., Wiswedel, B., de Figueiredo, L. F., Berthold, M., & Steinbeck, C. (2013). KNIME-CDK: Workflow-driven cheminformatics. *BMC Bioinformatics*, 14, 257–257.
- Berthold, M., Cebron, N., Dill, F., Gabriel, T., Kötter, T., Meinel, T., Ohl, P., Sieb, C., Thiel, K., & Wiswedel, B. (2008). KNIME: The Konstanz information miner. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, & R. Decker (Eds.), *Data analysis, machine learning and applications* (pp. 319–326). Berlin/Heidelberg: Springer.
- Boda, K., & Johnson, A. P. (2006). Molecular complexity analysis of de novo designed ligands. *Journal of Medicinal Chemistry*, 49(20), 5869–5879.
- Broughton, H., Hunt, P., & MacKey, M. (2003) Methods for classifying and searching chemical reactions. Google Patents.
- Brown, F. K. (1998). Chemoinformatics, what it is and how does it impact drug discovery. *Annual Reports in Medicinal Chemistry*, 33, 375–384.
- Burden, F. (1989). Molecular identification number for substructure searches. *Journal of Chemical Information and Computer Sciences*, 29, 225–227.
- Burden, F., Polley, M., & Winkler, D. (2009). Toward novel universal descriptors: Charge fingerprints. *Journal of Chemical Information and Modeling*, 49, 710–715.
- Carhart, R. E., Smith, D. H., & Venkataraghavan, R. (1985). Atom pairs as molecular features in structure-activity studies: Definition and applications. *Journal of Chemical Information and Computer Sciences*, 25(2), 64–73.
- Chatelain, G., Debing, Y., Burghgraef, T. D., Zmurko, J., Saudi, M., Rozenski, J., Neyts, J., & Aerschot, A. V. (2013). In search of flavivirus inhibitors: Evaluation of different tritylated nucleoside analogues. *European Journal of Medicinal Chemistry*, 65, 249–255.
- Cohen, Y., Rallo, R., Liu, R., & Liu, H. H. (2013). In silico analysis of nanomaterials hazard and risk. *Accounts of Chemical Research*, 46(3), 802–812.
- Ebenso, E. E., Kabanda, M. M., Murulana, L. C., Singh, A. K., & Shukla, S. K. (2012). Electrochemical and quantum chemical investigation of some azine and thiazine dyes as potential corrosion inhibitors for mild steel in hydrochloric acid solution. *Industrial and Engineering Chemistry Research*, 51, 12940–12958.
- El Ashry, E. S. H., El Nemer, A., & Ragab, S. (2012). Quantitative structure activity relationships of some pyridine derivatives as corrosion inhibitors of steel in acidic medium. *Journal of Molecular Modeling*, 18, 1173–1188.
- Epa, V. C., Burden, F. R., Tassa, C., Weissleder, R., Shaw, S., & Winkler, D. A. (2012). Modeling biological activities of nanoparticles. *Nano Letters*, 12(11), 5808–5812.
- Fechner, U., & Schneider, G. (2006). Flux (1): A virtual synthesis scheme for fragment-based de novo design. *Journal of Chemical Information and Modeling*, 46(2), 699–707.
- Fechner, U., & Schneider, G. (2007). Flux (2): Comparison of molecular mutation and crossover operators for ligand-based de novo design. *Journal of Chemical Information and Modeling*, 47(2), 656–667.
- Fourches, D., Pu, D., Tassa, C., Weissleder, R., Shaw, S. Y., Mumper, R. J., & Tropsha, A. (2010). Quantitative nanostructure–activity relationship modeling. *ACS Nano*, 4(10), 5703–5712.

- Franco-Lopez, H., Ek, A. R., & Bauer, M. E. (2001). Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. *Remote Sensing of Environment*, 77, 251–274.
- Gajewicz, A., Rasulev, B., Dinadayalane, T. C., Urbaszek, P., Puzyn, T., Leszczynska, D., & Leszczynski, J. (2012). Advancing risk assessment of engineered nanomaterials: Application of computational approaches. *Advanced Drug Delivery Reviews*, 64(15), 1663–1693.
- Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., & Overington, J. P. (2012). ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Research*, 40(Database issue), D1100–D1107.
- Gillet, V. J., & Johnson, A. P. (1998). *Structure generation for De Novo design*. Washington: American Chemical Society.
- Gutlein, M., Karwath, A., & Kramer, S. (2012). CheS-mapper – chemical space mapping and visualization in 3D. *Journal of Cheminformatics*, 4(1), 7.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *SIGKDD Explorations*, 11(1), 10–18.
- Hong, H., Xie, Q., Ge, W., Qian, F., Fang, H., Shi, L., Su, Z., Perkins, R., & Tong, W. (2008). Mold2, molecular descriptors from 2D structures for chemoinformatics and toxicoinformatics. *Journal of Chemical Information and Modeling*, 48, 1337–1344.
- Hu, Y., & Bajorath, J. (2012). Freely available compound data sets and software tools for chemoinformatics and computational medicinal chemistry applications [version 1; referees: 2 approved]. Vol. 1.
- Irwin, J. J., Sterling, T., Mysinger, M. M., Bolstad, E. S., & Coleman, R. G. (2012). ZINC: A free tool to discover chemistry for biology. *Journal of Chemical Information and Modeling*, 52(7), 1757–1768.
- Julander, J. G. (2013). Experimental therapies for yellow fever. *Antiviral Research*, 97(2), 169–179.
- Kavitha, B. C. V., & Rangappa, K. S. (2004). Simple and an efficient method for the synthesis of 1-[2-dimethylamino-1-(4-methoxy-phenyl)-ethyl]-cyclohexanol hydrochloride: (±) venlafaxine racemic mixtures. *Bioorganic & Medicinal Chemistry Letters*, 14(12), 3279–3281.
- Kleandrova, V. V., Luan, F., Gonzalez-Diaz, H., Ruso, J. M., Melo, A., Speck-Planche, A., & Cordeiro, M. N. (2014a). Computational ecotoxicology: Simultaneous prediction of ecotoxic effects of nanoparticles under different experimental conditions. *Environment International*, 73, 288–294.
- Kleandrova, V. V., Luan, F., Gonzalez-Diaz, H., Ruso, J. M., Speck-Planche, A., & Cordeiro, M. N. (2014b). Computational tool for risk assessment of nanomaterials: novel QSTR-perturbation model for simultaneous prediction of ecotoxicity and cytotoxicity of uncoated and coated nanoparticles under multiple experimental conditions. *Environmental Science and Technology*, 48(24), 14686–14694.
- Krečmerová, M., Holý, A., Pískala, A., Masojídková, M., Andrei, G., Naesens, L., Neyts, J., Balzarini, J., De Clercq, E., & Snoeck, R. (2007). Antiviral Activity of Triazine Analogues of 1-(S)-[3-Hydroxy-2-(phosphonomethoxy)propyl]cytosine (Cidofovir) and related compounds. *Journal of Medicinal Chemistry*, 50(5), 1069–1077.
- Lameijer, E.-W., Kok, J. N., Bäck, T., & Ijzerman, A. P. (2006). The molecule evaluator. An interactive evolutionary algorithm for the design of drug-like molecules. *Journal of Chemical Information and Modeling*, 46(2), 545–552.
- Leach, A. R., & Gillet, V. J. (2007). *An introduction to chemoinformatics* (Rev. ed.). Dordrecht: Springer.
- Leach, A. G., Jones, H. D., Cosgrove, D. A., Kenny, P. W., Ruston, L., MacFaul, P., Wood, J. M., Colclough, N., & Law, B. (2006). Matched molecular pairs as a guide in the optimization of pharmaceutical properties; A study of aqueous solubility, plasma protein binding and oral exposure. *Journal of Medicinal Chemistry*, 49(23), 6672–6682.
- Lee, A., Mercader, A. G., Duchowicz, P. R., Castro, E. A., & Pomilio, A. B. (2012). QSAR study of the DPPH radical scavenging activity of di(hetero)arylamines derivatives of

- benzo[b]thiophenes, halophenols and caffeic acid analogues. *Chemometrics and Intelligent Laboratory Systems*, 116, 33–40.
- Lewis, R., & Leach, A. (1994). Current methods for site-directed structure generation. *Journal of Computer-Aided Molecular Design*, 8(4), 467–475.
- Liu, T., Lin, Y., Wen, X., Jorissen, R. N., & Gilson, M. K. (2007). BindingDB: A web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Research*, 35(Database issue), D198–D201.
- Liu, R., Rallo, R., George, S., Ji, Z., Nair, S., Nel, A. E., & Cohen, Y. (2011). Classification nano-SAR development for cytotoxicity of metal oxide nanoparticles. *Small*, 7(8), 1118–1126.
- Liu, R., Rallo, R., Weissleder, R., Tassa, C., Shaw, S., & Cohen, Y. (2013). Nano-SAR development for bioactivity of nanoparticles with considerations of decision boundaries. *Small*, 9(9–10), 1842–1852.
- Melagraki, G., & Afantitis, A. (2013). Enalos KNIME nodes: Exploring corrosion inhibition of steel in acidic medium. *Chemometrics and Intelligent Laboratory Systems*, 123, 9–14.
- Melagraki, G., & Afantitis, A. (2014). Enalos InSilicoNano platform: An online decision support tool for the design and virtual screening of nanoparticles. *RSC Advances*, 4, 50713–50725.
- Melagraki, G., & Afantitis, A. (2015). A risk assessment tool for the virtual screening of metal oxide nanoparticles through enalos insiliconano platform. *Current Topics in Medicinal Chemistry*, 15(18), 1827–1836.
- Mouchlis, V. D., Melagraki, G., Mavromoustakos, T., Kollias, G., & Afantitis, A. (2012). Molecular modeling on pyrimidine-urea inhibitors of TNF- α production: An integrated approach using a combination of molecular docking, classification techniques, and 3D-QSAR CoMSIA. *Journal of Chemical Information and Modeling*, 52, 711–723.
- Narayana Moorthy, N. S. H., & Poongavanam, V. (2015). The KNIME based classification models for yellow fever virus inhibition. *RSC Advances*, 5(19), 14663–14669.
- O’Boyle, N., Banck, M., James, C., Morley, C., Vandermeersch, T., & Hutchison, G. (2011). Open Babel: An open chemical toolbox. *Journal of Cheminformatics*, 3(1), 33.
- Ojha, P. K., & Roy, K. (2011). Comparative QSARs for antimalarial endochins: Importance of descriptor-thinning and noise reduction prior to feature selection. *Chemometrics and Intelligent Laboratory Systems*, 109(2), 146–161.
- Papa, E., Kovarich, S., & Gramatica, P. (2009). Development, validation and inspection of the applicability domain of QSPR models for physicochemical properties of polybrominated diphenyl ethers. *QSAR Combinatorial Science*, 28, 790–796.
- Patel, H., Gillet V. J., Chen, B., & Bodkin, M. J. (2007). *Development of a de novo design tool using reaction vectors*. In Poster presented at the 4th Joint Sheffield Conference on Chemoinformatics Sheffield, UK.
- Patel, H., Bodkin, M. J., Chen, B., & Gillet, V. J. (2009). Knowledge-based approach to de novo design using reaction vectors. *Journal of Chemical Information and Modeling*, 49(5), 1163–1184.
- Puzyn, T., Rasulev, B., Gajewicz, A., Hu, X., Dasari, T. P., Michalkova, A., Hwang, H.-M., Toropov, A., Leszczynska, D., & Leszczynski, J. (2011). Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. *Nature Nanotechnology*, 6(3), 175–178.
- Russo, E. (2002). Chemistry plans a structural overhaul. *Nature*, 419(6903), 4–7.
- Schneider, G., & Fechner, U. (2005). Computer-based de novo design of drug-like molecules. *Nature Reviews. Drug Discovery*, 4(8), 649–663.
- Schürer, S. C., Tyagi, P., & Muskal, S. M. (2005). Prospective exploration of synthetically feasible, medically relevant chemical space. *Journal of Chemical Information and Modeling*, 45(2), 239–248.
- Shao, C. Y., Chen, S. Z., Su, B. H., Tseng, Y. J., Esposito, E. X., & Hopfinger, A. J. (2013). Dependence of QSAR models on the selection of trial descriptor sets: A demonstration using nanotoxicity endpoints of decorated nanotubes. *Journal of Chemical Information and Modeling*, 53(1), 142–158.
- Shaw, S. Y., Westly, E. C., Pittet, M. J., Subramanian, A., Schreiber, S. L., & Weissleder, R. (2008). Perturbational profiling of nanomaterial biologic activity. *Proceedings of the National Academy of Sciences of the United States of America*, 105(21), 7387–7392.

- Sheridan, R. P., Hunt, P., & Culberson, J. C. (2006). Molecular transformations as a way of finding and exploiting consistent local QSAR. *Journal of Chemical Information and Modeling*, 46(1), 180–192.
- Speck-Planche, A., Kleandrova, V. V., Luan, F., & Cordeiro, M. N. (2015). Computational modeling in nanomedicine: prediction of multiple antibacterial profiles of nanoparticles using a quantitative structure-activity relationship perturbation model. *Nanomedicine (London, England)*, 10(2), 193–204.
- Steinbeck, C., Han, Y. Q., Kuhn, S., Horlacher, O., Luttmann, E., & Willighagen, E. L. (2003). The Chemistry Development Kit (CDK): An open-source Java library for chemo- and bioinformatics. *Journal of Chemical Information and Computer Sciences*, 43(2), 493–500.
- Sushko, I., Novotarskyi, S., Körner, R., Pandey, A. K., Rupp, M., Teetz, W., Brandmaier, S., Abdelaziz, A., Prokopenko, V. V., Tanchuk, V. Y., Todeschini, R., Varnek, A., Marcou, G., Ertl, P., Potemkin, V., Grishina, M., Gasteiger, J., Schwab, C., Baskin, I. I., Palyulin, V. A., Radchenko, E. V., Welsh, W. J., Kholodovych, V., Chekmarev, D., Cherkasov, A., Aires-de-Sousa, J., Zhang, Q.-Y., Bender, A., Nigsch, F., Patiny, L., Williams, A., Tkachenko, V., & Tetko, I. V. (2011). Online chemical model environment (OCHEM): Web platform for data storage, model development and publishing of chemical information. *Journal of Computer-Aided Molecular Design*, 25(6), 533–554.
- Tetko, I. V. (2012). The perspectives of computational chemistry modeling. *Journal of Computer-Aided Molecular Design*, 26, 135–136.
- Todeschini, R., & Consonni, V. (2009). *Molecular descriptors for chemoinformatics*. Weinheim: Wiley.
- Toropov, A. A., Toropov, A. P., Martyanov, S. E., Benfenati, E., Gini, G., Leszczynska, D., & Leszczynski, J. (2012). CORAL: Predictions of rate constants of hydroxyl radical reaction using representation of the molecular structure obtained by combination of SMILES and Graph approaches. *Chemometrics and Intelligent Laboratory Systems*, 112, 65–70.
- Toropov, A. A., Toropova, A. P., Puzyn, T., Benfenati, E., Gini, G., Leszczynska, D., & Leszczynski, J. (2013). QSAR as a random event: Modeling of nanoparticles uptake in PaCa2 cancer cells. *Chemosphere*, 92(1), 31–37.
- Tropsha, A. (2010). Best Practices for QSAR model development, validation, and exploitation. *Molecular Informatics*, 29(6–7), 476–488.
- Vinkers, H. M., de Jonge, M. R., Daeyaert, F. F. D., Heeres, J., Koymans, L. M. H., van Lenthe, J. H., Lewi, P. J., Timmerman, H., Van Aken, K., & Janssen, P. A. J. (2003). Synopsis: Synthesize and optimize system in Silico. *Journal of Medicinal Chemistry*, 46(13), 2765–2773.
- Vrontaki, E., Mavromoustakos, T., Melagraki, G., & Afantitis, A. (2015). Quantitative nanostructure-activity relationship models for the risk assessment of nanomaterials. In K. Roy (Ed.), *Quantitative structure-activity relationships in drug design, predictive toxicology, and risk assessment* (pp. 537–561). Hershey, PA: IGI Global.
- Wang, L., Shen, J., Tang, Y., Chen, Y., Wang, W., Cai, Z., & Du, Z. (2007). Synthetic improvements in the preparation of clopidogrel. *Organic Process Research & Development*, 11(3), 487–489.
- Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J., & Bryant, S. H. (2009). PubChem: A public information system for analyzing bioactivities of small molecules. *Nucleic Acids Research*, 37(Web Server issue), W623–W633.
- Weissleder, R., Kelly, K., Sun, E. Y., Shtatland, T., & Josephson, L. (2005). Cell-specific targeting of nanoparticles by multivalent attachment of small molecules. *Nature Biotechnology*, 23, 1418–1423.
- Winkler, D. A., Mombelli, E., Pietroiusti, A., Tran, L., Worth, A., Fadeel, B., & McCall, M. J. (2013). Applying quantitative structure-activity relationship approaches to nanotoxicology: Current status and future potential. *Toxicology*, 313(1), 15–23.
- Winkler, D. A., Burden, F. R., Yan, B., Weissleder, R., Tassa, C., Shaw, S., & Epa, V. C. (2014). Modelling and predicting the biological effects of nanomaterials. *SAR and QSAR in Environmental Research*, 25(2), 161–172.
- Witten, I. H., Frank, E., & Hall, M. A. (2005). *Data mining, practical machine learning tools and techniques*. San Francisco, CA: Elsevier.

- Zhang, S., Golbraikh, A., Oloff, S., Kohn, H., & Tropsha, A. (1995). Novel Automated Lazy Learning QSAR (ALL-QSAR) approach: Method development, applications, and virtual screening of chemical databases using validated ALLQSAR models. *Journal of Chemical Information and Modeling*, *46*, 1984–1995.
- Zhang, H., Ji, Z., Xia, T., Meng, H., Low-Kam, C., Liu, R., Pokhrel, S., Lin, S., Wang, X., Liao, Y.-P., Wang, M., Li, L., Rallo, R., Damoiseaux, R., Telesca, D., Mädler, L., Cohen, Y., Zink, J. I., & Nel, A. E. (2012). Use of metal oxide nanoparticle band gap to develop a predictive paradigm for oxidative stress and acute pulmonary inflammation. *ACS Nano*, *6*(5), 4349–4368.

Paola Gramatica

Contents

Introduction	2232
QSAR Modeling for Prioritization	2233
Ranking Indexes: PC Scores as New Cumulative Endpoints for QSAR Models	2234
Multivariate Explorative Methods: PCA and HCA	2235
QSAR Modeling of Ranking Indexes and Classes	2236
Environmental Behavior	2236
Leaching of Pesticides	2236
Persistence	2239
Toxicity	2245
Endocrine Disruptors (EDs)	2245
Ecotoxicity	2247
Mammalian Toxicity on Rodents of Fluorinated Chemicals	2252
Persistence Bioaccumulation Toxicity (PBT)	2253
Conclusions	2257
Bibliography	2258

Abstract

Several different chemical properties/activities must be contemporaneously taken into account to prioritize compounds for their hazardous behavior. Examples of application of chemoinformatic methods, such as principal component analysis for obtaining ranking indexes and hierarchical cluster analysis for grouping chemicals with similar properties, are summarized for various classes of compounds of environmental concern. These cumulative endpoints are then modeled

P. Gramatica (✉)

QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Theoretical and Applied Sciences, University of Insubria, Varese, Italy
e-mail: paola.gramatica@uninsubria.it

by validated quantitative structure–activity relationships, based on theoretical molecular descriptors, to predict the potential hazard of new chemicals.

Introduction

The chemical universe is huge and is rapidly enlarging every day: the number of chemicals registered in the Chemical Abstract Service (CAS) registry (www.cas.org) gets nowadays over 100 million of chemicals, the majority of them are commercially available, and almost 345,000 are regulated and listed in various inventories (for instance, EU-EINECS, US-EPA TSCA, Canada-DSL). While many new chemicals are being developed continuously (many thousands each year), with increasing possibility to interact with humans and wildlife, information on physicochemical properties, reactivity, and biological activities are more slowly produced. The problem of lack of data and slow assessment procedures is highly significant: in fact, so far, we have extensive information about only a few chemicals, but the majority of compounds (>95 %), even high production volume (HPV) compounds, have been not sufficiently well characterized for their environmental behavior and potential to cause human or ecologic toxicity (Judson et al. 2009; Arnot et al. 2012). The filling of this data gap, in order to assess and control the chemicals effectively, is one of the main aims of several legislations worldwide, in particular of the recent European legislation REACH (Registration Evaluation Authorization and restriction of Chemicals) (EC Regulation 2006).

However, it is clearly impossible to measure all chemicals in all media to which humans and ecological receptors are exposed, as well as to test a plethora of endpoints. In order to reduce costs, time, and sacrificed animals, there is an urgent need to prioritize the use of testing resources toward those chemicals and endpoints that present the greatest potential of risk to human health and environment. Therefore, it is highly evident that the prioritization of chemicals is nowadays a big challenge, mainly for the identification of new emerging pollutants.

Chemoinformatic analysis has a clear and fundamental role in dealing with this issue. By chemoinformatic methods, it is possible to analyze and model the experimental information, which is already available for tested chemicals, and to exploit this information applying the developed tools to chemicals without experimental data or even before their synthesis. The main aim is to better use the existing knowledge for preventing, as soon as possible, potential dangerous properties of not yet tested compounds and also for planning *a priori* the synthesis of safe chemicals.

In recent years, many scientists have faced this important problem, studying various endpoints by different chemoinformatic approaches, but with the same common aim: to highlight the most hazardous chemicals by screening data sets of several compounds (here some representative of more recent examples: Gramatica and Di Guardo 2002; Salvito et al. 2002; Sanderson et al. 2003, 2004; Schmieder et al. 2003; Tong et al. 2003; Gramatica et al. 2004a, b, 2015, 2016a, b; Knekta et al. 2004; Öberg 2004, 2005, 2006; Muir and Howard 2006; Klasmeier et al. 2006; Hansson and Rudén 2006; Liu et al. 2006, 2007; Dix et al. 2007; Gramatica and

Papa 2007; Brown and Wania 2008; Papa and Gramatica 2008, 2010; Wegmann et al. 2009; Judson et al. 2009; Stenberg et al. 2009; Kavlock and Dix 2010; Li and Gramatica 2010a, b; Bhatarai and Gramatica 2010, 2011a; Howard and Muir 2010; Kovarich et al. 2011, 2012; Öberg and Iqbal 2012; Stempel et al. 2012; Roos et al. 2012; Sanderson 2012; Zarfl et al. 2012; Scherlinger et al. 2012; Arnot et al. 2012; Guillen et al. 2012; Singh et al. 2014; Cassani and Gramatica 2015; Wedebye et al. 2015; Sangion and Gramatica 2016a, b).

In this chapter, the crucial topic of screening chemicals of heterogeneous molecular structure is faced, studying some specific endpoints and focusing on compounds of environmental concern. These screening studies have two different aims: (1) to rank, highlight, and prioritize the most hazardous compounds among the already used chemicals, also those without experimental data, and (2) to predict the potential dangerous behavior of not yet synthesized compounds, in an *a priori* approach of the “benign by design” strategy of green chemistry. The focus here will be on the potential hazard intrinsically related to the chemical structure, thus on the utility of QSAR (quantitative structure–activity relationship) modeling, in particular based on a preliminary chemoinformatic analysis. A chemometric method of explorative analysis, such as principal component analysis (PCA), is applied for defining trends and ranking indices, as well as for the *a priori* data set splitting for external validation of QSAR models. Hierarchical cluster analysis (HCA) is used for grouping chemicals, according to some properties, and for defining the *a priori* classes for a subsequent classification by various classification methods. Particular emphasis is here devoted to multivariate linear regression (MLR) models, in particular ordinary least squares (OLS) models, based on genetic algorithm (GA) for variable selection and developed by the software QSARINS (QSAR-INSubria) for QSAR model development and validation (Gramatica et al. 2013). Some of the QSAR models of cumulative ranking endpoints here presented, applied to several classes of chemicals of emerging concern (CEC), are implemented in the module QSARINS-Chem (Gramatica et al. 2014) for easy applicability. Moreover, results of prioritization of endocrine disruptors (EDs), performed by various classification models, are also here commented.

A previous review of these chemoinformatic approaches, presenting the basis of externally validated QSAR modeling, illustrated according to the OECD principles for QSAR in regulation (OECD 2004) was published in Chapter 12 of the book *Recent Advances in QSAR Studies* (Gramatica 2009), edited by the same editors of the present book. Therefore, this chapter is mostly an updating of the previous one on the approaches widely applied in the Insubria QSAR research group, with special emphasis on the QSAR modeling of ranking indices obtained by PCA.

QSAR Modeling for Prioritization

QSAR (quantitative structure–activity relationship) modeling is based on the assumption that the molecular structure of a chemical (i.e., its geometric, steric, and electronic properties) contains the features responsible for its physical, chemical, and biological properties. Such modeling techniques are the best chemoinformatic

approaches for finding and exploiting the information inherent in the molecular structure related to the intrinsic hazard of any chemical. In fact, by QSAR models, based on theoretical molecular descriptors and validated chemometric methods, both of regression and classification, the biological activity (or property, reactivity, etc.) of new or untested chemicals can be inferred from the molecular structure of compounds whose activities (properties, reactivities, etc.) have already been assessed.

There is no need, in this chapter, to enter into details on QSAR modeling, commented elsewhere in this book; this has been also the topic of my chapter of the above-cited book (Gramatica 2009) and of all my papers in these last 20 years, some cited also here.

However, it is important to stress the point that, particularly for prioritization aim in screening big data sets, it is not sufficient to be able to reproduce well the available data and it should not be of primary importance to understand underlying mechanisms. To know the real predictivity of QSAR models and to which chemicals the model could be more reliably applied is of crucial and fundamental relevance. For this reason, all my QSAR works, which are mainly devoted to screening and prioritization, are focused on external validation on chemicals never used for model development and on model applicability domain (AD) check (Gramatica 2007, 2014; Gramatica et al. 2012). The proposal of concordance correlation coefficient (CCC) as validation parameter for QSAR models and its comparison with other used statistical parameters (Chirico and Gramatica 2011, 2012), highlighting some drawbacks and proposing intercomparable thresholds for real predictivity, were also done for this purpose. Recently, all these compared external validation parameters, and the Insubria graph for checking the applicability domain of QSAR models to chemicals without experimental data, have been implemented in our software QSARINS, freely available on request (www.qsar.it) for academia and research centers.

Ranking Indexes: PC Scores as New Cumulative Endpoints for QSAR Models

The behavior of chemicals in the environment and their impact on humans and wildlife are dependent on many different variables such as physicochemical properties, chemical reactivity, biological activity, etc. of the compounds. Since many parameters could be of contemporary importance, it is crucial to understand, rationalize, and interpret the covariance, which is inherent in this environmental complexity. Explorative methods of multivariate analysis, applied to various topics of environmental concern, give a combined view that generates ranking of the studied chemicals and highlight variable relationships. Then, based on these chemoinformatic tools, a more focused investigation can be made into chemicals of higher concern, guiding experimental tests on the prioritized compounds.

A multivariate explorative technique, such as principal component analysis (PCA), is used by several researchers (e.g., Knekt et al. 2004; Öberg and Iqbal

2012) to visualize the distribution of chemicals, represented by structural descriptors or environmental properties, with the aim to select representative compounds for experimental testing or to highlight the structural properties more related to specific hazardous behavior.

Some of my researches, summarized in this chapter, led to the proposal of a cumulative index, based on the outcome of the application of PCA for chemical screening and ranking. This index (the PC1 score) condenses the main information related to the studied properties. If it explains a reasonably significant variance of the studied variables, it can be usefully modeled as a new aggregate endpoint by QSAR approaches. The developed QSAR models exploit the already available information and can be used to predict the potential behavior of chemicals without experimental values or of new chemicals even before their synthesis. In fact, the QSAR approach, which is based on theoretical molecular descriptors that can be calculated for whatever drawn chemicals, can be applied without knowledge of any experimental parameter.

Similarly, another chemoinformatic method, such as hierarchical cluster analysis, can be applied for grouping chemicals according to their similarity based on several properties. The obtained groups can then be modeled by QSAR classification methods.

Many studies and published papers of the author 20-year researches at Insubria University are focused on this combined approach of chemoinformatic analysis (ranking indexes from PCA plus regression QSAR models of these cumulative indexes; groups modeled as classes by classification QSAR models) for prioritization aims. These will be summarized in the following paragraphs, organized according to the studied endpoints.

Multivariate Explorative Methods: PCA and HCA

The multivariate explorative techniques have the principal aim to condense the information, present in any multivariate data set, into a more easily interpretable view.

Principal component analysis (PCA) (Jackson 1995; Jolliffe 2002) is probably the most widely known and used explorative multivariate method. In PCA, the studied variables are linearly combined so that the obtained combinations (the principal components, PCs) explain the variation in the original data with decreasing explained variance. The first principal component (PC1) condenses the maximum amount of possible data variance in a single variable, while the following orthogonal PCs account for successively smaller quantities of the original variance.

To be useful and be considered sufficiently representative of the main information included in the data, it is desirable that the first two PCs account for a substantial proportion of the variance in the original data, while the remaining PCs condense irrelevant information and noise and could be disregarded. The more common representations of PCA are score plot, loading plot, and biplot, defined as the joint representation of the rows and columns of a data matrix: the points (scores) represent

the chemicals, while the vectors or lines represent the variables (loadings). The length of each vector indicates the information associated with that specific variable, while the angle between the vectors reflects their correlation.

Hierarchical cluster analysis (HCA) (Kaufman and Rousseeuw 1990) is a clustering technique with the purpose to build a binary tree of the data that successively merges similar groups of points. HCA creates clusters according to the measure of distance or similarity between data points, based on measured characteristics, connecting, in an iterative process, the nearest groups of objects. It is based on the idea that objects are more related to nearby objects than to objects farther away. The main output of the HCA is the dendrogram that summarizes the relationships between objects in a visual binary tree. Clusters can be identified cutting the dendrogram at different similarity levels. Hierarchical cluster analysis is one of the best ways to observe how homogeneous groups of objects with similar properties are formed and to identify classes.

QSAR Modeling of Ranking Indexes and Classes

In our environmental chemistry studies, PCA for obtaining ranking indexes and regression QSAR modeling of these ranking indexes, as new cumulative endpoints, and HCA for obtaining *a priori* classes and classification QSAR models have been widely used in my group for screening, ranking, and priority setting in many contexts. I'll cite here and comment below only some of the most significant and/or recent: (a) environmental partitioning and leaching of pesticides (Gramatica and Di Guardo 2002; Gramatica et al. 2004a) and benzotriazoles (Bhatarai and Gramatica 2011b), (b) degradability of volatile organic compounds (VOCs) by tropospheric oxidants (Gramatica et al. 2004b), (c) persistence of POPs by global half-life index (GHLI) (Gramatica and Papa 2007; Papa and Gramatica 2008), (d) rat/mouse toxicity of perfluorinated compounds (PFCs) (Bhatarai and Gramatica 2010, 2011a), (e) aquatic toxicity of personal care products (PCPs) (Gramatica et al. 2016b) and pharmaceuticals (Sangion et al. 2015; Sangion and Gramatica 2016b), and (f) PBT screening of various compounds by PBT Index (Papa and Gramatica 2010; Gramatica et al. 2015, 2016a; Cassani and Gramatica 2015; Sangion and Gramatica 2016a).

Environmental Behavior

Leaching of Pesticides

The tendency of pesticides to pollute groundwaters and, in general, the partitioning of pesticides into different environmental compartments depend mainly on physicochemical properties, such as soil organic carbon partition coefficient (K_{oc}), *n*-octanol/water partition coefficient (K_{ow}), water solubility (S_w), vapor pressure

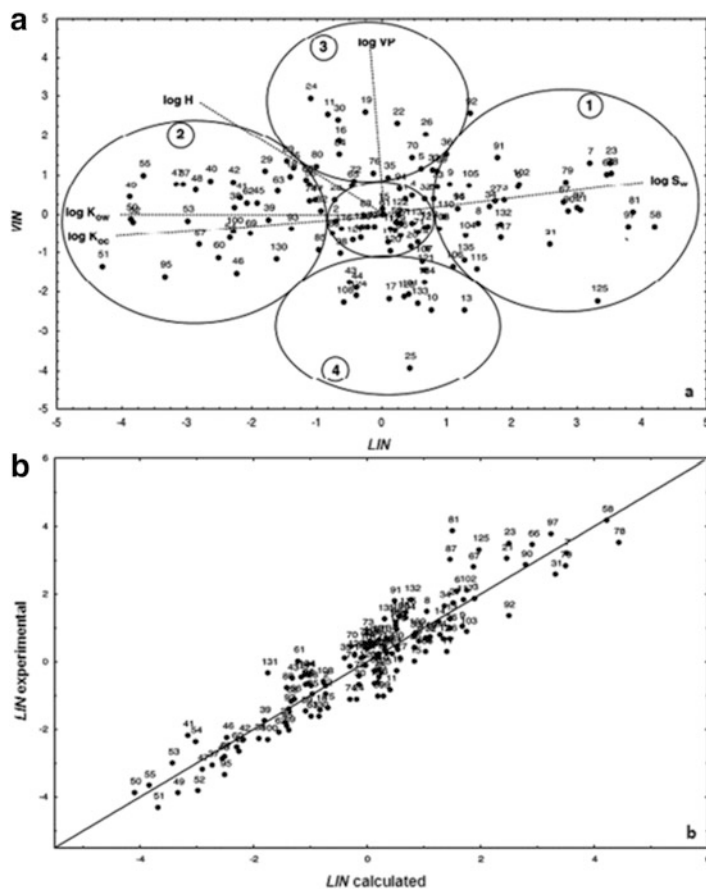


Fig. 1 (a) PCA of environmental physicochemical properties of 135 pesticides and definition of leaching index (*LIN*) and volatility index (*VIN*). (b) Scatter plot of the OLS model of *LIN* (Permission from Gramatica and Di Guardo, *Chemosphere*, 2002)

(VP), and Henry's law constant (H). We have applied PCA on these various environmental partitioning properties of a heterogeneous and highly representative data set of 135 pesticides of different chemical classes (acetanilides, carbamates, dinitroanilines, organochlorines, organophosphates, phenylureas, triazines, triazoles) to study the tendency to leach from soil into the surface and subsurface waters (Gramatica and Di Guardo 2002) (Fig. 1a).

The resultant macrovariables, the PC1 and PC2 scores, were called leaching index (*LIN*) and volatility index (*VIN*) and were proposed as cumulative partitioning indexes in different environmental media. The component *LIN* tends to discriminate between the relatively more sorbed/less soluble (on the left of Fig. 1a) and the less sorbed/more soluble pesticides (on the right), while *VIN* appears to differentiate between volatile (upper part of Fig. 1a) and nonvolatile compounds.

Both indexes were modeled by OLS QSAR model, using theoretical DRAGON molecular descriptors (Talete 2007) selected by genetic algorithms: they are mainly atom and group count parameters as a number of halogens, nitro groups and sulfur, plus Ms the mean electrotopological state of the molecule related to the polarizability, and a topological descriptor ICR (the radial centric information index) (Bonchev and Rouvray 1991). The model robustness and internal predictive power are satisfactory; below the QSAR model equation for LIN, the cumulative index which explains 65 % of the data variance and the corresponding plot in Fig. 1b are reported:

$$\text{LIN} = -3.04 - 0.96nX - 2.28n\text{NO}_2 + 3.42\text{Ms} - 1.74\text{ICR} - 0.45n\text{S}$$
$$n = 135, R^2 = 87.0 \%, Q^2_{\text{LOO}} = 85.8 \%, Q^2_{\text{LMO}} = 85.7 \%,$$
$$s = 0.66, F = 172.22, \text{SDEP} = 0.68 \text{ and } \text{SDEC} = 0.65.$$

A combination of two chemoinformatic methods, principal component analysis for ranking and hierarchical cluster analysis for the definition of four *a priori* classes (Fig. 2a), according to the environmental behavior as soluble, sorbed, volatile, or nonvolatile/medium class, was applied to the environmental physicochemical properties of 54 pesticides of various chemical categories (Gramatica et al. 2004b). The pesticides were finally assigned to the defined four classes by three different classification methods (classification and regression tree (CART, the classification tree in Fig. 2b), k-nearest neighbors (k-NN), and regularized discriminant analysis (RDA)) with misclassification risk in cross validation ranging from 17 % to 18 %. The discriminant variables were simple theoretical molecular descriptors, such as MW, nHDon, and topological Balaban Index (Balaban 1983) (named J in DRAGON). MW, which encodes information on molecule dimension, is able to discriminate the chemicals that are contemporaneously most sorbed in organic soils and least soluble in water; in fact it is well known that biggest molecules have the greatest tendency to bind, by van der Waals forces, to the organic component of the soil (mainly humic acids). The more soluble pesticides, which have the higher possibility to form hydrogen bonds with water molecules, are discriminated by nHDon, the number of groups able to donate hydrogen in the hydrogen bonds. Furthermore, the chemicals with fewer intramolecular hydrogen bonds are the most volatile.

Similar PCA ranking was also useful to highlight which benzotriazoles could be most dangerous for the aquatic compartment (Bhatarai and Gramatica 2011a): these chemicals, used in the past mainly as pesticides, are now recognized as new contaminants of emerging concern (CEC) for the environment. In fact they are nowadays used as deicing additives and are a major source of pollution predominantly of aquatic resources near the airports of major cities.

The presented chemoinformatic approaches allow the screening of the environmental distribution of pesticides and a rapid predetermination of their possibility to

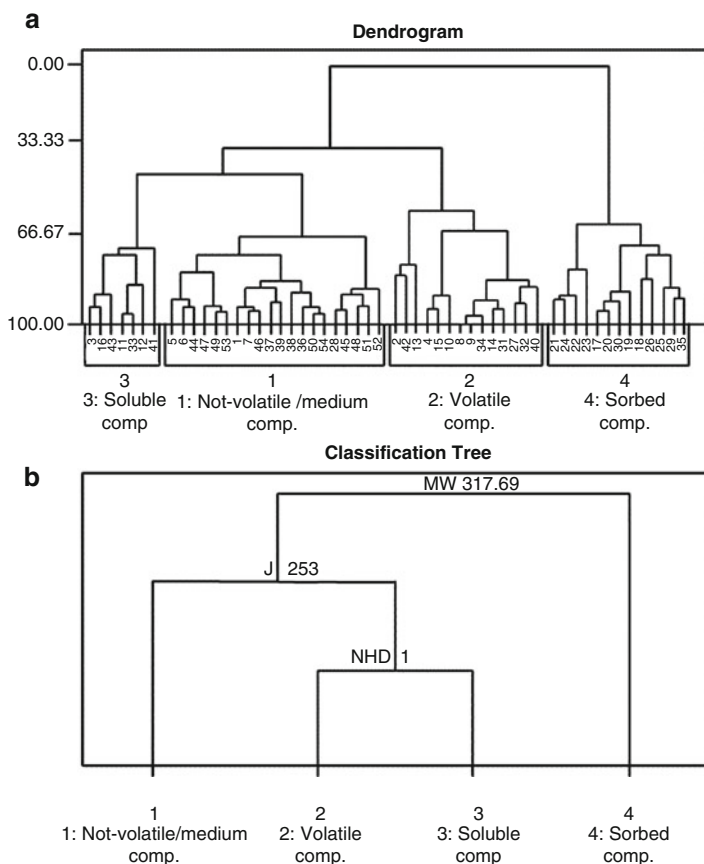


Fig. 2 (a) Hierarchical cluster analysis of the environmental physicochemical properties of 54 pesticides for the definition of four *a priori* classes. (b) Classification tree of CART model of the four classes (Permission from Gramatica et al. *Int. J. Environ. Anal. Chem.*, 2004b)

pollute both surface and groundwaters, starting only from the molecular structure without any *a priori* knowledge of the physicochemical properties.

Persistence

Persistence in the environment is an important criterion in prioritizing hazardous chemicals and in identifying new persistent organic pollutants (POPs).

Various studies, based on various approaches, have been performed on this topic (Öberg 2005, 2006; Muir and Howard 2006; Klasmeier et al. 2006; Wegmann et al. 2009; Howard and Muir 2010; Puzyn et al. 2011; Scheringer et al. 2012). Here some

of my studies regarding the screening and prioritization using cumulative indexes are summarized.

Degradability of Volatile Organic Compounds (VOCs) by Tropospheric Oxidants

An indirect measure of the persistence of volatile organic compounds (VOCs) in the atmosphere, and therefore a necessary preliminary parameter in environmental exposure assessment, is the degradability, measured by the reaction rates with the main tropospheric oxidants: hydroxyl radical and ozone during the daytime and nitrate radical at night. The contemporaneous variation and influence of the rate constants of the degradation by three oxidants (kOH, kNO₃, and kO₃) of several VOCs, in determining their inherent tendency to degradability, were explored by principal component analysis (Fig. 3a). The first component, along which the variables are grouped in the same direction, discriminates between the less degradable compounds, so the relatively more persistent (chemicals to the right in the PCA graph) and the more degradable chemicals (chemicals to the left).

Thus, the PC1 score, which explained 81 % of the data variance for 399 chemicals, was defined as ATDIN, an atmospheric degradability index, and was modeled by OLS, based on theoretical molecular descriptors and externally validated (Gramatica et al. 2004a) (scatter plot in Fig. 3b).

It is also interesting to note that the PC2 is able to highlight the different reactivity of chemicals with ozone and with OH and NO₃ radicals, respectively.

$$\text{ATDIN} = -17.59 - 1.80\text{HOMO} + 2.87\text{nBnz} - 0.51\text{BEHe4}$$

$$n_{\text{training}} = 227, R^2 = 93.9 \%, Q^2_{\text{LOO}} = 93.7 \%, Q^2_{\text{LMO}(50\%)} = 93.5 \%;$$

$$n_{\text{prediction}} = 172; Q^2_{\text{EXT}} = 92.3 \%;$$

$$s = 0.387; \text{SDEC} = 0.384; \text{SDEP} = 0.391$$

The molecular descriptors of this model are informative of different aspects of the studied reaction. The best descriptor is the energy of the highest occupied molecular orbital (HOMO), as a measure of the molecular reactivity. The number of aromatic rings (nBnz) is probably selected in the model to encode for the different reactivity of aliphatic and aromatic chemicals in relation to the attack sites for the three different oxidants. 2D-BCUT descriptors, BEHe4, weighted by the atomic electronegativity of Sanderson (Burden 1989) encode for charge distribution factors.

This model can be useful in avoiding the release into the environment of potential persistent volatile compounds which are not inherently degradable in troposphere, causing risk to humans and wildlife for their persistence at that atmospheric level. Moreover, these chemicals could reach the stratosphere with potential dangerous behavior on the ozone layer.

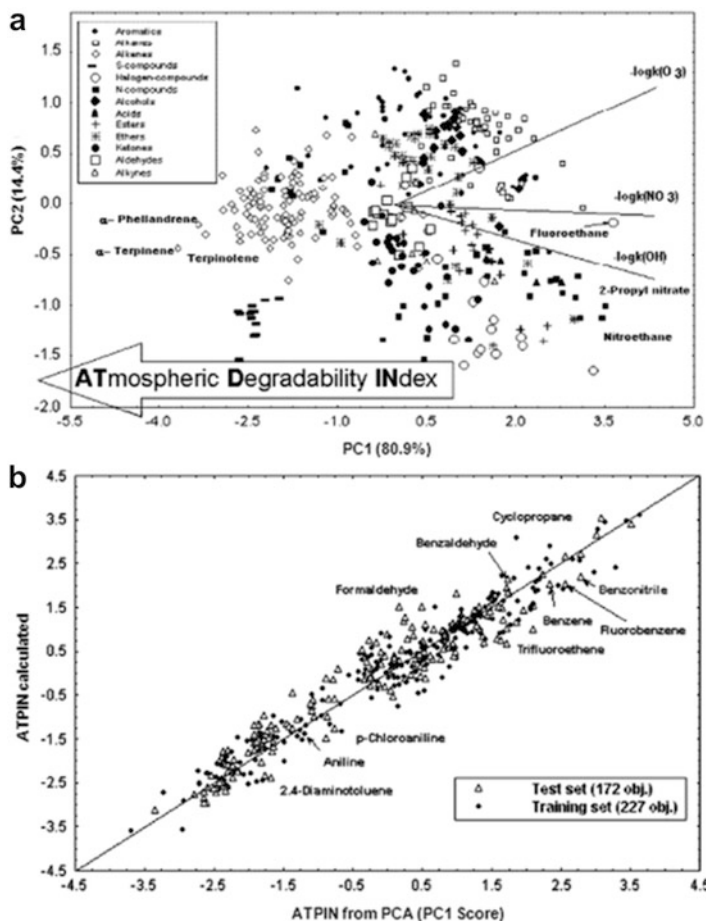


Fig. 3 (a) PCA of the kinetic rate constants of degradation by tropospheric oxidants of 399 VOCs and definition of ATDIN. (b) Scatter plot of QSAR model of ATDIN (Permission from Gramatica et al. *Atmos Environ*2004a)

Screening of POPs by Environmental Half-Life (HL)

The degradation half-lives in various compartments are among the more commonly used criteria for studying environmental persistence. Available half-life data for degradation in air, water, sediment, and soil, for a set of 250 organic chemicals, were combined in multivariate approach by principal component analysis. A ranking of the studied organic pollutants according to their relative overall half-life is obtained in this way: we named this global half-life index (GHLI) (Gramatica and Papa 2007).

The biplot relative to the first and second components is reported in Fig. 4a, where the chemicals (points) are distributed according to their environmental persistence, represented by the linear combination of their half-lives in the four selected

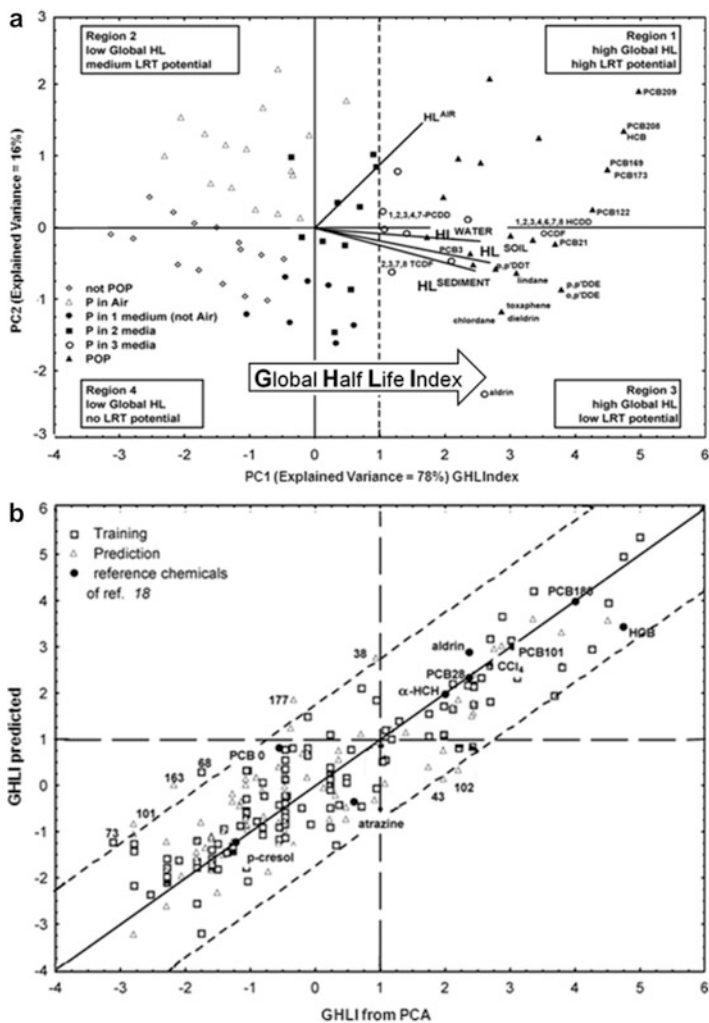


Fig. 4 (a) PCA of the half-lives in four environmental compartments of 250 heterogeneous chemicals and definition of global half-life index (*GHLI*) for POPs. (b) Scatter plot of the QSAR model of *GHLI* (Permission from Gramatica and Papa, *Environ. Sci. Technol.*, 2007)

media (the loading lines show the importance of each variable in the first two PCs). The cumulative explained variance of the first two PCs is 94 %; the PC1 alone providing a very significant part of the total information is 78 %. Since all the half-lives in different media (the lines) are oriented in the same direction along the first principal component, PC1 is a new macro-variable representing cumulative half-life and condensing chemical tendency to environmental persistence. Therefore, PC1 is useful to discriminate chemicals with regard to persistence: chemicals with high

half-life values in all the media are located to the right of the PCA plot (Fig. 4a), in the zone of global higher persistence (very persistent chemicals anywhere); chemicals with a lower global half-life fall to the left of the graph, not being persistent in any medium.

PC2, although less informative (E.V. 16 %), is also interesting: it separates the compounds more persistent in air (upper parts in Fig. 4a, region 1), i.e., those with higher long-range transfer (LRT) potential, from those more persistent in water, soil, and sediment (region 3 in Fig. 4a).

A deeper analysis of the distribution of the studied chemicals confirms experimental evidences: to the right, among the very persistent chemicals in all the compartments (full triangles in Fig. 4a), most of the compounds recognized as POPs by the Stockholm Convention (UNEP 2014) are located. Highly chlorinated PCBs and hexachlorobenzene are among the most persistent compounds in this reference scenario; they are grouped in region 1 owing to their global high persistence, especially in air. The less chlorinated PCBs (PCB 3 and PCB 21), p,p'-DDT, p,p'-DDE and o,p'-DDE, highly chlorinated dioxins and some dioxin-like compounds, as well as pesticides toxaphene, lindane, chlordane, dieldrin, and aldrin fall in region 3 of highly persistent chemicals, mainly in compartments different from air.

This global index GHLI was then modeled as a cumulative endpoint using a QSAR approach based on theoretical DRAGON molecular descriptors. The original set of available data was first randomly split into training and prediction sets: 50 % of the compounds (125 compounds) were used for OLS model development, while the other 50 % was put into the prediction set to validate the QSPR model. Given below is the best model, selected by statistical approaches, and its statistical parameters, which confirm model robustness and real external predictivity. Figure 4b shows the plot of GHLI values from PCA versus the predicted GHLI values:

$$\text{GHL Index} = -3.12 + 0.33X_{0v} + 5.1Mv - 0.32MAXDP - 0.61nHDon \\ - 0.5CIC0 - 0.61O - 060$$

$$n_{\text{training}} = 125; R^2 = 0.85; Q^2_{\text{LOO}} = 0.83; Q^2_{\text{BOOT}} = 0.83,$$

$$\text{RMSE} = 0.76; \text{RMSE}_{\text{cv}} = 0.70;$$

$$n_{\text{prediction}} = 125, R^2_{\text{EXT}} = 0.79; \text{RMSEP} = 0.78.$$

A similar highly predictive model for GHLI, based on PaDEL-Descriptors (Yap 2011), has been recently implemented in the module QSARINS-Chem of the software QSARINS (www.qsar.it) for easy applicability on new compounds, in order to help in avoiding the synthesis of chemicals that have, inherent in their molecular structure, the potentiality to be POPs.

The chemical environmental half-lives were also used for POP pre-screening (Papa and Gramatica 2008) developing predictive classification models, based on k-NN, CART, and counter-propagation artificial neural networks (CP-ANN). In this approach, the three *a priori* classes of different degrees of persistence (high, medium, and low) were determined by hierarchical cluster analysis (Fig. 5a) applied

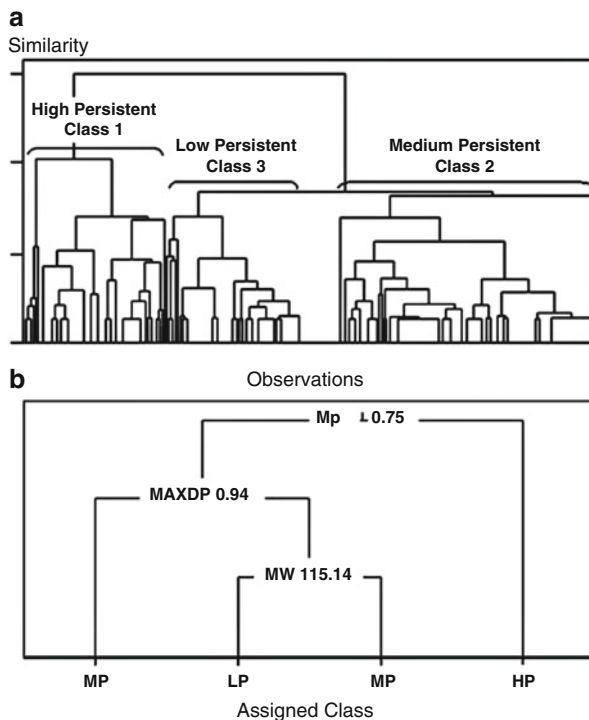


Fig. 5 (a) Hierarchical cluster analysis of environmental half-lives for defining three classes of persistence; (b) CART tree of three classes for POP ranking (Permission from Papa and Gramatica *J. Mol. Graph. Mod.*, 2008)

to environmental half-lives. The range of overall external predictivity of the three classification models was high, 75–85 %. The three discriminant structural variables selected in this study (mean polarizability, Mp; maximum electrotopological variation, MAXDP; and molecular weight, MW, in the decision nodes of the CART tree in Fig. 5b) are all bidimensional descriptors independent of chemical conformation, thus easily calculable from the bidimensional structural graph of a compound. The calculations of these variables can be performed starting from the simple SMILES string of a chemical.

The application of both kind of models, regression of GHLI and classification of persistence classes, using only a few structural descriptors, could allow a fast preliminary identification and prioritization of not yet known POPs, just from the knowledge of their molecular structure. The proposed multivariate approach is particularly useful not only to screen and to make an early prioritization of environmental persistence for pollutants already on the market, but also for not yet synthesized compounds, which could represent safer alternative and replacement solutions for recognized POPs. No method other than QSAR is applicable to detect the potential persistence of completely new compounds.

Toxicity

Endocrine Disruptors (EDs)

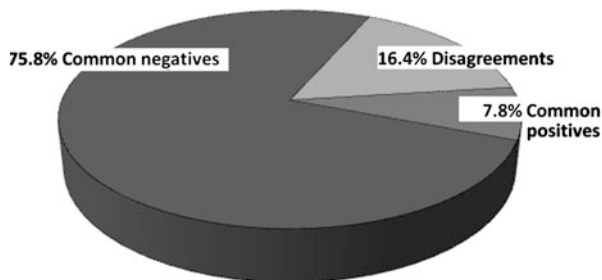
A large number of environmental chemicals are suspected to disrupt endocrine systems by mimicking or antagonizing natural hormones. Such chemicals, named endocrine disruptors (EDs), may have dangerous effects on the health of humans and wildlife. Under REACH, the chemicals with demonstrated endocrine disruption activity require authorization to be produced and used; in addition, safer alternatives should be proposed. However, it is practically impossible to perform a variety of toxicological tests on all potential EDs; thus, QSAR modeling for prioritization has been applied by many authors in these last years (e.g., Shi et al. 2001; Hong et al. 2002; Fang et al. 2003; Schmieder et al. 2003; Tong et al. 2003; Roncaglioni et al. 2004; Asikainen et al. 2004, 2006; Saliner et al. 2006; Devillers et al. 2007, 2015; Dybdahl et al. 2012; Vuorinen et al. 2013; Browne et al. 2015) providing promising methods for the screening of a set of chemicals for potential estrogenic activity.

QSAR models of the estrogen receptor binding affinity of a data set of 128 NCTR heterogeneous compounds (Ding et al. 2010) were built also in our laboratory by OLS method using theoretical DRAGON descriptors (Liu et al. 2006), giving full consideration to the OECD principles regulating QSAR acceptability (OECD 2004) during model construction and assessment.

The results of several validation paths using different splitting methods (D-optimal design, self-organizing maps (SOM), random on activity sampling) give proof that the proposed QSAR model is robust and satisfactory (Q^2_{pred} range, 0.76–0.81), thus providing a feasible and practical tool for the rapid screening of the estrogen activity of organic compounds, supposed EDs. A similar regression model, based on the same training set, and on PaDEL-Descriptors (Yap 2011), is now available in the QSARINS-Chem module of the QSARINS software for easy application on new chemicals (www.qsar.it).

On the same topic, satisfactory predictive models for EDs' classification, based on different classification methods, were proposed (Liu et al. 2007). In this study, QSAR models, based on 232 structurally diverse chemicals from the NCTR database as training set and on theoretical structural descriptors, were developed to quickly and effectively identify possible estrogen-like chemicals by using nonlinear classification methodologies (least squares support vector machine (LS-SVM), counter-propagation artificial neural network (CP-ANN), and k-nearest neighbor (k-NN)). The main descriptor, able to classify alone, with a concordance percentage near 80 %, was nArOH, the number of phenolic groups. Therefore, chemicals with phenolic groups have a great potentiality to be endocrine disruptors. The models were externally validated on 87 chemicals (prediction set) not included in the training set. All three methods gave satisfactory prediction results both for training and prediction sets (accuracy range, 82.8–89.7 %); the most accurate model was obtained by the LS-SVM approach. In addition, our models were also applied for screening a big data set from US-EPA (more than 58,000 discrete organic chemicals)

Fig. 6 Classification obtained by consensus of three methods on the screening of more than 50,000 chemicals as potential EDs

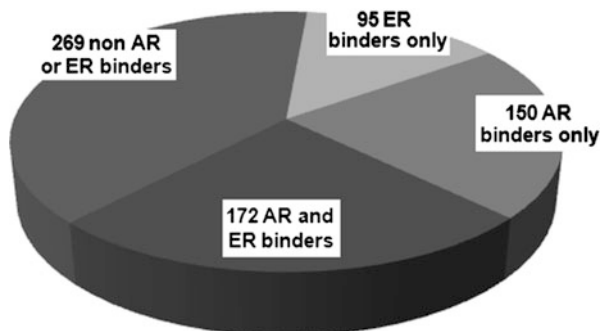


to verify the predictions by consensus: about 76 % of the screened chemicals were predicted not to bind to an estrogen receptor (Fig. 6). The obtained results indicate that the proposed QSAR models could provide a feasible and practical tool for the rapid screening of huge data sets and a very useful prioritization approach for focusing experimental tests only on potential estrogens. In fact, the common 40,300 negative compounds could be excluded from the potential ED list without experiments, because the models have high accuracy and low false-negative rate (3–9 %), while costly and lengthy experimental tests should be concentrated on the common positives (7.8 %, meaning >4000) and eventually on the disagreements, which are predicted as EDs by one classification method and not by the others.

An additional screening work was done to classify a big data set of EDs, both estrogen receptor (ER) binders and androgen receptor (AR) antagonists, mainly aiming to improve the external sensitivity in comparison to the literature models (Vinggaard et al. 2008) and to screen for potential AR binders (Li and Gramatica 2010a). The k-NN, the local lazy method and alternating decision tree methods, and the consensus approach were used to build different models, which improved the sensitivity on external chemicals from 57.1 % (Vinggaard et al. 2008) to 76.4 %. The models' predictive abilities were further validated on a blind data set: the sensitivity was even higher >85 %. Then the proposed classifiers were used: (i) to distinguish a set of AR binders into antagonists and agonists, (ii) to screen a combined estrogen receptor binder database to find out possible chemicals that can bind to both AR and ER, and (iii) to virtually screen our in-house environmental chemical database. The *in silico* screening results suggested that: (i) some compounds can affect the normal endocrine system through a complex mechanism because they could bind both to ER and AR; (ii) new EDs, which are non-ER binders, are recognized *in silico* as binders to AR; and (iii) about 20 % of compounds in a big data set of environmental chemicals are predicted as new AR antagonists. Therefore, the priority should be given to them for experimentally testing their binding activities with AR. The results of this complex screening study are summarized in Fig. 7.

Other QSAR models for screening and prioritization of potential EDs were also developed in Insubria group on brominated flame retardants (Kovarich et al. 2011), on perfluorinated chemicals (Kovarich et al. 2012), as well as on other heterogeneous chemicals (Li and Gramatica 2010b, c).

Fig. 7 Results of classification models for endocrine disruptors and prediction of potential receptor binding (estrogen or androgen) (Permission from Li and Gramatica, *J. Chem. Inf. Mod.*, 2010a)



Ecotoxicity

A lot of chemicals that enter into the environment could have dangerous toxic impact on different wild species. Aquatic organisms, such as algae, crustacean, and fish, are normally used as test organism to determine ecotoxicological data. However, also in relation to ecotoxicity, there is a relevant lack of experimental data; therefore, several chemoinformatic approaches have been, and are continuously, applied to model the existing data and to predict ecotoxicological endpoints for various classes of chemicals of high concern (e.g., Vighi et al. 2001; Salvito et al. 2002; Sanderson et al. 2003, 2004; Öberg 2004; Lo Piparo et al. 2006; Roy 2006; Mazzatorta et al. 2006; Sanderson and Thomsen 2009; Gramatica et al. 2012; Sanderson 2012; Kar and Roy 2012; Cassani et al. 2013, 2014; Singh et al. 2014; Roy et al. 2015). The majority of chemical regulators apply ECOSAR models (US EPA 2012), but these models are not always applicable with satisfactory reliability to all kind of chemicals. This problem, related to applicability domain, has been recently highlighted for an important class of emerging environmental pollutants, the pharmaceuticals (Madden et al. 2009), but it is relevant also for other chemicals, such as personal care products (PCPs). Below, some examples of new ecotoxicity studies, recently performed by the Insubria group on these contaminants of emerging concern for the environment (PCPs and pharmaceuticals), are reported.

Aquatic Toxicity of Personal Care Products

PCP ingredients, widely used all over the world, during the last years became chemicals of increasing environmental concern, mainly because they are detected in water and may harm wildlife. Due to their high structural heterogeneity, to the big number of endpoints and the huge lack of experimental data, *in silico* tools, as QSAR models based on structural molecular descriptors, are highly useful to quickly highlight the most hazardous and toxic compounds, prioritizing existing or even not yet synthesized chemicals. In a recent study (Gramatica et al. 2016a), new externally validated QSAR models, specific to predict acute PCP toxicity in three key organisms of aquatic trophic level, i.e., algae (*Pseudokirchneriella subcapitata*), crustacean (*Daphnia magna*), and fish (*Pimephales promelas*), were

developed according to the OECD principles for the validation of QSARs, using the QSARINS software (Gramatica et al. 2013). These OLS models, based on theoretical molecular descriptors calculated by free PaDEL-Descriptor, selected by genetic algorithm, are statistically robust and externally predictive (CCC range:89–95 %). They were applied to predict the three modeled acute toxicities for 534 PCPs without experimental data, verifying the wide structural applicability domain of each model by the Insubria graphs (more than 95 % of 534 screened PCPs were into the AD). The root mean squared error (RMSE) of each model for chemicals in the prediction sets, not used for model development, was compared with the corresponding RMSE of the ECOSAR models (US EPA 2012): the RMSE of the Insubria models is always around half logarithmic units, while ECOSAR models exhibit always RMSE values higher than one logarithmic unit. Then, according to the consolidated approach in our group, a trend of cumulative acute aquatic toxicity was highlighted by PCA of the three endpoints of ecotoxicity, allowing the ranking of the overall most toxic for all three trophic levels of the aquatic compartment with more than 90 % of data explained variance (Fig. 8). In the biplot of Fig. 8, the most dangerous PCPs, in the right zone, are highlighted as filled squares, using the multi-criteria decision making (MCDM), included in QSARINS-Chem (Gramatica 2014). MCDM is a technique that summarizes the performances of a certain number of criteria simultaneously, as a single number (score) between 0 and 1. This is done associating to every criteria, in our case different predictions for the studied endpoints, a desirability function which values range from 0 to 1 (where 0 represents the less toxic compound and 1 the most toxic), and giving different weights to the selected criteria. The sum of the weights of the criteria must be 1, and in our case, we used the same weight for each criterion: 0.333, which is 1/3 (total/number

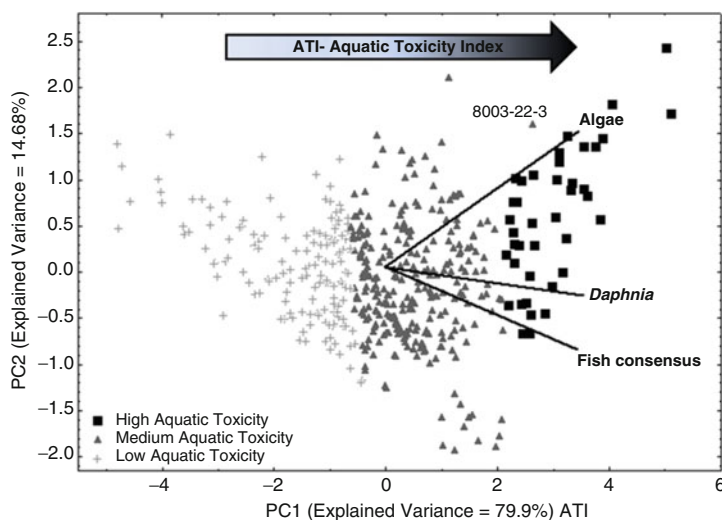


Fig. 8 PCA of experimental and predicted aquatic toxicity data for 484 PCPs and definition of the aquatic toxicity index (ATI) (Permission from RCS: Gramatica et al., Green Chem. 2016b)

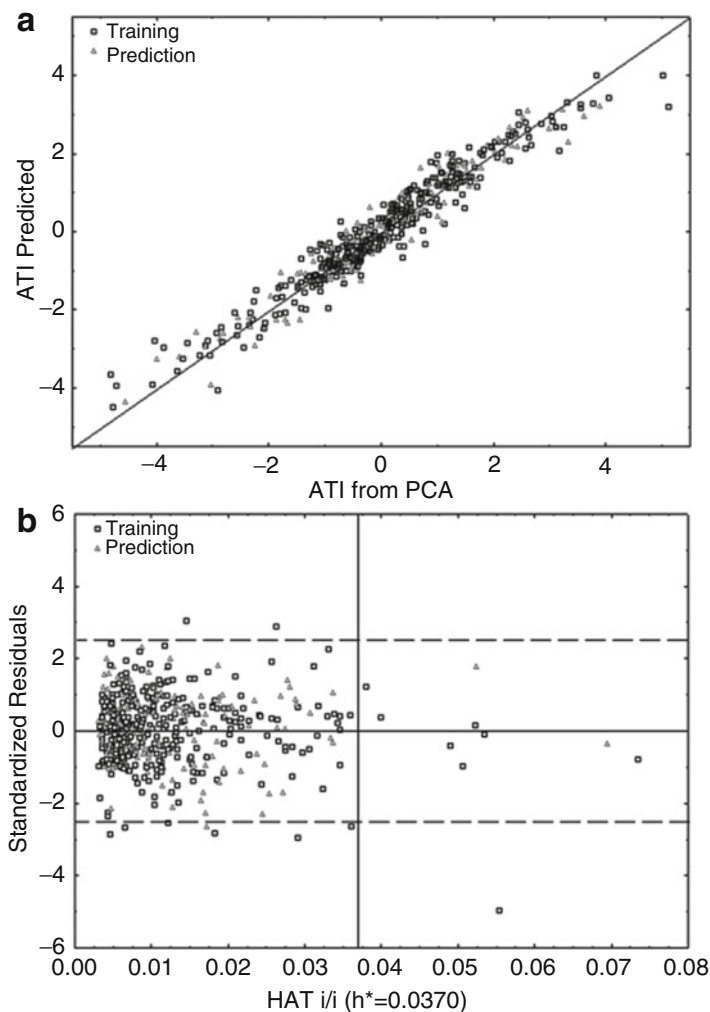


Fig. 9 (a) Scatter plot of the OLS model of ATI; (b) Williams plot for the AD of ATI model (Permission from RCS: Gramatica et al., Green Chem. 2016b)

of criteria). The geometric average of all the values obtained from the desirability functions gives the MCDM value (i.e., the ranking).

A priority list of 40 most hazardous PCPs was then proposed: it includes mainly UV filters (in particular benzotriazoles), some phthalates, and also some fragrances.

Finally, the trend of PC1, which explains about 80 % of the data variance, was proposed as an aquatic toxicity index (ATI). A QSAR model for the prediction of ATI was developed, by OLS on 484 data and using the PaDEL descriptors (scatter plot in Fig. 9a), to be applicable in QSARINS (www.qsar.it) for the *a priori* detection of not yet tested PCPs and also for the preliminary chemical design

of possibly not environmentally hazardous PCPs. The model, with the following equation, is highly predictive and has a good applicability domain verified by the Williams plot of Fig. 9b:

$$\text{ATI} = -14.27 + 0.33 \text{ XlogP} + 18.33 \text{ Mp} + 0.02 \text{ TIC1}$$

$$n_{\text{training}} = 324; R^2 = 0.94; Q^2_{\text{LOO}} = 0.93; Q^2_{\text{LMO}} = 0.93, \text{ RMSE} = 0.40;$$

$$\text{RMSE}_{\text{cv}} = 0.40;$$

$$n_{\text{prediction}} = 160; Q^2_{\text{Fn}} = 0.93 - 0.94; \text{CCC}_{\text{EXT}} = 0.97; \text{RMSE}_{\text{EXT}} = 0.39.$$

The model is mainly driven by LogP descriptor, here represented by XlogP, which has, as expected, a positive sign in the equation, increasing the toxicity of a chemical. The remaining two descriptors, Mp and TIC1, both with a positive sign in the model equation and thus with a positive impact on the cumulative toxicity trend of PCPs, encode respectively for mean atomic polarizability and total information content index (neighborhood symmetry of 1 order). These three descriptors are mainly related to the complexity and the dimension of the molecule, but also to the presence of electronegative atoms, giving higher values for big and halo-substituted compounds.

Aquatic Toxicity of Pharmaceuticals

Similarly to the modeling and screening developed for PCPs, active pharmaceutical ingredients (APIs) were also studied. Pharmaceuticals have become ubiquitous present in the environment; for this reason in 2006, the European Medicines Evaluation Agency (EMA) published guidelines for the environmental risk assessment of human pharmaceuticals. Every environmental risk assessment (ERA) requires large amounts of data for each chemical, but, unfortunately, information on ecotoxicological data is available only for a little percentage of APIs in literature and databases. From literature, we collected and carefully curated data sets for the limited quantity of ecotoxicity data of species at different trophic levels (algae, *Daphnia*, fish). For each species, we developed *ad hoc* QSAR acute toxicity models, based on PaDEL molecular descriptors, using the OLS method and genetic algorithm for variable subset selection in QSARINS software (www.qsar.it). All models are robust ($R^2 > 0.75$, $Q^2_{\text{LOO}} > 0.70$) and externally validated ($\text{CCC} > 0.85$) on different splitting schemes, thus allowing reliable application to new pharmaceuticals. The structural applicability domain (AD) of the proposed models to pharmaceuticals without experimental data was verified and demonstrated to be very high with 74 % of chemicals inside the AD for all the toxicity models. The predictions from Insubria models were always better than those obtained by ECOSAR (US EPA 2012) (average RMSE of 0.5 for Insubria models against an average RMSE of 1.3 for ECOSAR). Moreover, reliable predictions from the different models, applied on a set of more than 1000 pharmaceuticals, were combined by PCA to find an ecotoxicity trend representative of the pharmaceuticals' toxicity in the

whole aquatic ecosystem. This trend, called overall aquatic toxicity index (ATI) for pharmaceuticals, was then modeled by molecular descriptors to obtain a QSAR model useful to highlight, directly from the chemical structure, the pharmaceuticals potentially most hazardous for the environment. This index, and the predictions obtained by it, could be used to refine procedures of input prevention and control at consumer level as well as *a priori* in the rational design of environmentally safer pharmaceuticals (Sangion et al. 2015; Sangion and Gramatica 2016b; Fig. 10).

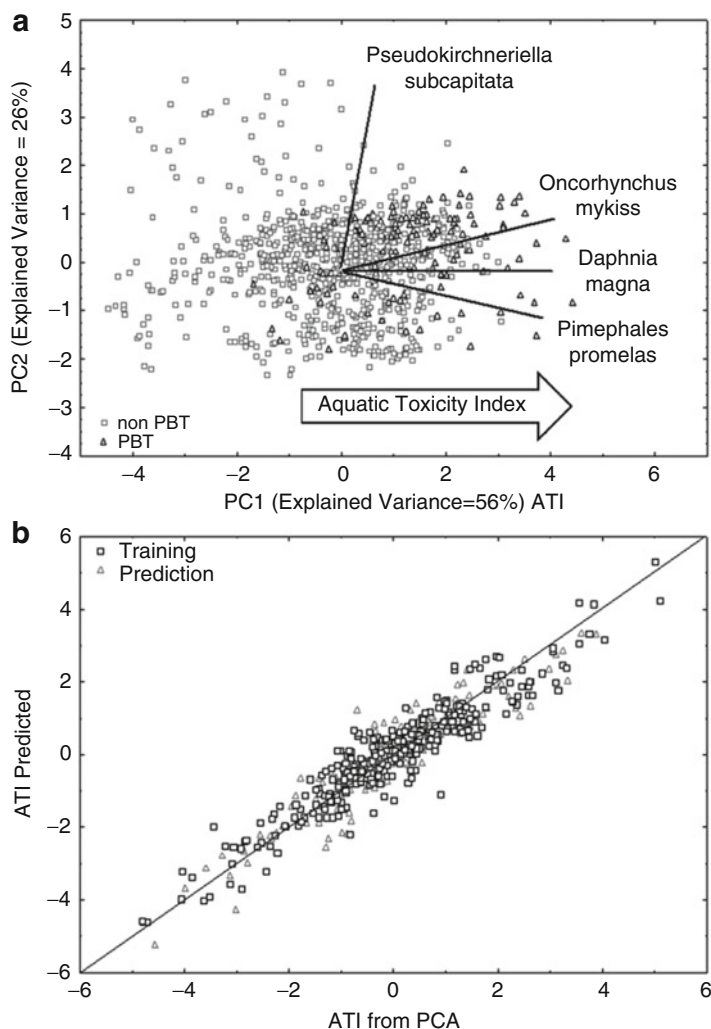


Fig. 10 (a) PCA of experimental and predicted aquatic toxicity data for 984 pharmaceuticals and definition of the aquatic toxicity index (*ATI*); (b) scatter plot of the OLS model of *ATI* for pharmaceuticals

Mammalian Toxicity on Rodents of Fluorinated Chemicals

Fully or highly fluorinated compounds, known as per- and polyfluorinated chemicals (PFCs), are widely distributed and released in the environment, because of their use in different household and industrial products. Some long-chain PFCs are classified as emerging pollutants; in fact, they are found undegraded worldwide and even in polar regions, but their environmental and toxicological effects are mostly not well known. In our lab, under the CADASTER FP7 EU Project (www.cadaster.eu), QSAR models of the mammalian toxicity of PFCs on two species of rodents, rat and mouse, were developed, on two endpoints: LC50 data for inhalation (Bhhatarai and Gramatica 2010) and LD50 oral toxicity (Bhhatarai and Gramatica 2011b).

The OLS models, based on DRAGON molecular descriptors selected by genetic algorithm, were always developed on data sets split in different ways (random, by structural similarity using self-organizing maps (SOM)) to verify the satisfactory external predictivity.

Furthermore, to understand the contribution of each toxicity endpoint and the mutual effect of rodent toxicity, the four models were applied individually to a combined data set of 376 compounds, obtaining predictions for those without experimental data. The compounds were checked for their belonging to the structural AD of each model by Insubria graphs, and 204 compounds were found to be within AD. The PCA of experimental and predicted inhalation and oral toxicity on both rodents (rat and mouse) of PFCs, within the AD of all the four models, allowed to rank PFCs according to their cumulated toxicity on rodents. In Fig. 11 the biplot

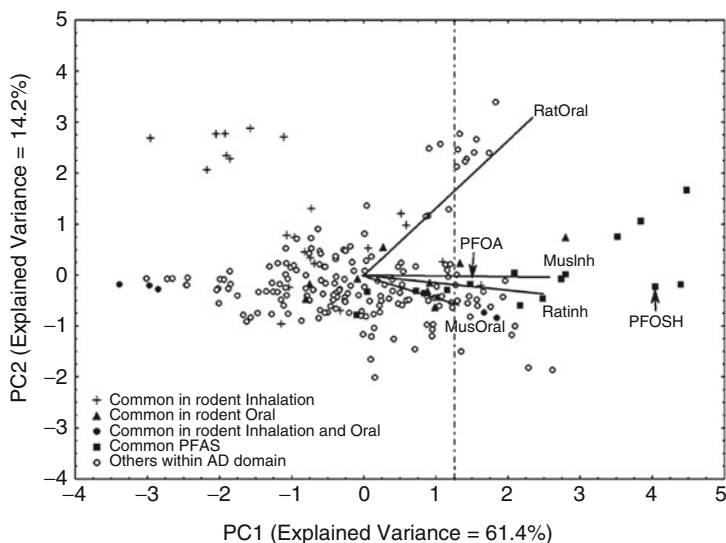


Fig. 11 PCA of inhalation and oral toxicity for two rodent species of 204 PFCs (Permission from Bhhatarai and Gramatica *Mol. Divers.* 2011b)

of this PCA, which explains 75.6 % of data variance, is reported. The compounds with experimental values of rodent inhalation (cross), rodent oral (filled triangles), both rodent inhalation and oral (filled circles), and most common alkylated PFCs (filled squares) are tagged. PC1 ranks the PFCs according to their cumulative rodent toxicity, while PC2 differentiates the PFCs more toxic for rat oral toxicity from those more toxic on the other endpoints. In the right zone of the PCA, the most dangerous and already banned PFCs, as perfluorooctanoic acid (PFOA) and perfluorosulfonic acid (PFSOA), are correctly located: this is a proof of the reliability of the obtained screening and ranking results. Therefore, some other PFCs, which are located in the same zone of concern, should be considered of potential toxicity and prioritized for experimental tests.

Persistence Bioaccumulation Toxicity (PBT)

The chemicals that are contemporaneously persistent, bioaccumulative, and toxic (PBT) are priority chemicals due to the potential risk they pose to humans and ecosystems; therefore, they are considered substances of very high concern (SVHC), which require authorization for use and plan for safer alternatives by REACH. Therefore, in accordance to the precautionary principle, chemicals have to be screened and evaluated for their overall PBT behavior. Unfortunately, for many of the existing substances, even for high production volume (HPV), it is not known, currently, if they could have a potential PBT-like behavior.

Several screening works (Muir and Howard 2006; Howard and Muir 2010; Öberg and Iqbal 2012; Stempel et al. 2012) have highlighted that, among commercial chemicals, many might be PBTs. Therefore, it is evident that not only different approaches for priority setting should be compared to identify the existing potential PBTs but also the “benign by design” approach of green chemistry should be applied for planning the synthesis of safer alternatives to these dangerous compounds (Papa and Gramatica 2010; Stempel et al. 2012; Gramatica et al. 2015, 2016a; Cassani and Gramatica 2015; Sangion and Gramatica 2016a).

Currently, identifying substances as potential PBT or POP candidates relies mainly on determining if specific properties of a chemical exceed threshold values for each property related to PBT behavior (commonly, half-life in various compartments for P, BCF for B, and a number of toxicity evidences for T) (Muir and Howard 2006; Howard and Muir 2010). The most widely used tool for PBT assessment is the US-EPA PBT Profiler (2006), because it is easily applicable from the web (US EPA 2006). The PBT Profiler screens chemicals on the basis of individual P, B, and T properties, calculated by QSAR models and compared to cut off values for each endpoint.

As an alternative approach for PBT prioritization, we developed and proposed a new tool for the screening of chemicals for their potential cumulative PBT behavior (Papa and Gramatica 2010), as an inherent property of a compound that makes it potentially hazardous. PCA of overall half-life in various compartments (the global half-life index, commented above (Gramatica and Papa 2007)), for taking into

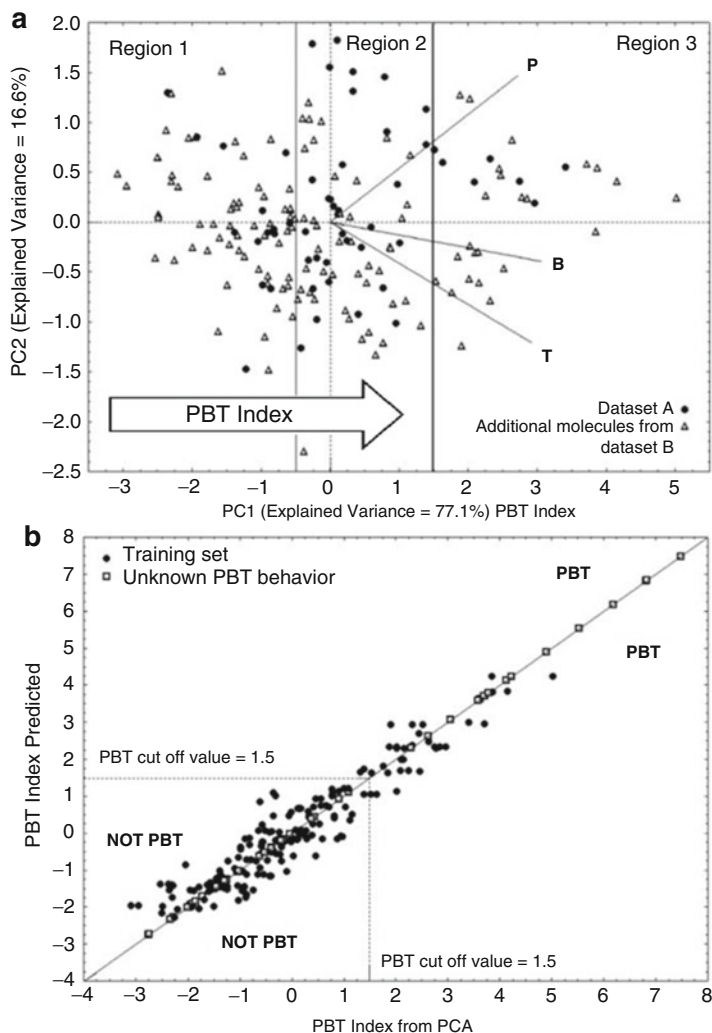


Fig. 12 (a) PCA of persistence, bioaccumulation, and toxicity for 180 heterogeneous chemicals and definition of the PBT Index; (b) scatter plot of the QSAR model of the PBT Index (Permission from Papa and Gramatica, *Green Chem.* 2010)

account the chemical persistence (P), bioconcentration factor (BCF) (Gramatica and Papa 2005) for bioaccumulation (B), and aquatic toxicity on *Pimephales promelas* (Papa et al. 2005) for toxicity (T), allowed to rank 180 representative non-PBT and PBT chemicals according to their cumulative PBT behavior. In fact, since the loadings of the three properties (P, B, and T) are oriented in the same direction in the PCA biplot (Fig. 12a), the PBTs are ranked on the right of the plot. Therefore, the

PC1 score, which explains more than 77 % of the data variance, can be considered as a PBT Index. The cutoff for PBTs was arbitrary fixed at 1.5 score value, by comparison with the thresholds of the criteria normally applied for PBTs and very P very B (vPvB). It is interesting to note that this Index is precautionary; in fact, also some chemicals with only two thresholds of criteria exceeded are located in the right zone (>1.5 of PC1 score): these compounds would be not recognized as PBT by the normally applied criteria (as those in the US-PBT Profiler), while all the chemicals located in region 3 are considered PBTs by the PBT Index.

This Insubria PBT Index was then modeled by OLS QSAR model, using four simple molecular descriptors, with high verified external predictivity ($Q^2_{\text{EXT}} = 80\%$). Thus, our work was based on two different steps according to our consolidated approach: (a) the application of a multivariate tool, such as PCA, for screening chemicals according to their cumulative PBT properties and for the definition of the PC1 score as a PBT Index (biplot in Fig. 12a) and (b) the development of a QSAR model of the PBT Index (scatter plot in Fig. 12b).

Our PBT Index model was recently redeveloped using the PaDEL-Descriptor (Yap 2011), freely calculable online: therefore, the model is now easily applicable in the module QSARINS-Chem of the software QSARINS (Gramatica et al. 2013, 2014) (www.qsar.it).

$$\text{PBT Index} = -1.42 + 0.65 \text{ nX} + 0.22 \text{ nBondsM} - 0.41 \text{ nHBDon_Lipinski} \\ - 0.09 \text{ MAXDP2}$$

$$n_{\text{training}} = 92, R^2 = 0.89, \text{RMSE} = 0.52, Q^2_{\text{LOO}} = 0.88, Q^2_{\text{LMO30\%}} = 0.87, \\ \text{RMSE}_{\text{CV}} = 0.55,$$

$$n_{\text{prediction}} = 88, Q^2_{\text{FI}} = 0.89; \text{CCC}_{\text{EXT}} = 0.94; \text{RMSE}_{\text{EXT}} = 0.49.$$

The descriptors, selected for the best model by the genetic algorithm procedure, are (in descending order of importance) nX (number of halogen atoms), nBondsM (number of multiple bonds), nHDon_Lipinski (number of donor atoms for H bonds), and MAXDP2 (maximal electrotopological positive variation). All of these parameters are mono- or bidimensional and independent of chemical conformation, thus easily calculable from the topological graph (2D sketch) or even from the SMILES code. These variables take into account different chemical properties. The most important descriptors, nX and nBM, which encode for substitution with halogens and unsaturation, are known to increase the PBT behavior of chemicals. On the contrary, MAXDP and nHDon are inversely related to the PBT Index. These last two descriptors are related to a compound's ability to form electrostatic and dipole-dipole interactions, as well as hydrogen bonds in the surrounding media.

Recently, we have extensively applied our PBT Index model for the screening of large data sets of hundreds of heterogeneous chemicals (part 1 of the series

“Early PBT assessment and prioritization of emerging environmental contaminants” in Gramatica et al. 2015), to personal care product (PCP) ingredients (part 2 in Cassani and Gramatica 2015), to flame retardants (part 3 in Gramatica et al. 2016a), and to pharmaceuticals (part 4 in Sangion and Gramatica 2016a). In all these screening studies, we have compared the PBT Index results with those obtained by the US-EPA PBT Profiler, and we have proposed to consider as highly reliable the predictions obtained in agreement by two methods (plot for flame retardants in Fig. 13a).

It is interesting to note that in the screening of flame retardants (FRs), some supposed “safer alternatives” to the banned FRs, which are already in commerce, are detected as intrinsically hazardous for their PBT properties. They are “regrettable substitutions.” If reliable predictive models, based on the chemical structure, would be more often applied *a priori* in a green chemistry approach, from the very beginning of the product development process, it would be possible to avoid the continuous placing on the market, and consequently in the environment, of compounds that will be identified as PBTs, only several years after their use.

Regarding the Insubria PBT Index, we have verified that it is, in the majority of the screenings, more precautionary in highlighting more compounds as PBTs, and the analysis of the disagreements, based on experimental evidences, supports our results for the majority of the cases. It is also important to highlight that the prediction by PBT Index for new chemicals can be verified for the model applicability domain by the Insubria graph (Fig. 13b).

From these screening and prioritization studies, we have verified the reliability of the Insubria PBT Index and, in general, the satisfactory agreement with the PBT Profiler (always >70 %). Main interesting results are:

- (a) In the screening of the Insubria data set of 2780 chemicals of environmental concern (Gramatica et al. 2015), included in the QSARINS-Chem module of the software QSARINS (Gramatica et al. 2013, 2014) (www.qsar.it), the compounds predicted as PBTs by consensus (agreement >82 %) are more than 300.
- (b) In the screening of 534 PCP ingredients, only eight are prioritized as PBTs by consensus of two methods: they are mainly UV filters as benzotriazoles (Cassani and Gramatica 2015).
- (c) In the screening of 128 flame retardants (FRs) (Fig. 13), some already banned and some on the market as substitutes, 30 FRs, which are supposed “safer alternatives,” are predicted as PBTs by both modeling tools in agreement (Gramatica et al. 2016a).
- (d) In the screening of 1267 pharmaceutical ingredients, only 35, of various therapeutic categories, were included in a priority list of potential PBTs, while 83 % of the screened pharmaceuticals are predicted as non-PBTs by consensus (Sangion and Gramatica 2016a).

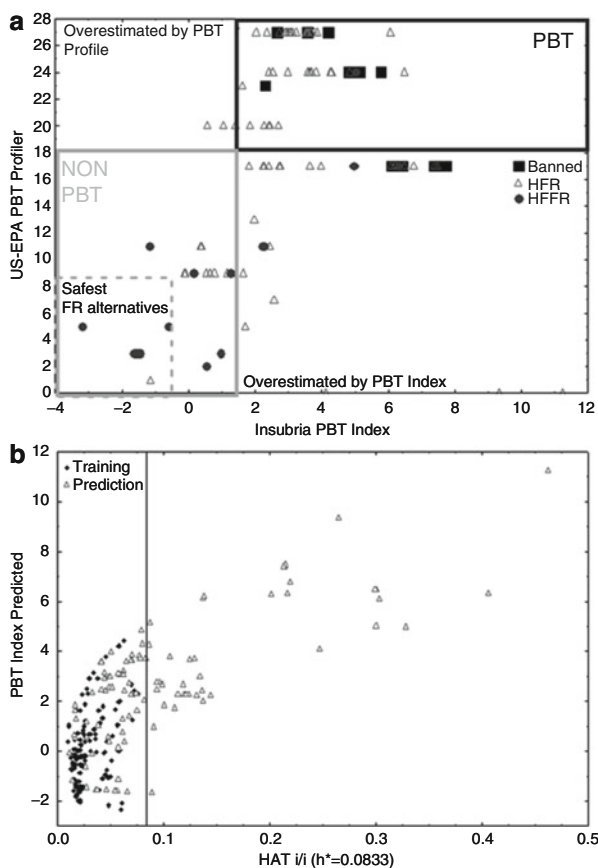


Fig. 13 (a) Graph of the agreement between Insubria PBT Index and US-EPA PBT Profiler for the flame retardants study (in the graph are labeled banned flame retardant, halogen flame retardants (HFRs), and halogen-free flame retardants (HFFRs)) (Permission from Gramatica et al., *J. Hazard. Mater.* 2016a), (b) Insubria graph for the AD of Flame Retardant model

Conclusions

In this chapter, a review of the most recent and significant studies performed for prioritization of chemicals in the Insubria QSAR research group is presented. These studies are based on the fundamental assumption that the hazard of any chemical is an inherent property of the molecular structure; therefore, that QSAR models are an incomparable tool to extract and exploit the information related to any physicochemical or biological property of compounds with available experimental data. The application of externally validated predictive QSAR models to chemicals without experimental data, which are into the model applicability domain, is useful

for highlighting the potentially most hazardous compounds in the screening and priority setting of big data sets of chemicals. This prioritization can allow to concentrate experiments on prioritized chemicals, thus reducing time, costs, and animal test, but also to avoid the synthesis, and introduction to the market and into the environment, of harmful compounds, which could be recognized dangerous only after evidence of human health concerns has been manifested. This is the basis of the “benign by design” approach of green chemistry.

In these studies, we have demonstrated that, taking into account that the chemical behavior derives from a contemporaneous combination of variables, the application of various chemoinformatic methods, such as explorative analysis by PCA and HCA, is useful in ranking and grouping chemicals according to their cumulative behavior, based on more than a single property or activity. In this way, ranking indexes can be proposed for priority-setting purposes and can be also modeled by QSAR to exploit the fundamental information inherent in the chemical structure. The possibility to continuously contaminate the environment with “regrettable substitutions” could be highly reduced if *a priori* screenings, by combining QSAR models and other chemometric approaches, will be more widely applied. Unfortunately, so far, we do not learn enough from the past knowledge, and we do not take advantage, in an extensive way, from the existing information on the inherent hazard that is included in the chemical structure. This should be done, but more expertise on chemoinformatic method for prioritization should be present also at chemical regulation level.

Acknowledgments All my collaborators in the researches performed at the QSAR Research Unit of Insubria University are greatly acknowledged, in particular now Alessandro Sangion for his help in the preparation of the manuscript.

Bibliography

- Arnot, J. A., Brown, T. N., Wania, F., et al. (2012). Prioritizing chemicals and data requirements for screening-level exposure and risk assessment. *Environmental Health Perspectives*, *120*, 1565–1570. doi:10.1289/ehp.1205355.
- Asikainen, A. H., Ruuskanen, J., & Tuppurainen, K. A. (2004). Consensus kNN QSAR: A versatile method for predicting the estrogenic activity of organic compounds in silico. A comparative study with five estrogen receptors and a large, diverse set of ligands. *Environmental Science and Technology*, *38*, 6724–6729. doi:10.1021/es049665h.
- Asikainen, A., Kolehmainen, M., Ruuskanen, J., & Tuppurainen, K. (2006). Structure-based classification of active and inactive estrogenic compounds by decision tree, LVQ and kNN methods. *Chemosphere*, *62*, 658–673. doi:10.1016/j.chemosphere.2005.04.115.
- Balaban, A. T. (1983). Topological indices based on topological distances in molecular graphs. *Pure and Applied Chemistry*, *55*, 199–206.
- Bhatarai, B., & Gramatica, P. (2010). Per- and polyfluoro toxicity (LC50 Inhalation) study in rat and mouse using QSAR modeling. *Chemical Research in Toxicology*, *23*, 528–539. doi:10.1021/tx900252h.
- Bhatarai, B., & Gramatica, P. (2011a). Modelling physico-chemical properties of (benzo)triazoles, and screening for environmental partitioning. *Water Research*, *45*, 1463–1471. doi:10.1016/j.watres.2010.11.006.

- Bhatarai, B., & Gramatica, P. (2011b). Oral LD50 toxicity modeling and prediction of per- and polyfluorinated chemicals on rat and mouse. *Molecular Diversity*, *15*, 467–476. doi:[10.1007/s11030-010-9268-z](https://doi.org/10.1007/s11030-010-9268-z).
- Bonchev, D., & Rouvray, D. H. (1991). *Chemical graph theory*. New York: Gordon & Breach.
- Brown, T. N., & Wania, F. (2008). Screening chemicals for the potential to the persistent organic pollutants: A case study of Arctic contaminants. *Environmental Science and Technology*, *42*, 5202–5209. doi:[10.1021/es8004514](https://doi.org/10.1021/es8004514).
- Browne, P., Judson, R. S., Casey, W. M., et al. (2015). Screening chemicals for estrogen receptor bioactivity using a computational model. *Environmental Science and Technology*, *49*, 8804–8814. doi:[10.1021/acs.est.5b02641](https://doi.org/10.1021/acs.est.5b02641).
- Burden, F. (1989). Molecular-identification number for substructure searches. *Journal of Chemical Information and Computer Science*, *29*, 225–227. doi:[10.1021/ci00063a011](https://doi.org/10.1021/ci00063a011).
- Cassani, S., & Gramatica, P. (2015). Identification of potential PBT behavior of personal care products by structural approaches. *Sustainable Chemistry and Pharmacy*, *1*, 19–27. doi:[10.1016/j.scp.2015.10.002](https://doi.org/10.1016/j.scp.2015.10.002).
- Cassani, S., Kovarich, S., Papa, E., et al. (2013). Daphnia and fish toxicity of (benzo)triazoles: Validated QSAR models, and interspecies quantitative activity–activity modelling. *Journal of Hazardous Materials*, *258–259*, 50–60. doi:[10.1016/j.jhazmat.2013.04.025](https://doi.org/10.1016/j.jhazmat.2013.04.025).
- Cassotti, M., Ballabio, D., Consonni, V., et al. (2014). Prediction of acute aquatic toxicity toward daphnia magna by using the GA-kNN method. *Alternatives to Laboratory Animals*, *42*, 31–41.
- Chirico, N., & Gramatica, P. (2011). Real external predictivity of QSAR models: How to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient. *Journal of Chemical Information and Modeling*, *51*, 2320–2335. doi:[10.1021/ci200211n](https://doi.org/10.1021/ci200211n).
- Chirico, N., & Gramatica, P. (2012). Real external predictivity of QSAR models. Part 2. New intercomparable thresholds for different validation criteria and the need for scatter plot inspection. *Journal of Chemical Information and Modeling*, *52*, 2044–2058. doi:[10.1021/ci300084j](https://doi.org/10.1021/ci300084j).
- Devillers, J., Marchand-Geneste, N., Dore, J. C., et al. (2007). Endocrine disruption profile analysis of 11,416 chemicals from chemometrical tools? *SAR and QSAR in Environmental Research*, *18*, 181–193. doi:[10.1080/10629360701303669](https://doi.org/10.1080/10629360701303669).
- Devillers, J., Bro, E., & Millot, F. (2015). Prediction of the endocrine disruption profile of pesticides. *SAR and QSAR in Environmental Research*, *26*, 831–852. doi:[10.1080/1062936X.2015.1104809](https://doi.org/10.1080/1062936X.2015.1104809).
- Ding, D., Xu, L., Fang, H., et al. (2010). The EDKB: An established knowledge base for endocrine disrupting chemicals. *BMC Bioinformatics*, *11*, S5. doi:[10.1186/1471-2105-11-S6-S5](https://doi.org/10.1186/1471-2105-11-S6-S5).
- Dix, D. J., Houck, K. A., Martin, M. T., et al. (2007). The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicological Sciences*, *95*, 5–12. doi:[10.1093/toxsci/kfl103](https://doi.org/10.1093/toxsci/kfl103).
- Dybdahl, M., Nikolov, N. G., Wedebye, E. B., et al. (2012). QSAR model for human pregnane X receptor (PXR) binding: Screening of environmental chemicals and correlations with genotoxicity, endocrine disruption and teratogenicity. *Toxicology and Applied Pharmacology*, *262*, 301–309. doi:[10.1016/j.taap.2012.05.008](https://doi.org/10.1016/j.taap.2012.05.008).
- EC Regulation. (2006). Registration, evaluation, authorisation and restriction of chemicals (REACH). Regulation (EC) No. 1907/2006 of the European Parliament and of the Council.
- Fang, H., Tong, W. D., Branham, W. S., et al. (2003). Study of 202 natural, synthetic, and environmental chemicals for binding to the androgen receptor. *Chemical Research in Toxicology*, *16*, 1338–1358. doi:[10.1021/tx030011g](https://doi.org/10.1021/tx030011g).
- Gramatica, P. (2007). Principles of QSAR models validation: Internal and external. *QSAR and Combinatorial Science*, *26*, 694–701. doi:[10.1002/qsar.200610151](https://doi.org/10.1002/qsar.200610151).
- Gramatica, P. (2009). Chemometric methods and theoretical molecular descriptors in predictive QSAR Modeling of the environmental behavior of organic pollutants, Chapter 12. In T. Puzyn, J. Leszczynski, & M. T. Cronin (Eds.), *Recent advances in QSAR studies* (pp. 327–366). New York: Springer.

- Gramatica, P. (2014). External evaluation of QSAR models, in addition to cross validation: Verification of predictive capability on totally new chemicals. *Molecular Informatics*, *33*, 311–314. doi:10.1002/minf.201400030.
- Gramatica, P., & Di Guardo, A. (2002). Screening of pesticides for environmental partitioning tendency. *Chemosphere*, *47*, 947–956. doi:10.1016/S0045-6535(02)00007-3.
- Gramatica, P., & Papa, E. (2005). An update of the BCF QSAR model based on theoretical molecular descriptors. *QSAR & Combinatorial Science*, *24*, 953–960. doi:10.1002/qsar.200530123.
- Gramatica, P., & Papa, E. (2007). Screening and ranking of POPs for global half-life: QSAR approaches for prioritization based on molecular structure. *Environmental Science and Technology*, *41*, 2833–2839. doi:10.1021/es061773b.
- Gramatica, P., Pilutti, P., & Papa, E. (2004a). A tool for the assessment of VOC degradability by tropospheric oxidants starting from chemical structure. *Atmospheric Environment*, *38*, 6167–6175. doi:10.1016/j.atmosenv.2004.07.026.
- Gramatica, P., Papa, E., & Battaini, B. (2004b). Ranking and classification of non-ionic organic pesticides for environmental distribution: A QSAR approach. *International Journal of Environmental Analytical Chemistry*, *84*, 65–74. doi:10.1080/0306731031000149732.
- Gramatica, P., Cassani, S., Roy, P. P., et al. (2012). QSAR modeling is not “push a button and find a correlation”: A case study of toxicity of (benzo-)triazoles on algae. *Molecular Informatics*, *31*, 817–835. doi:10.1002/minf.201200075.
- Gramatica, P., Chirico, N., Papa, E., et al. (2013). QSARINS: A new software for the development, analysis and validation of QSAR MLR models. *Journal of Computational Chemistry*, *34*, 2121–2132. doi:10.1002/jcc.23361.
- Gramatica, P., Cassani, S., & Chirico, N. (2014). QSARINS-Chem: Insurbia datasets and new QSAR/QSPR models for environmental pollutants in QSARINS. *Journal of Computational Chemistry*, *35*, 1036–1044. doi:10.1002/jcc.23576.
- Gramatica, P., Cassani, S., & Sangion, A. (2015). PBT assessment and prioritization by PBT Index and consensus modeling: Comparison of screening results from structural models. *Environmental International*, *77*, 25–34. doi:10.1016/j.envint.2014.12.012.
- Gramatica, P., Cassani, S., & Sangion, A. (2016a). Are some “safer alternatives” hazardous as PBTs? The case study of new flame retardants. *Journal of Hazardous Materials*, *306*, 237–246. doi:10.1016/j.jhazmat.2015.12.017.
- Gramatica, P., Cassani, S., & Sangion, A. (2016b). Aquatic Ecotoxicity of Personal Care Products: QSAR models and ranking for prioritization and safer alternatives’ design. (*Green Chemistry, Advance Article, online*. doi: 10.1039/C5GC02818C.)
- Guillen, D., Ginebreda, A., Farre, M., et al. (2012). Prioritization of chemicals in the aquatic environment based on risk assessment: Analytical, modeling and regulatory perspective. *Science of the Total Environment*, *440*, 236–252. doi:10.1016/j.scitotenv.2012.06.064.
- Hansson, S. O., & Rudén, C. (2006). Priority setting in the REACH system. *Toxicological Sciences: An Official Journal of the Society of Toxicology*, *90*, 304–308. doi:10.1093/toxsci/kfj071.
- Hong, H. X., Tong, W. D., Fang, H., et al. (2002). Prediction of estrogen receptor binding for 58,000 chemicals using an integrated system of a tree-based model with structural alerts. *Environmental Health Perspectives*, *110*, 29–36. doi:10.1289/ehp.0211029.
- Howard, P. H., & Muir, D. C. G. (2010). Identifying new persistent and bioaccumulative organics among chemicals in commerce. *Environmental Science and Technology*, *44*, 2277–2285. doi:10.1021/es903383a.
- Jackson, J. E. (1995). Review of a user’s guide to principal components. *Journal of Educational and Behavioral Statistics*, *20*, 105–107. doi:10.2307/1165392.
- Jolliffe, I. T. (2002). *Principal component analysis*. New York: Springer.
- Judson, R., Richard, A., Dix, D. J., et al. (2009). The toxicity data landscape for environmental chemicals. *Environmental Health Perspectives*, *117*, 685–695. doi:10.1289/ehp.0800168.
- Kar, S., & Roy, K. (2012). Risk assessment for ecotoxicity of pharmaceuticals – An emerging issue. *Expert Opinion on Drug Safety*, *11*, 235–274. doi:10.1517/14740338.2012.644272.
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis*. Hoboken: Wiley.

- Kavlock, R., & Dix, D. (2010). Computational toxicology as implemented by the Us Epa: Providing high throughput decision support tools for screening and assessing chemical exposure, hazard and risk. *Journal of Toxicology and Environmental Health, Part B: Critical Reviews*, 13, 197–217. doi:10.1080/10937404.2010.483935.
- Klasmeier, J., Matthies, M., Macleod, M., et al. (2006). Application of multimedia models for screening assessment of long-range transport potential and overall persistence. *Environmental Science and Technology*, 40, 53–60. doi:10.1021/es0512024.
- Knekta, E., Andersson, P. L., Johansson, M., & Tysklind, M. (2004). An overview of OSPAR priority compounds and selection of a representative training set. *Chemosphere*, 57, 1495–1503. doi:10.1016/j.chemosphere.2004.07.056.
- Kovarich, S., Papa, E., & Gramatica, P. (2011). QSAR classification models for the prediction of endocrine disrupting activity of brominated flame retardants. *Journal of Hazardous Materials*, 190, 106–112. doi:10.1016/j.jhazmat.2011.03.008.
- Kovarich, S., Papa, E., Li, J., & Gramatica, P. (2012). QSAR classification models for the screening of the endocrine-disrupting activity of perfluorinated compounds. *SAR and QSAR in Environmental Research*, 23, 207–220. doi:10.1080/1062936X.2012.657235.
- Li, J., & Gramatica, P. (2010a). Classification and virtual screening of androgen receptor antagonists. *Journal of Chemical Information and Modeling*, 50, 861–874. doi:10.1021/ci100078u.
- Li, J., & Gramatica, P. (2010b). The importance of molecular structures, endpoints' values, and predictivity parameters in QSAR research: QSAR analysis of a series of estrogen receptor binders. *Molecular Diversity*, 14, 687–696. doi:10.1007/s11030-009-9212-2.
- Li, J., & Gramatica, P. (2010c). QSAR classification of estrogen receptor binders and identification of pleiotropic EDCs. *SAR and QSAR in Environmental Research*, 21, 657–669. doi:10.1080/1062936X.2010.528254.
- Liu, H., Papa, E., & Gramatica, P. (2006). QSAR prediction of estrogen activity for a large set of diverse chemicals under the guidance of OECD principles. *Chemical Research in Toxicology*, 19, 1540–1548. doi:10.1021/tx0601509.
- Liu, H., Papa, E., Walker, J. D., & Gramatica, P. (2007). In silico screening of estrogen-like chemicals based on different nonlinear classification models. *Journal of Molecular Graphics and Modelling*, 26, 135–144. doi:10.1016/j.jmgm.2007.01.003.
- Lo Piparo, E., Smiesko, M., Mazzatorta, P., et al. (2006). Preliminary analysis of toxicity of benzoxazinones and their metabolites for *Folsomia candida*. *Journal of Agricultural and Food Chemistry*, 54, 1099–1104. doi:10.1021/jf.050916v.
- Madden, J. C., Enoch, S. J., Hewitt, M., & Cronin, M. T. D. (2009). Pharmaceuticals in the environment: Good practice in predicting acute ecotoxicological effects. *Toxicology Letters*, 185, 85–101. doi:10.1016/j.toxlet.2008.12.005.
- Mazzatorta, P., Cronin, M. T. D., & Benfenati, E. (2006). A QSAR study of avian oral toxicity using support vector machines and genetic algorithms. *QSAR and Combinatorial Science*, 25, 616–628. doi:10.1002/qsar.200530189.
- Muir, D. C., & Howard, P. H. (2006). Are there other persistent organic pollutants? A challenge for environmental chemists. *Environmental Science and Technology*, 40, 7157–7166. doi:10.1021/es061677a.
- Öberg, T. (2004). A QSAR for baseline toxicity: Validation, domain of application, and prediction. *Chemical Research in Toxicology*, 17, 1630–1637. doi:10.1021/tx0498253.
- Öberg, T. (2005). A QSAR for the hydroxyl radical reaction rate constant: Validation, domain of application, and prediction. *Atmospheric Environment*, 39, 2189–2200. doi:10.1016/j.atmosenv.2005.01.007.
- Öberg, T. (2006). Virtual screening for environmental pollutants: Structure-activity relationships applied to a database of industrial chemicals. *Environmental Toxicology and Chemistry*, 25, 1178–1183. doi:10.1897/05-326R.1.
- Öberg, T., & Iqbal, M. S. (2012). The chemical and environmental property space of REACH chemicals. *Chemosphere*, 87, 975–981. doi:10.1016/j.chemosphere.2012.02.034.
- OECD. (2004). Principles for the validation, for regulatory purposes, of (Quantitative) Structure-Activity Relationship Models. <http://www.oecd.org/chemicalsafety/risk-assessment/validationofqsarmodels.htm>

- Papa, E., & Gramatica, P. (2008). Screening of persistent organic pollutants by QSPR classification models: A comparative study. *Journal of Molecular Graphics and Modelling*, *27*, 59–65. doi:10.1016/j.jmgm.2008.02.004.
- Papa, E., & Gramatica, P. (2010). QSPR as a support for the EU REACH regulation and rational design of environmentally safer chemicals: PBT identification from molecular structure. *Green Chemistry*, *12*, 836–843. doi:10.1039/B923843C.
- Papa, E., Villa, F., & Gramatica, P. (2005). Statistically validated QSARs, based on theoretical descriptors, for modeling aquatic toxicity of organic chemicals in Pimephales promelas (fathead minnow). *Journal of Chemical Information and Modeling*, *45*, 1256–1266. doi:10.1021/ci050212l.
- Puzyn, T., Gajewicz, A., Rybacka, A., & Haranczyk, M. (2011). Global versus local QSPR models for persistent organic pollutants: Balancing between predictivity and economy. *Structural Chemistry*, *22*, 873–884. doi:10.1007/s11224-011-9764-5.
- Roncaglioni, A., Novic, M., Vracko, M., & Benfenati, E. (2004). Classification of potential endocrine disrupters on the basis of molecular structure using a nonlinear modeling method. *Journal of Chemical Information and Computer Science*, *44*, 300–309. doi:10.1021/ci030421a.
- Roos, V., Gunnarsson, L., Fick, J., et al. (2012). Prioritising pharmaceuticals for environmental risk assessment: Towards adequate and feasible first-tier selection. *Science of the Total Environment*, *421*, 102–110. doi:10.1016/j.scitotenv.2012.01.039.
- Roy, K. (2006). Ecotoxicological modeling and risk assessment using chemometric tools. *Molecular Diversity*, *10*, 93–94. doi:10.1007/s11030-006-9025-5.
- Roy, K., Kar, S., & Das, R. (2015). Understanding the basics of QSAR for applications. In *Pharmaceutical sciences and risk assessment* (1st ed.). Amsterdam/Boston: Academic.
- Saliner, A. G., Netzeva, T. I., & Worth, A. P. (2006). Prediction of estrogenicity: Validation of a classification model. *SAR and QSAR in Environmental Research*, *17*, 195–223. doi:10.1080/10659360600636022.
- Salvito, D. T., Senna, R. J., & Federle, T. W. (2002). A framework for prioritizing fragrance materials for aquatic risk assessment. *Environmental Toxicology and Chemistry*, *21*, 1301–1308. doi:10.1897/1551-5028(2002)021<1301:AFFPFM>2.0.CO;2.
- Sanderson, H. (2012). Challenges and directions for regulatory use of QSARs for predicting active pharmaceutical ingredients environmental toxicity. *Current Drug Safety*, *7*, 309–312.
- Sanderson, H., & Thomsen, M. (2009). Comparative analysis of pharmaceuticals versus industrial chemicals acute aquatic toxicity classification according to the United Nations classification system for chemicals. Assessment of the (Q)SAR predictability of pharmaceuticals acute aquatic toxicity and their predominant acute toxic mode-of-action. *Toxicology Letters*, *187*, 84–93. doi:10.1016/j.toxlet.2009.02.003.
- Sanderson, H., Johnson, D. J., Wilson, C. J., et al. (2003). Probabilistic hazard assessment of environmentally occurring pharmaceuticals toxicity to fish, daphnids and algae by ECOSAR screening. *Toxicology Letters*, *144*, 383–395. doi:10.1016/S0378-4274(03)00257-1.
- Sanderson, H., Johnson, D. J., Reitsma, T., et al. (2004). Ranking and prioritization of environmental risks of pharmaceuticals in surface waters. *Regulatory Toxicology and Pharmacology*, *39*, 158–183. doi:10.1016/j.yrtph.2003.12.006.
- Sangion, A., & Gramatica, P. (2016a). PBT assessment and prioritization of selected Pharmaceuticals. *Environmental Research*, *147*, 297–306. DOI: 10.1016/j.envres.2016.02.021.
- Sangion, A., & Gramatica, P. (2016b). Prioritization of Pharmaceuticals of higher concern for water pollution: structural approaches for modeling and prediction of ecotoxicity (Under revision on Environment International).
- Sangion, A., Cassani, S., Papa, E., & Gramatica, P. (2015). *Identification of potential environmentally hazardous pharmaceuticals by QSAR modeling*. In SETAC Europe 25th annual meeting Barcelona (Spain).
- Scheringer, M., Stempel, S., Hukari, S., et al. (2012). How many persistent organic pollutants should we expect? *Atmospheric Pollution Research*, *3*, 383–391. doi:10.5094/APR.2012.044.
- Schmieder, P., Mekenyan, O., Bradbury, S., & Veith, G. (2003). QSAR prioritization of chemical inventories for endocrine disruptor testing. *Pure and Applied Chemistry*, *75*, 2389–2396. doi:10.1351/pac200375112389.

- Shi, L. M., Fang, H., Tong, W. D., et al. (2001). QSAR models using a large diverse set of estrogens. *Journal of Chemical Information and Computer Science*, 41, 186–195. doi:10.1021/ci000066d.
- Singh, K. P., Gupta, S., Kumar, A., & Mohan, D. (2014). Multispecies QSAR modeling for predicting the aquatic toxicity of diverse organic chemicals for regulatory toxicology. *Chemical Research in Toxicology*, 27, 741–753. doi:10.1021/tx400371w.
- Stenberg, M., Linusson, A., Tysklind, M., & Andersson, P. L. (2009). A multivariate chemical map of industrial chemicals – Assessment of various protocols for identification of chemicals of potential concern. *Chemosphere*, 76, 878–884. doi:10.1016/j.chemosphere.2009.05.011.
- Stempel, S., Scheringer, M., Ng, C. A., & Hungerbühler, K. (2012). Screening for PBT Chemicals among the “Existing” and “New” Chemicals of the EU. *Environmental Science & Technology*, 46, 5680–5687.
- Talete. (2007). DRAGON for Windows (Software for Molecular Descriptor Calculations). Talete srl. Milano, Italy. www.talete.mi.it
- Tong, W. D., Fang, H., Hong, H. X., et al. (2003). Regulatory application of SAR/QSAR for priority setting of endocrine disruptors: A perspective. *Pure and Applied Chemistry*, 75, 2375–2388. doi:10.1351/pac200375112375.
- UNEP (2014). Stockholm convention on persistent organic pollutants (POPs). Stockholm, Sweden.
- US EPA. (2006). PBT profiler; persistent, bioaccumulative, and toxic profiles estimated for organic chemicals on-line. <http://www.pbtprofiler.net/>. Accessed 14 Jan 2016.
- US EPA. (2012). The ECOSAR (ECOLOGICAL Structure Activity Relationship) class program. www.epa.gov/tsca-screening-tools/ecological-structure-activity-relationships-ecosar-predictive-model. Accessed 14 Jan 2016.
- Vighi, M., Gramatica, P., Consolaro, F., & Todeschini, R. (2001). QSAR and chemometric approaches for setting water quality objectives for dangerous chemicals. *Ecotoxicology and Environmental Safety*, 49, 206–220. doi:10.1006/eesa.2001.2064.
- Vinggaard, A. M., Niemela, J., Wedebye, E. B., & Jensen, G. E. (2008). Screening of 397 chemicals and development of a quantitative structure-activity relationship model for androgen receptor antagonism. *Chemical Research in Toxicology*, 21, 813–823. doi:10.1021/tx7002382.
- Vuorinen, A., Odermatt, A., & Schuster, D. (2013). In silico methods in the discovery of endocrine disrupting chemicals. *Journal of Steroid Biochemistry and Molecular Biology*, 137, 18–26. doi:10.1016/j.jsbmb.2013.04.009.
- Wedebye, E. B., Dybdahl, M., Nikolov, N. G., et al. (2015). QSAR screening of 70,983 REACH substances for genotoxic carcinogenicity, mutagenicity and developmental toxicity in the ChemScreen project. *Reproductive Toxicology*, 55, 64–72. doi:10.1016/j.reprotox.2015.03.002.
- Wegmann, F., Cavin, L., MacLeod, M., et al. (2009). The OECD software tool for screening chemicals for persistence and long-range transport Potential. *Environmental Modelling and Software*, 24, 228–237. doi:10.1016/j.envsoft.2008.06.014.
- Yap, C. W. (2011). PaDEL-Descriptor: An open source software to calculate molecular descriptors and fingerprints. *Journal of Computational Chemistry*, 32, 1466–1474.
- Zarfl, C., Hotopp, I., Kehrein, N., & Matthies, M. (2012). Identification of substances with potential for long-range transport as possible substances of very high concern. *Environmental Science and Pollution Research*, 19, 3152–3161. doi:10.1007/s11356-012-1046-2.

Hyun Kil Shin, Young-Mook Kang, and Kyoung Tai No

Contents

Introduction	2266
PART I: Physicochemical Property Prediction Models	2268
Solvation Free Energy (ΔG_{solv})	2268
Ionization Constant ($\text{p}K_a$)	2271
Lipophilicity ($\log P$)	2273
Solubility ($\log S$)	2273
PART II: ADME Prediction Model	2274
Absorption	2278
Distribution	2279
Metabolism	2282
Excretion	2287
PART III: Physiologically Based Pharmacokinetic (PBPK) Models	2288
Gut Physiological Model	2289
Tissue Physiological Model	2291
Liver Physiological Model	2292
Kidney Physiological Model	2294
Summary	2295
Bibliography	2295

H.K. Shin (✉)

Department of Biotechnology, Yonsei University, Seoul, Republic of Korea

e-mail: shkdidrlf@gmail.com

e-mail: shkdidrlf@gmail.com; ktno@bmdrc.org

Y-M. Kang

Bioinformatics and Molecular Design Research Center, Seoul, Republic of Korea

e-mail: youngmook@thylove.org

K.T. No

Bioinformatics and Molecular Design Research Center, Seoul, Republic of Korea

Department of Biotechnology, Yonsei University, Seoul, Republic of Korea

e-mail: ktno@bmdrc.org

Abstract

Since many drug development projects fail during clinical trials due to poor ADME properties, it is a wise practice to introduce ADME tests at the early stage of drug discovery. Various experimental and computational methods have been developed to obtain ADME properties in an economical manner in terms of time and cost. As *in vitro* and *in vivo* experimental data on ADME have accumulated, the accuracy of *in silico* models in ADME increases and thus, many *in silico* models are now widely used in drug discovery. Because of the demands from drug discovery researchers, the development of *in silico* models in ADME has become more active. In this chapter, the definitions of ADME endpoints are summarized, and *in silico* models related to ADME are introduced for each endpoint. Part I discusses the prediction models of the physicochemical properties of compounds, which influence much of the pharmacokinetics of pharmaceuticals. The prediction models of physical properties are developed based mainly on thermodynamics and are knowledge based, especially QSAR (quantitative structure activity relationship) methods. Part II covers the prediction models of the endpoints in ADME which include both *in vitro* and *in vivo* assay results. Most models are QSAR based and various kinds of descriptors (topology, 1D, 2D, and 3D descriptors) are used. Part III reviews physiologically based pharmacokinetic (PBPK) models.

Introduction

ADME describes the disposition and fate of pharmaceutical compounds within an organism, especially in the human body. The primary cause of failure during the drug development phase is poor pharmacokinetics (PK) and toxicity rather than poor efficacy of the candidate compound. In particular, PK has been a major source of failure in drug development (Cheng et al. 2013). Termination of drug development during the clinical phase results in significant losses of both time and cost. The overview of ADME of a chemical compound in the human body is described in Fig. 1. The integration of all components of ADME, and the introduction of time and concentration of chemicals (mother compound and its metabolites) as variables, leads to the development of a PK model. A PK model can be used to determine the proper dose and dosing interval (Gleeson et al. 2011) of a drug to avoid excess administration that can lead to undesirable side effects.

The recent trend in drug discovery has been to optimize both efficacy and PK properties concurrently to reduce the failures aforementioned. Therefore, to obtain ADME or PK properties of a large number of compounds during the early stage of drug discovery, both high-throughput experimental and computational (*in silico*) methods have been developed and have become popular in current drug discovery. Each method has pros and cons. *In vitro* assays are faster than *in vivo* assays; however, the results of both methods do not always correlate. *In vivo* assays performed with surrogate animals have relatively good correlation with results from humans; however, they are slow and expensive. *In silico* models are easy to use,

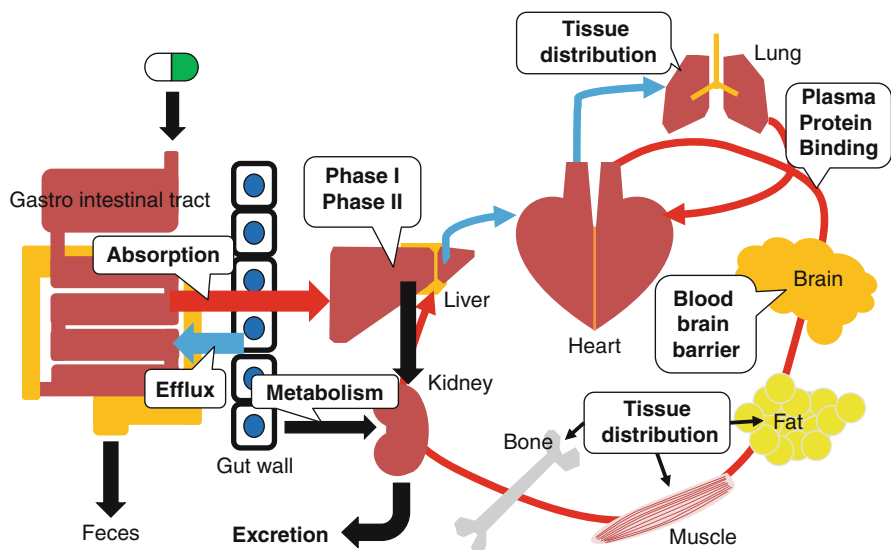


Fig. 1 This figure illustrates ADME for oral administration of drugs. Orally administered tablets or capsules are disintegrated in the gastrointestinal (GI) tract. Dissolved drug molecules can be absorbed through the gut wall, while precipitated drug particles are excreted through the GI tract. When the drug molecules pass through the gut wall, it can be expelled by transporters or metabolized by enzymes. Drugs that successfully pass the gut wall will reach the liver. The liver has many enzymes that metabolize foreign compounds. Reactions can be classified into phase I and phase II according to the enzymes involved. As a result of both reactions, the hydrophilicity of xenobiotics is increased and easily excreted through the kidney. Drugs that are not chemically modified during the metabolism process enter the systemic circulation. A portion of drugs are unable to be delivered to the target organ, tissue, or cell because they are caught by proteins in the blood. The unbound free form of drugs and metabolites can be delivered to the target cell and target biomolecules, and some are rapidly removed through the kidney. Therapeutic effects can be achieved when the concentration of drug is adequate at the site of action after the drugs go through the interplays with the physiological system

fast, and inexpensive, and an important advantage is that one can predict ADME properties using only the molecular structure, even before synthesis of the molecule.

Most *in silico* models developed for ADME endpoint prediction are data-driven (knowledge-based) models that use experimental data as training objects. As a consequence, the quality (predictability and robustness) of the model depends on the quality of the data and the physical feasibility of the model. The model is the mathematical representation of the relationship between the molecular structure and the target property, but in a qualitative or quantitative way. These methods are commonly designated as quantitative structure activity relationships (QSARs), which are widely used for the prediction of chemical toxicity in both humans and in the environment. REACH (Registration, Evaluation, Authorisation and Restriction of Chemical), a regulation of the European Union, adopted to improve the protection of human health and the environment from the risks that can be posed by chemicals while enhancing the competitiveness of the EU chemical industry. It

also promotes alternative methods for the assessment of hazards in substances in order to reduce the amount of animal testing. For this purpose, it is necessary to provide detailed guidance that explains the validation principles of the different types of QSAR models. In an OECD-published book, (Guidance document on the validation of (Quantitative) Structure Activity Relationships [(Q)SAR] Models 2007), the principles for the construction of QSAR models, computational methods, and model validation methods are described in detail.

Due to the increase in demand of QSAR models in various research and regulatory fields, many researchers have been working on developing models for new endpoints, expanding chemical space (application domain), and increasing the accuracy of *in silico* models. The growing market of ADME software is evidence of the increased use of *in silico* models in the pharmaceutical industry.

This chapter will explain the prediction of ADME properties of chemicals in three parts. Part I covers the prediction models of physicochemical properties, which influence much of the pharmacokinetics of pharmaceuticals; Part II covers the prediction models of the endpoints of ADME, which include both *in vitro* and *in vivo* assay results; and Part III covers physiologically based pharmacokinetic (PBPK) models. Details of the methods of model development can be found in Tao et al. (2015) as it is beyond the scope of our work.

PART I: Physicochemical Property Prediction Models

Certain physical properties of compounds have a high correlation with the endpoints of ADME – for example, $\log P_{w/oct}$ (the partition coefficient of a compound between water and 1-octanol). These highly correlated physical properties frequently appear in QSAR models as an important variable, the descriptor. The Lipinski rule of five is a popular rule-based method to check whether chemicals have pharmacological activity. The Lipinski rule is as follows: Molecular weight <500 Da, $\log P_{w/oct} < 5$; with no more than 5 hydrogen bond donors, and no more than 10 hydrogen bond acceptors. The physical properties in the Lipinski rule can be replaced with other physical properties, for example, surface area, polar surface area fraction, solubility, solvation free energy, and so forth, because most physical properties of a compound can be described with other physical properties of the compound. For this reason, the set of variables, or descriptors, of one QSAR model may differ from that of another QSAR model despite both models using the same data set for model construction. Here, we introduce the physical properties that are very important variables to construct ADME endpoint models – solvation free energy, ionization constant, lipophilicity, and solubility of compounds. Available software that predicts physical properties is summarized in Table 1.

Solvation Free Energy (ΔG_{solv})

Solvation free energy is used for the estimation of the ionization constant, lipophilicity, and solubility of compounds and is one of the most frequently appearing

Table 1 Available software for the prediction of physical properties

End points	Type	Program	Service
Solubility	SW	ACD/percepta, preADMET, volsurf+, Stardrop, ADMETPredictor, ADMEWORKS Predictor, QikProp	C
		MarvinSketch(A), EPIsuite(F), ePhysChem(W)	A/F/W
	DB	OCHEM	W
logP	SW	ACD/percepta, preADMET, PrologP, MoKa, StarDrop, ADMETPredictor, QikProp	C
		MarvinSketch(A), EPIsuite(F), ePhysChem(W)	A/F/W
	DB	OCHEM	W
logD	SW	ACD/percepta, PrologD, MoKa, ADMETPredictor	C
		MarvinSketch	A
	DB	OCHEM	W
pK_a	SW	ACD/percepta, pKcalc, MoKa, Jaguar, Epik, ADMETPredictor	C
		MarvinSketch	A
	DB	OCHEM	W

C commercial, F freeware, A academic license, W web service

descriptors in QSAR models. Based on explicit or implicit expressions of the solvent, various methods have been developed for calculating the solvation free energy of compounds and solutes. Since explicit solvent models have excessively high computational costs, less expensive and practical models based on free-energy relationship or continuum solvation method have been intensively developed and widely used.

The solvent-accessible surface area (SASA) model is the simplest and most widely used model for the ΔG_{solv} calculation of large molecules like proteins (Lee and Richards 1971). SASA is the surface of a molecule that is accessible to a solvent. The solvation free energy of a solute is given by

$$\Delta G_{solv} = \sum_i \sigma_i ASA_i \quad (1)$$

where σ_i is the empirical parameter of the solvation free energy of the i th atom and ASA_i is the SASA of the i th atom.

The linear free energy relationship (LFER)-based model is simple and described in a realistic physical way. The linear term contributing to free energy is described by the product of the properties of the solvent and the solute of a solution. The Abraham LFER (Abraham et al. 1988) model is expressed as:

$$\log SP = c + eE + sS + aA + bB + IL \quad (2)$$

where SP is a free energy-related physical property; E is excess molar refraction; S is the ability of a solute to stabilize a neighboring dipole; A and B are the

solute's effective hydrogen bond acidity and basicity, respectively; and L is the gas–liquid partition constant of the solute. The linear coefficients, c , e , s , a , b , and l are determined with experimental data through an optimization procedure, and the coefficients represent the magnitude of interaction of the solvent with the solute, for example, “sS” represents the magnitude of the contribution to $\log SP$ by the interaction of the solvent with solute dipole, “S.”

Continuum solvation models (Cramer and Truhlar 1995) calculate solvation free energy considering the following three terms:

$$\Delta G_{\text{solv}} = \Delta G_{\text{cav}} + \Delta G_{\text{vdw}} + \Delta G_{\text{pol}} \quad (3)$$

ΔG_{cav} is the free energy to form a cavity in the solvent with the SASA of solute, ΔG_{vdw} is the van der Waals interaction between the solute and solvent, and ΔG_{pol} is the electrostatic interaction between the polar solute with polarizable solvent. Continuum solvation models describe a solute as a collection of point charges, and solvent as a dielectric continuum medium. The charge distribution of the solute electrically polarizes the solvent, and the electric field of induced dipole of the polarized solvent interacts with the dipole of the solute. In the classical approaches, the Poisson-Boltzmann (PB) equation is used to calculate ΔG_{pol} .

In quantum mechanical approaches, the self-consistent reaction field (SCRF) (Tomasi et al. 2005) is applied to account for the mutual polarization. It is further applied in a polarizable continuum model (PCM) (Mennucci 2012). PCM uses SCRF equations in a boundary element problem with apparent surface charges spread on the cavity surface. SMx models (Marenich et al. 2013) calculate solvation energy with this equation:

$$\Delta G_{\text{solv}} = \Delta E_E + G_P + G_{\text{CDS}} + \Delta G_N + \Delta G_{\text{conc}} \quad (4)$$

ΔE_E is the solute's internal electronic energy change upon transfer of the solute from the gas to the liquid. G_P is the polarization energy of the solute-solvent system from introduction of the solute; it is defined by the general Born (GB) approximation. G_{CDS} is the energy for cavitation, dispersion, and solvent structure. ΔG_N is energy variation due to change in nuclear coordinates. ΔG_{conc} is the difference in concentration between the gas phase standard state and the solution phase.

In the conductor-like screening model for realistic solvation (COSMO-RS) (Klamt et al. 2010), the polarization charges of the continuum are estimated from scaled-conductor approximation. COSMO-RS derives the polarization charges from the continuum, caused by distribution of screening charges (σ -profile) of the solute, from a scaled-conductor approximation. In the model, the hydrogen bonding term $E_{hb}(\sigma, \sigma')$ is considered as one component of the interaction between solute and solvent,

$$E_{hb}(\sigma, \sigma') = c_{hb}(T) \max[0, \sigma_{acc} - \sigma_{hb}] \min[0, \sigma_{don} + \sigma_{hb}] \quad (5)$$

where σ and σ' refer to the screening charge densities of the two surface patches in contact, and σ_{acc} and σ_{don} are the screening charge densities of the hydrogen bond acceptor and donor, respectively. The hydrogen-bonding threshold σ_{hb} and the prefactor c_{hb} are adjustable parameters.

The generalized solvation free energy density (G-SFED) model (Lee et al. 2013) calculates ΔG_{solv} as:

$$\Delta G_{solv} = \Delta G_{inter} + \Delta G_{HB} + \Delta G_{cav} \quad (6)$$

The solute molecule is considered as a point charge, and SASA is discretized into small polarizable fractions. The summation of interactions between SASA and solute gives interaction energy (ΔG_{inter}) between the solvent and solute. ΔG_{HB} is the contribution of a hydrogen bond described with the acidity and basicity of both solute and solvent. ΔG_{cav} is calculated based on the surface tension of the solvent.

Ionization Constant (pK_a)

pK_a is the negative logarithm of the acid dissociation constant, K_a , which is a quantitative measurement of the strength of acid in solution. pK_a prediction models have been developed by many researchers because it is one of the most important features of molecules in water solution, e.g., proteins in water. Broadly, pK_a can be estimated using the thermodynamic cycle or QSAR. In the thermodynamic approach (Fig. 2), the Gibbs free energy changes of a system is used as

$$pK_a = \Delta G_{aq}/2.303RT \quad (7)$$

In the thermodynamic cycle, ΔG_{aq} is calculated through

$$\begin{aligned} \Delta G_{aq} = & [\Delta G_{acid \text{ or base}} - \Delta G_{water} (H - OH_2^+)] \\ & + [\Delta G_{solv} (A^- \text{ or B}) - \Delta G_{solv} (HA \text{ or } BH^+)] \\ & + [\Delta G_{solv} (H_3O^+) - \Delta G_{solv} (H_2O)] \end{aligned} \quad (8)$$

The first term, $\Delta G_{acid} - \Delta G_{water} (H - OH_2^+)$, is the free energy of bond dissociation. The second and third terms relate to the solvation energy difference between neutral and ionized species.

Ionization of a molecule in the gas phase can be predicted using gas phase acidity (GPA) and gas phase basicity (GPB) prediction models (Gasteiger and Hutchings 1983). Several models were developed with expanded application domains. Recently, a GPA/B prediction model was developed and published with a brief introduction to other GPA/B models (You et al. 2014).

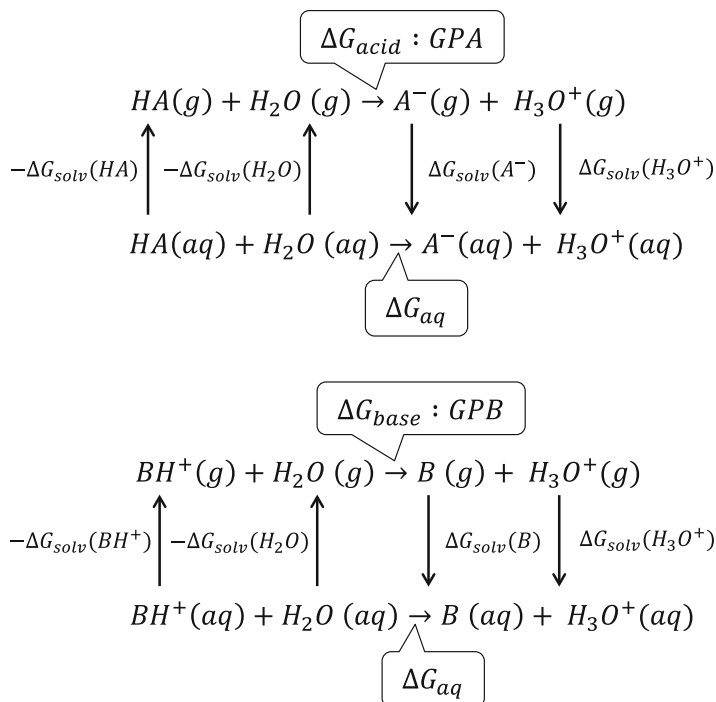


Fig. 2 The thermodynamic cycle is applied to calculate the Gibbs free energy difference in proton dissociation in water. GPA and GPB is gas phase acidity and basicity

LFER-based models are developed to predict pK_a using an experimental database. This approach is based on the discovery that similar molecular fragments have similar pK_a values. Databases that contain the pK_a of simple molecules and their molecular fragments are used to predict pK_a by searching the experimental pK_a of the closest molecular structure. Recent work in this approach used ab initio equilibrium bond length as a descriptor. In this model, the closest neighbor of the input molecule is searched in the database, then pK_a is predicted based on the experimental pK_a of the closest neighbor and bond lengths of the input molecule (Alkorta and Popelier 2015).

QSAR models for pK_a prediction mainly introduce QM descriptors and some thermodynamic descriptors for building models – HOMO, polarizability, free energies, relative proton transfer energy, minimum surface local ionization energy, partial atomic charges, group philicity, and molecular electrostatic potential (Svobodova Varekova et al. 2011). Simulations Plus, commercial software, updated their pK_a prediction model with Bayer's data to expand the application domain of QSAR. This model covers both public and industrial domains (Fraczkiewicz et al. 2015) and gives generally good prediction results on test sets. Several *in silico* models are now available. Details of pK_a prediction software can be found in the paper by Fraczkiewicz (Fraczkiewicz 2013).

Lipophilicity (logP)

To predict lipophilicity, three main approaches were used – group contribution methods, QSAR, and solvation energy-based models. Group contribution methods use additivity contribution approximation as follows:

$$\log P_{w/oct} = \sum a_i f_i + \sum c_i f_i + \text{constant} \quad (9)$$

f_i is the number of i th fragments in a molecule, and a_i is the coefficient for i th fragments. These coefficients show the magnitude of contribution of each fragment of the molecule to $\log P_{w/oct}$. c_i is the correcting factor for the i th fragment. These group contribution methods are broadly applied in available software.

QSAR approaches are applied with various descriptors, such as lattice energy, QM descriptors, Abraham's descriptors, and so forth. Details of descriptors and methods applied to predict $\log P$ were reviewed by Mannhold et al. (Mannhold et al. 2009). The solvation energy-based model predicts $\log P$ by estimating the solvation Gibbs free energies of a molecule in two solvents. The $\log P_{w/oct}$ is calculated directly from the difference between solvation free energies in water (ΔG_w) and in 1-octanol (ΔG_{oct}):

$$\log P_{w/oct} = -(\Delta G_{oct} - \Delta G_w) / 2.303RT \quad (10)$$

In physiological conditions, pH differs from compartment to compartment in the body. The Henderson-Hasselbach (HH) (Hansen et al. 2006) equation is used for the calculation of the $\log P$ of ionizable species, $\log D$,

$$\text{pH} = \text{p}K_a + \log_{10} \frac{[A^-]}{[HA]} \quad (\text{for acids}) \quad (11)$$

using the HH equation, $\log D$ is expressed as,

$$\log D = \log P_{w/oct} + \log \left[1 / \left(1 + 10^{\text{pH} - \text{p}K_a(\text{acid})} \right) \right] \text{ pag} \quad (12)$$

Solubility (logS)

For solubility prediction models, group contribution methods are broadly used for $\log P$ calculation. Recently, a deep learning algorithm has been adopted to build a solubility prediction model (Lusci et al. 2013). The thermodynamic approach to predict solubility requires the following additional parameters: crystalline molar volume (V_m) (Skyner et al. 2015), vapor pressure (VP) (Thompson et al. 2003), and melting point (MP) (Ran and Yalkowsky 2001).

A model using V_m calculates ΔG_{solv} with intrinsic solubility (S_o),

$$\Delta G_{solv} = -RT \ln (S_o V_m) \quad (13)$$

A model using vapor pressure,

$$S_o = (P_{VP} / P_{ideal}) \cdot \exp (-\Delta G_w / RT) \quad (14)$$

P_{ideal} is the pressure of an ideal gas (24.45 atm). The model using MP is known as the general solubility equation (GSE), and it uses estimated $\log P_{w/oct}$ with the model.

$$\log S_o = 0.5 - 0.01 (MP - 25) - \log P_{w/oct} \quad (15)$$

Solubility of ionized species is also calculated using the HH equation (Hansen et al. 2006).

$$\log S_{pH} = \log S (1 + 10^{pH-pK_a}), \log S_{pH} = \log S (1 + 10^{pK_a-pH}) \quad (16)$$

and the HH equation is applied twice in the equation for an ampholyte

$$\log S_{pH} = \log S \left(1 + 10^{pH-pK_a(\text{acid})} + 10^{pK_a(\text{base})-pH} \right) \quad (17)$$

The HH equation is applied to describe the ionization of species, although accuracy issues have been raised (Bergstrom et al. 2004). GSE also can be used to predict $\log S_{pH}$ by applying logD.

$$\log S_{pH} = 0.5 - 0.01 (MP - 25) - \log D \quad (18)$$

PART II: ADME Prediction Model

In the past several decades, a significant amount of in vitro and in vivo assay data has been accumulated as a byproduct of pharmaceutical development. This accumulated data enabled the development of ADME prediction software with high reliability and accuracy. In Part II, we will explain the endpoints of ADME and introduce prediction models of each endpoint. Various ADME prediction softwares are currently available thanks to the efforts of many model and software development groups. A list of widely used ADME prediction software is presented in Table 2.

Table 2 Available software for ADME endpoints

Endpoints	Type	Program	Service
Absorption			
PAMPA	SW	MolCode toolbox	C
caco-2	SW	preADMET, MolCode toolbox, Discovery Studio, volsurf+, QikProp	C
		preADMET, admetSAR	W
	DB	the ADME database (100)	W
MDCK	SW	preADMET, QikProp, ADMETpredictor	C
		preADMET	W
HIA	SW	ACD/percepta, preADMET, MolCode toolbox, Discovery Studio, ADMEWORKS Predictor, Stardrop	C
		PK/DB, admetSAR, preADMET	W
	DB	the ADME database (647), PK/DB (687)	W
Skin permeability	SW	preADMET, ADMETpredictor	C
		preADMET	W
Corneal permeability	SW	ADMETpredictor	C
BA	SW	ACD/percepta, Impact-F	C
		PK/DB	W
	DB	Pact-F the ADME database (1013), PK/DB (660)	C W
Distribution			
BBB	SW	ACD/percepta, preADMET, MolCode toolbox, Discovery Studio, ADMEWORKS Predictor, volsurf+, Stardrop, QikProp, ADMETpredictor	C
		PK/DB, admetSAR, preADMET	W
	DB	the ADME database (109), Chembench (159), PK/DB (200)	W
VD	SW	ACD/percepta, volsurf+, ADMETpredictor	C
	DB	PK/DB (738)	W
PPB	SW	ACD/percepta, preADMET, MolCode toolbox, Discovery Studio, volsurf+, stardrop, QikProp, ADMETpredictor	C
		PK/DB	W
	DB	Chembench (1244), PK/DB (726)	W
Partition coefficient	SW	MolCode toolbox, ADMETpredictor	C

(continued)

Table 2 Available software for ADME endpoints

Endpoints	Type	Program	Service
<i>Metabolism/Excretion</i>			
P-gP inhibition	SW	ACD/percepta, ADMETpredictor	C
		preADMET, PK/DB, admetSAR	W
P-gP substrates	DB	the ADME database (1302)	W
		ADMEWORKS Predictor	C
Regioselectivityphase 1	SW	Chembench, admetSAR	W
		Chembench (195)	W
Regioselectivity 1A2	SW	QikProp	C
		FAME	W
Regioselectivity 2A6	SW	ACD/percepta, Stardrop, ADMETpredictor	C
		RS-webpredictor	W
Regioselectivity 2B6	SW	RS-webpredictor, ADMETpredictor	W
Regioselectivity 2C8	SW	Stardrop, ADMETpredictor	C
		RS-webpredictor	W
Regioselectivity 2C9	SW	ACD/percepta, Stardrop, ADMETpredictor	C
		RS-webpredictor, SMARTcyp	W
Regioselectivity 2C19	SW	ACD/percepta, Stardrop	C
		RS-webpredictor	W
Regioselectivity 2D6	SW	ACD/percepta, Stardrop, ADMETpredictor	C
		RS-webpredictor, SMARTcyp	W
Regioselectivity 2E1	SW	Stardrop, ADMETpredictor	C
		RS-webpredictor	W
Regioselectivity 3A4	SW	ACD/percepta, Stardrop, ADMETpredictor	C
		RS-webpredictor, SMARTcyp	W
Regioselectivity phase 2	SW	MEXAlert	C
		FAME	W
1A2 inhibitor	SW	ACD/percepta, ADMETpredictor, MetaDrug	C
		ePhysChem, admetSAR	W
2C9 inhibitor	SW	ACD/percepta, ADMETpredictor, MetaDrug	C
		preADMET, ePhysChem, admetSAR, preMetabo	W
2C19 inhibitor	SW	ACD/percepta, ADMETpredictor, MetaDrug	C
		preADMET, ePhysChem, admetSAR, preMetabo	W

(continued)

Table 2 Available software for ADME endpoints

Endpoints	Type	Program	Service
2D6 inhibitor	SW	ACD/percepta, ADMETpredictor, MetaDrug	C
		preADMET, eADMET, admetSAR, preMetabo	W
3A4 inhibitor	SW	ACD/percepta, ADMEWORKS Predictor, ADMETpredictor, MetaDrug	C
		preADMET, eADMET, admetSAR, preMetabo	W
1A2 substrates	SW	ACD/percepta, ADMETpredictor, MetaDrug	C
2A6 substrates	SW	ADMETpredictor	C
2B6 substrates	SW	ADMETpredictor, MetaDdrug	C
2C8 substrates	SW	ADMETpredictor	C
2C9 substrates	SW	ACD/percepta, ADMETpredictor	C
		admetSAR	W
2C19 substrates	SW	ACD/percepta, ADMETpredictor	C
2D6 substrates	SW	ACD/percepta, Discovery Studio, ADMETpredictor, MetaDrug	C
		preADMET, admetSAR, preMetabo	W
2E1 substrates	SW	ADMETpredictor	C
3A4 substrates	SW	ACD/percepta, ADMETpredictor, MetaDrug	C
		preADMET, admetSAR, preMetabo	W
1A2 Km	SW	ADMETpredictor	C
2C9 Km	SW	ADMETpredictor	C
2C19 Km	SW	ADMETpredictor	C
2D Km	SW	ADMEWORKS Predictor	C
2D6 Km	SW	ADMETpredictor	C
3A4 Ki	SW	ADMEWORKS Predictor	C
3A4 Km	SW	ADMEWORKS Predictor, ADMETpredictor	C
Vmax	SW	ADMETpredictor	C
Metabolic stability	SW	volsurf+	C
Metabolites (phase I)	SW	Meteor Nexus, MetaSite, MetaDrug	C
		Metaprint2D-react	W
	DB	webmetabase, Fujitsu ADME database	C
		PK/DB (295)	W

(continued)

Table 2 Available software for ADME endpoints

Endpoints	Type	Program	Service
Metabolites (phase II)	SW	MetabolExpert, MetaSite, MetaDrug, Meteor Nexus	C
	SW	Metabolizer, Metaprint2D-react	A/W
	DB	webmetabase, Fujitsu ADME database	C
OATP1B1 inhibition	SW	ADMETpredictor	C
Intrinsic clearance	SW	ADMETpredictor	C
Half-life	DB	PK/DB (770)	W
UGT substrate	SW	ADMETpredictor	C
Phase II substrate/inhibitor	SW	preMetabo	W
OATP, OAT, OCT binder/nonbinder	SW	preMetabo	W
SOM	SW	PASS, ADMETpredictor	C
		Xenosite, Metaprint 2D	W
Renal clearance	DB	PK/DB (360)	W

Absorption

Absorption in drug PK refers to the process of how drugs go through the organs of the body to reach the systemic circulation. Drugs have various routes of administration – oral, intravenous, intramuscular, intraperitoneal, ocular, nasal, rectal, intratracheal, and subcutaneous. In this chapter, we focus mainly on the oral administration route since it is the most common route of administration.

The most popular in vitro absorption assays are the parallel artificial membrane permeability assay (PAMPA), the cell-based assay, and the human intestinal absorption (HIA) assay. Several QSAR models have been developed using experimental data from in vitro methods. Since PAMPA uses artificial lipid membranes, it is convenient and offers good reproducibility; therefore, it is broadly used for absorption tests. However, the membrane does not have transporters or drug metabolizing enzymes (DMEs). Thus, it is an adequate assay method for testing passive transport but it cannot be used for the prediction of active transport. Since transporters and DMEs are expressed on the cell membrane, cell-based assays, to some extent, reflect active transport as well as passive transport. Caco-2 cells derived from human colon adenocarcinomas are broadly used for drug absorption assays. Caco-2 cells exhibit transporters like OCT, OAT, and PEPT1 and CYP3A enzymes. Another popular cell-based assay is the Madin-Darby canine kidney (MDCK) cell assay. Absorption (or permeation) is represented by the permeability coefficient (P_{eff}) defined as,

$$P_{eff} = C_t / (A \cdot C_i \cdot t) \quad (19)$$

P_{eff} is obtained by dividing the amount of drug transported through the membrane (C_t) by the product of the surface area (A) of the membrane monolayer, the time (t) taken until the system reaches steady state, and the initial concentration of the drug (C_i).

HIA is measured by dividing the absorbed drug by the drug dose. In vitro cell assays, such as those in caco-2 cells, show good correlation with human absorption. In vivo animal studies are also used to predict HIA.

$$\text{HIA (\%FA)} = \text{total mass absorbed/dose} \quad (20)$$

Bioavailability (BA) is the fraction of drug absorbed into the systemic circulation. BA is calculated based on the blood concentration profile obtained through in vivo testing.

$$\text{BA (\%F)} = [AUC_{\text{oral}} \times \text{dose}_{IV} / AUC_{IV} \times \text{dose}_{\text{oral}}] \times 100 \quad (21)$$

AUC is the area under the curve of the blood concentration profile in oral administration (AUC_{oral}) and in intravenous (IV) administration (AUC_{IV}), respectively, and dose_{IV} and $\text{dose}_{\text{oral}}$ represent the dose with IV and oral administration, respectively.

The major method used for the prediction of absorption is QSAR. Various regression QSAR models have been developed to predict the PAMPA assay (Akamatsu et al. 2009), cell-based assay (Hou et al. 2006), and BA (Hou et al. 2009). Since the absorption of cell-based assay results vary depending on how the cells are cultured, these results are not suitable for quantitative model construction; therefore, the assay results are converted to qualitative data with certain criteria in order to combine data from several laboratories (Pham The et al. 2011). The interaction of drugs with the membrane is studied and applied in the MDCK permeation prediction model (Chen et al. 2005). Information from drug interactions with transporters (Kim et al. 2014a) and DMEs (Xu et al. 2012) modifies the results of BA prediction. For BA prediction, Zhu (Zhu et al. 2011) reviewed the relationship between physical properties and oral BA. Not only physicochemical and structural properties but also in vitro data (in vitro permeability, intrinsic clearance, and animal data) and *in silico* prediction results (permeability and intrinsic clearance) were used to build the model (Wang and Hou 2015).

Distribution

Distribution in PK is defined as the transportation of drugs from one tissue/organ to another tissue/organ. Among all tissues, blood delivers drugs mainly into other tissues in affinity-dependent manner. Generally, drugs interact with proteins in the blood and membranes of the tissues. These interactions are crucial for determining the amount of drug delivered to the site of action; therefore, assay systems are

designed to measure free drug in the blood and in the volume of tissue, where drugs can be transferred. These two measurements are the parameters that predict drug distribution.

The blood brain barrier (BBB) is the set of endothelial cells that protect the brain by allowing the absorption of only those chemicals that are necessary for brain metabolism. The BBB fills cell margins with tight junctions; therefore, it blocks penetration of hydrophilic or charged compounds. Distribution of drugs into the brain is an important focus in the discovery of drugs that target the central nervous system (CNS). QSAR and molecular modeling have been applied to predict BBB penetration.

Plasma Protein Binding (PPB)

The measured percentage of a drug bound to plasma protein (%PPB) can be converted to the binding parameter, logK,

$$\log K = \log \left(\frac{\%PPB}{101 - \%PPB} \right) = \log \%bound - \log \%free \quad (22)$$

The constant value 101 is used based on the definition that log K is two for a 100 % bound compound. Most QSAR models for PPB are focused on human serum albumin (HSA) binding because serum albumin is the most abundant plasma protein in human blood. The binding parameter for HSA ($\log K_{HSA}$) can be obtained with the above equation as,

$$\log K_{HSA} = \log K - \log [HSA] = \log K + 3.22 \quad (23)$$

The value of 3.22 is calculated based on the concentration of HSA, $0.6 \times 10^{-3} M$. Normally, three PPB parameters (%PPB, logK, and $\log K_{HSA}$, measured by high performance affinity chromatography specifically for HSA) are used as endpoints in the QSAR model. Various QSAR models have been developed and details of the descriptors of these models are explained by Colmenarejo (Colmenarejo 2007). Even though the HSA binding prediction model uses a major part of the QSAR model, few α_1 -acid glycoprotein (AGP) binding prediction models have been developed. Molecular modeling is also applied to study drug binding with HSA and AGP using crystal structures. Details of the QSAR model description and molecular modeling with available crystal structures for HSA and AGP are reviewed by Lambrinidis (Lambrinidis et al. 2015).

Volume of Distribution (VD)

VD is the ratio between the total amount of drug in the body and the concentration of drug in blood plasma. The unit of VD is the liter. Several types of VD have been defined. The volume of the central compartment (VD_c) is defined as:

$$VD_c = D/C_0 \quad (24)$$

D is the IV administered dose, and C_0 is the initial plasma concentration. In a system that has multiple compartments in distribution, VD decays multiexponentially. In this case, two types of VD can be used: VD_{area} and steady-state volume of distribution (VD_{ss}).

$$VD_{area} = \frac{CL}{\beta} = \frac{D}{\beta AUC}, VD_{ss} = VD_c \left(1 + \frac{k_{12}}{k_{21}}\right) \quad (25)$$

VD_{area} is the ratio between plasma clearance (CL) and the elimination rate constant β . VD_{ss} is the equilibrium distribution of a drug between the systemic circulation and tissue compartments. k_{12} and k_{21} are the constants of distribution between the systemic circulation and tissues.

VD_c and VD_{ss} are used as endpoints in ADME prediction models. In a VD_c prediction model, plasma protein binding information can be used, since VD_c is the volume of the systemic circulation. Taravat Ghafourian et al. established models for all data, acids, and bases (Ghafourian et al. 2006). VD_{ss} per unit weight (VD_{ss}/MW) (Zhivkova and Doytchinova 2012) is used as an endpoint for acidic drugs. In another work, the Øie-Tozer equation (Øie and Tozer 1979) was applied to calculate VD_{ss} with immobilized artificial membrane (IAM) partitioning coefficients ($\log k_{IAM}$), the predicted fraction unbound in tissues (f_{ut}), and the experimental fraction unbound in plasma (f_u) (Sui et al. 2009).

$$VD_{ss} = V_P (1 + R_{E/I}) + f_u V_P \left(\frac{V_E}{V_P} - R_{E/I}\right) + \frac{V_R f_u}{f_{ut}} \quad (26)$$

$R_{E/I}$ is the ratio of binding proteins in the extracellular fluid to those in plasma. V_P , V_E , and V_R are the volumes of plasma, extracellular fluid, and the remaining fluid, respectively. These are considered as constants in the human body. The $\log f_{ut}$ prediction model was built to predict VD_{ss} .

Blood Brain Barrier (BBB)

In vitro assays for BBB prediction are focused on membrane penetration using PAMPA, caco-2 cell, and MDCK cell lines. For the QSAR model, the log scale of the brain/plasma ratio obtained by in vivo assay is used as:

$$\log BB = AUC_{brain}/AUC_{plasma} \quad (27)$$

where AUC_{brain} and AUC_{plasma} are the areas under the curve for brain and plasma concentrations, respectively. $\log BB$ is also used to classify compounds as BBB+ (BBB permeable) and BBB- (not permeable), based on certain standards. A classification model has been developed to predict the BBB permeability of drug (Mensch et al. 2009). Molecular dynamics (MD) simulations are applied to study BBB penetration of protonated sumatriptan (pSMT) using NAMD2 (Wood and

Pickholz 2013). In the simulation, 1-palmitoil-2-oleoil-sn-glicero-3-phosphatidylcholine (POPC) is used as the BBB. Interactions between POPC and pSMT are investigated under physiological conditions.

Metabolism

Metabolism is the biotransformation or chemical modification of exogenous compounds to increase their hydrophilicity and water solubility to facilitate their excretion. A compound that increases in hydrophilicity is removed from the circulatory system through the kidney. Metabolism can be classified into two main phases, phase I and phase II. Phase I metabolism is governed mainly by cytochrome P450 enzymes (CYP) and includes oxidation, reduction, and hydrolysis reactions. The CYP enzymes involved in phase I drug metabolism are summarized in Table 3 (Dowty et al. 2011). Phase II metabolism consists mainly of conjugation reactions. Moieties increasing hydrophilicity are added to exogenous compounds or its metabolites. Metabolites obtained as a result of phase-I and/or phase-II reactions in the liver may go through further biotransformation reactions by interacting with other enzymes. Tissue distribution of phase II enzymes is summarized in Table 4, and their substrates, inhibitors, and inducers have been well summarized elsewhere (Jancova et al. 2010). The phase I reaction usually generates the conjugation reaction sites for phase II; however, phase I does not necessarily precede phase II.

Drugs and their metabolites from phase I and/or phase II metabolism interact with transporters and/or nuclear receptors. Transporters are expressed in various organs and pump compounds into or out of the cell across the cell membrane, a process also called active transport. The major organs of transporter expression are summarized in Table 5 (Matsson and Bergström 2015). In the liver, transporters pump drugs into hepatocytes and pump out drugs and metabolites into the systemic circulation or bile duct. Drugs and metabolites are pumped out through transporters into the systemic circulation, and then circulate through blood vessels, are distributed among organs, and are finally excreted through the kidney. Drugs and metabolites transferred to the bile duct are excreted through the gastrointestinal (GI) tract.

Nuclear receptors in the liver are activated by the binding of drugs or metabolites and then transferred into the nucleus where they increase gene expression of phase I and phase II enzymes and transporters. This pathway leads to the increase of the metabolic reaction rate due to the increase of the concentration of enzymes.

In vitro methods of drug metabolism are performed mainly with hepatic microsomes that contain hepatic CYP enzymes in order to test the metabolic stability of a drug to the CYP enzymes. Liquid chromatography/mass spectrometry (LC/MS) or liquid chromatography/tandem mass spectrometry (LC/MS/MS) allows identification of the site of metabolism (SoM). NMR is used to elucidate the structure of metabolites. However, these techniques cannot be performed in high-throughput screening (HTS); therefore, it is necessary to develop *in silico* methods for rapid screening of compounds with moderate metabolic stability.

Table 3 Tissue distribution of CYP enzymes (Anzenbacher and Anzenbacherova 2001) and their substrates (Lewis and Ito 2009)

CYP	Organ	Substrates
1A1	Lung, liver, brain, GI, heart, lymphocytes	Polycyclic aromatic hydrocarbons
1A2	Liver	Small, planar lipophilic compounds
1B1	Skin, brain, heart, lung, liver, kidney, GI, spleen, placenta	Polycyclic aromatic hydrocarbons
2A6	Liver	Relatively small neutral molecules
2B1/2	Brain	Morphine (typical substrate)
2B6	Liver, heart	Unionized basic compounds
2C8	Liver, kidney	Medium to large molecular weight acidic and lipophilic compounds
2C9	Liver	Hydrophobic and weakly acidic compounds
2C10	Liver	Tolbutamide, diclofenac (typical substrate)
2C19	Liver, heart	Medium to large molecular weight lipophilic amines and amides
2D6	Liver, brain, heart	Ionized basic compounds
2E1	Liver, lung, brain, endothelium, heart, bone marrow	Small neutral molecules
2 F	Lung	Coumarins (typical substrate)
3A4/5	Liver, GI, kidney, lung, brain, endothelium, placenta, lymphocytes	Broad substrate specificity
3A7	Fetus, placenta, liver	Broad substrate specificity
4A9/11	Kidney	Fatty acids
4B1	Lung, placenta	?
4 F2/3	Kidney	Arachidonic acid derivatives

Phase I Metabolism

Phase-I metabolic reactions are well studied because CYPs play a major role in drug metabolism. *In silico* models have been developed to predict enzyme selectivity, to predict the SoM, to predict metabolite structure, and to illuminate the interaction of a compound with CYP enzymes (Andrade et al. 2014).

Classification models give yes or no answers to the question of whether a compound is a substrate or inhibitor of a certain enzyme. The models mainly use machine learning algorithms like SVM (Yap and Chen 2005) or decision tree (Choi et al. 2009).

Various methods have been introduced to develop models for SoM prediction. The most widely used methods are: (i) calculation of the activation energy (representative of reactivity) of phase-I metabolism reactions with the molecular orbital calculation or empirical rule and (ii) knowledge-based training of SoM with topological and electronic descriptors. Reactivity can be scored using a model developed by semiempirical MO-derived descriptors (Hennemann et al. 2009). Activation energies (E_a) calculated at SoMs are good estimators of the activity

Table 4 Tissue distribution of phase II enzymes in isoform level

Enzymes	Isoforms	Organ
UGTs	1A1, 1A3, 1A4, 1A6, 1A9, 2B7, 2B15	Liver
	1A7, 1A8, 1A10	GI
SULTs	1A subfamily (1A1: highest expression)	Liver, brain, breast, intestine, jejunum, lung, adrenal gland, endometrium, placenta, kidney, blood platelets.
	1B1	Liver, small intestine, colon, leukocytes
	1C subfamily	Fetal kidney, lung, heart, GI
	1E1	Liver, jejunum, fetal liver, lung, kidney
	2A1	Liver, adrenal, duodenum, fetal adrenal gland
	2B	Prostate, placenta, adrenal gland, ovary, lung, kidney, colon
	4A1	Brain
	6B1	Testis, kidney
NATs	NAT1	Ubiquitous tissue distribution
	NAT2	Liver, colon, intestinal epithelium
GSTs	Alpha, Kappa, Mu, Pi, Sigma, Theta, Zeta, Omega	Soluble GSTs : cytoplasm, nucleus, mitochondria, peroxisomes
TPMP		Liver, kidney (highest) / brain, lung (low)
COMTs	Cytoplasmic soluble form	Peripheral tissues
	ER membrane-bound form	Brain

UGTs uridine diphosphate (UDP)-glucuronosyltransferases, *SULTs* sulfotransferases, *NATs* N-acetyltransferases, *GSTs* glutathione S-transferases, *TPMT* thiopurine S-methyltransferase, *COMT* catechol O-methyl transferase

of substrate. The activation energy of the reaction between heme in CYP and a substrate bound to the reaction pocket/pockets of CYP can be calculated with the transition state (TS) MO calculation. EaMEAD is an empirical model developed to estimate E_a of the reactions using ab initio MO transition states as constraints for training the model (Kim et al. 2009). SmartCyp uses density functional theory (DFT) to calculate transition state and to build a E_a prediction model with 2D descriptors (Rydberg et al. 2010). Prediction models for CYP1A2 (Jung et al. 2008) and CYP3A4 (Oh et al. 2008) were developed by using both CYP-substrate docking and E_a s at SoMs. In MetaSite, a grid of pockets is considered (Cruciani et al. 2006). DR-predictor (Huang et al. 2013) and IDSite (Li et al. 2011) introduced docking for activity prediction. FAME is a reactivity-based SoM prediction model that is not limited to specific enzymes and is applicable to both phase I and phase II metabolism (Kirchmair et al. 2013).

For training QSAR, database of both parent and metabolite structures (containing SoM information) must be constructed. MetaPrint2D uses fingerprints to predict SoM. Simulations Plus developed the artificial neural network ensembles (ANNe) model. PASS uses labeled multilevel neighborhoods of an atom (LMNA) in the

Table 5 The distribution of transporters among organs that contribute mainly to ADME and their expression level. The expression level is quantified (3: high, 2: moderate, 1: low, 0: no data or not detected)

Transporter	Small intestine	Liver	Kidney	Brain
MDR1	2	2.5	1	3
BSEP	0	3	0	1
MRP1	2	0	3	1
MRP2	1.5	2.5	2	2
MRP3	2	0	2	1
MRP4	0	0	0	1.5
MRP5	1	0	3	1
BCRP	2	1.5	1	2
PEPT1	3	2	2	2
OCT1	2	2	2	1
OCT2	0	0	3	1
OCT3	2	2	3	2
OAT3	0	0	2	1
MATE1	2	1	3	1
MATE2	1	0	1	2
OATP1B1	0	2	0	0
OATP1B3	0	3	0	0
OATP2B1	0	1	0	2

compound descriptors to predict probability of SoM (Rudik et al. 2014). Xenosite applied a deep learning algorithm to predict SOM (Hughes et al. 2015). Expert systems predict metabolic sites and products of metabolism using dictionaries of biotransformation according to rules in the system. MetabolExpert, Metabolizer, and Meteor are examples of expert systems. Metaprint2D-React is a QSAR model with fingerprint descriptors to predict metabolite structures.

The interaction of compounds with enzymes was studied using crystal structures of CYP enzymes. Recently, Kirchmair (Kirchmair et al. 2015) reviewed metabolism prediction models and recently available CYP crystal structures. Increased information of CYP crystal structures enables us to use various computer simulation methods, such as protein-ligand docking, molecular dynamics (MD), and quantum mechanics/molecular mechanics (QM/MM) (Lonsdale et al. 2013).

Phase II Metabolism

The number of phase II prediction models is significantly fewer than that of phase I prediction models since CYP enzymes play a major role in drug metabolism. Despite the relatively minor role of phase II metabolism, it was found that some toxicants are produced by phase II metabolism.

Uridine diphosphate (UDP)-glucuronosyltransferases (UGTs) are key enzymes in phase II metabolism, and ligand-based approaches like 2D-, 3D-QSAR, pharmacophore models, and classification models are applied to predict UDP-mediated

metabolism (Miners et al. 2004). Recently, a SoM by UGT prediction model was published. In this model, data are grouped into four classes according to reactions: aliphatic hydroxyl, aromatic hydroxyl, carboxylic acid, or amino nitrogen. The SVM model has been trained on each data set (Peng et al. 2013).

The experimentally identified UGT crystal structure is available; however, the binding domain is not resolved in the crystal structure (Miley et al. 2007). A number of human sulfotransferase (SULTs) crystal structures are known (1A1, 2, 3, 4/1B1/1C1, 2, 3/1E1/2A1/2B1a, b/4A1/6B1) (Moroy et al. 2012). However, studies on SULTs crystal structures like docking, MD (Stjerschantz et al. 2010), and virtual screenings (Campagna-Slater and Schapira 2009) are few due to difficulties in handling flexible binding pockets of SULTs.

Transporter

Hepatic transporters have a major role in determining drug concentration in hepatocytes. However, few *in silico* models are available to predict whether a compound's role is a substrate, inhibitor, or has no interaction. As listed in Table 3, though several transporters are working as hepatic transporters, no human transporter structures are known. An even more difficult point is that the data reported from different laboratories on the classification between substrate and inhibitor does not match. Therefore, it is difficult to develop a QSAR model with high reliability. P-glycoprotein (P-gP), also known as multidrug resistance protein 1 (MDR1) and ABCB1, is one of the most important transporters in drug resistance; however, its x-ray structure is unknown and the drug pumping mechanism is unclear. Human P-gP structure was obtained by homology modeling using mouse P-gP crystal structure, and this structure was used for the study of the docking of the P-gP substrate with P-gP (Pajeva et al. 2009). P-gP has several binding sites, and defining binding sites from a large pocket is a formidable task. Homology model-based study provides theoretical suggestions on probable binding poses and sites (Klepsch et al. 2011). Recently, with the binding pocket structure obtained from the homology modeled P-gP structure, four probable binding sites were proposed (Kim et al. 2014b).

The classification model which determines whether a compound is binder or nonbinder of OATP, OAT, and OCT was developed; however, the predictability is not satisfactory because of the high ambiguity of the training data set (You et al. 2015).

Nuclear Receptor

Pregnane X receptor (PXR) and constitutive androstane receptor (CAR) are the major nuclear receptors (NRs) in hepatocytes. These NRs are activated by drugs or metabolites and work with other proteins to regulate gene expression, especially the genes of the proteins that are related to drug metabolism. Such protein expression variation is related to drug-drug interactions (DDI).

Very few *in silico* methods have been applied for PXR and CAR. PXR ligand classification model was developed with known PXR binders as constraints (Ma et al. 2012). Crystal structures of PXR have been used for pharmacophore modeling (Chen et al. 2011) and docking (Ai et al. 2015). Two CAR crystal structures of

ligand-binding domains have been identified, and a homology modeled ligand-binding domain of CAR based on PXR and the vitamin D receptor (VDR) was used for MD simulations (Windshugel et al. 2005).

Half-Life of Drug

Drug half-life is the time required to reach 50 % of plasma concentration after reaching pseudoequilibrium, not half of the administered dose. Therefore, the half-life is influenced by each ADME endpoint of a drug. The half-life is an important guideline for dosing regimens. The half-life prediction model was built using experimental physicochemical properties of drugs. A local QSAR model was developed where data were divided into classes according to pK_a values, and a multiple linear regression model was trained with the data from each class of drugs (Durairaj et al. 2009). Self-organizing molecular field analysis (SOMFA) was used to get contribution of 3D conformations and electrostatic potential unto half-life (Goel et al. 2011).

Excretion

Excretion is the removal of chemicals from the body, and clearance (Cl) is a parameter to measure the elimination of drug. QSAR models have been developed for clearance prediction. Endpoints of these models are intrinsic (Cl_{int}), hepatic (Cl_H), renal (Cl_R), biliary (Cl_b), and total clearance in animals or humans (Cl_{animal} or Cl_{human}) (Yu 2010). Cl_{int} is measured using in vitro microsomal assay, and Cl_H is calculated using Cl_{int} as (Varma et al. 2015):

$$Cl_H = Q_H \cdot \frac{f_{ub} \cdot (PS_{influx} + PS_{diff}) \cdot Cl_{int}}{Q_H \cdot (PS_{efflux} + PS_{diff} + Cl_{int}) + f_{ub} \cdot (PS_{influx} + PS_{diff}) \cdot Cl_{int}} \quad (28)$$

where Q_H is the blood flow in liver, f_{ub} is the unbound fraction of drug, PS_{influx} , PS_{efflux} , and PS_{diff} are transportation of drug by efflux, influx, and passive diffusion, respectively. The above equation can be reduced to predict clearance by CYPs and other enzymes without considering transporters as follows:

$$Cl_H = Q_H \cdot [f_{ub} \cdot Cl_{int} / (Q_H + f_{ub} \cdot Cl_{int})] \quad (29)$$

Cl_R is dependent on the activity of glomerular filtration (GFR) and of transporters. It is defined as (Varma et al. 2015):

$$Cl_R = (f_{ub} \cdot GFR + Cl_{sec}) \cdot (1 - F_{reabs}) \quad (30)$$

Cl_{sec} is renal clearance by secretion and F_{reabs} is the fraction of reabsorbed drugs. Influx and efflux transporters govern these processes. To consider the mechanism of excretion in the kidney, the Cl_R prediction model was developed by considering

Table 6 Available software for PBPK models

Distributor	Program
Bayer Technology Services	PK-Sim
Cyprotex	chemPK
Certara	Simcyp
Simulations Plus	PBPKPlus™

the elimination mechanism (net secretion and net reabsorption) and information from 5 transporter subfamilies (P-glycoprotein (P-gP), multidrug resistance associated proteins (MRP) 1/2/4, organic anion transporters (OAT) 1/3, organic cation transporters (OCT) 2, and multidrug and toxin extrusion (MATE) 1/2 K) (Dave and Morris 2015).

Cl_b is calculated as,

$$Cl_b = X_{\text{bile}}/AUC_{\text{plasma}} \quad (31)$$

X_{bile} is the amount of drug excreted into bile, and AUC_{plasma} is the area under the curve of plasma concentration profile. A prediction model that was trained with two types of hepatocyte transporters (MRP2 and P-gP) was published to make prediction more reasonable (Yang et al. 2009).

Total clearance is the sum of all aforementioned CIs. CI can be described as:

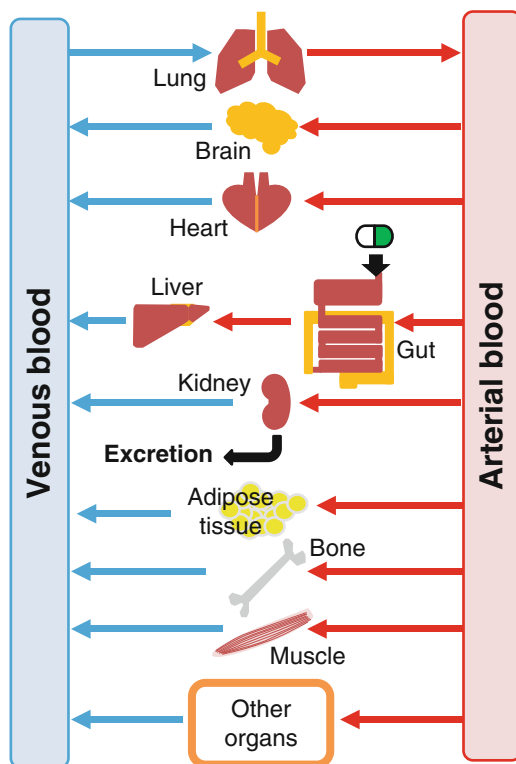
$$Cl = Cl_{\text{liver}} + Cl_{\text{kidney}} + Cl_{\text{other}} \quad (32)$$

QSAR models were developed to directly predict Cl_{human} with descriptors. Interspecies correlation in clearance is also an important research topic to predict Cl_{human} from preclinical results. QSAR models using Cl_{animal} and molecular descriptors together to predict Cl_{human} have also been developed (Huang et al. 2015).

PART III: Physiologically Based Pharmacokinetic (PBPK) Models

PBPK models describe an organism as a closed circulatory system composed of organs that contribute to the PK of drugs (Edginton et al. 2008). Figure 3 shows the architecture of the PBPK model. In this model, organs are described with differential equations to calculate the changes of drug concentration in each organ as a function of time. PBPK model architecture is varied according to the modeling purpose and hypothesis. To develop the model, one needs information on: (i) physical properties of the drug molecule and formulation properties; (ii) properties of the biological system interacting with the drug molecule (physiological property); and (iii) the anatomical structure of tissues and organs (Rowland et al. 2011). Here, we introduce the organ-based simulation models, which are necessary for PK prediction of oral administration. Available software for PBPK models is listed in Table 6. Details for implementation of the PBPK model were explained by Peters (Peters 2012).

Fig. 3 In the physiologically based pharmacokinetic (PBPK) model, organs important for ADME are included. The organs in PBPK can be defined by a differential equation. For deterministic organs for the blood concentration profile, such as the gut, liver, and kidney, details of the mechanism are considered to implement the model. Additional tissues and organs can be added according to modeling purpose and hypothesis



Gut Physiological Model

The gut physiological model is broadly composed of three types of prediction models: dissolution, permeation, and absorption. The dissolution rate (DR) is the starting point of gut simulation models of orally administered drugs that are formulated as a tablet or capsule (Fig. 4). Various models have been developed for predicting the dissolution profile. Each model considers various formulation conditions to predict varied dissolution profile (Costa and Lobo 2001). Noyes-Whitney coined the DR equation first, and models are modified by defining the dissolution constant (k) differently (Dokoumetzidis and Macheras 2006).

$$dC/dt = k(C_S - C) \quad (33)$$

C_S is saturation solubility and C is the concentration at time t . The concentration gradient is the driving force in dissolution, and it is described by the difference between C_S and C .

DR varies based on the condition of the GI fluid. GI fluid properties are influenced by factors such as agitation strength, pH, bile concentration, and food. These factors influence the diffusion coefficient and pH-dependent solubility of

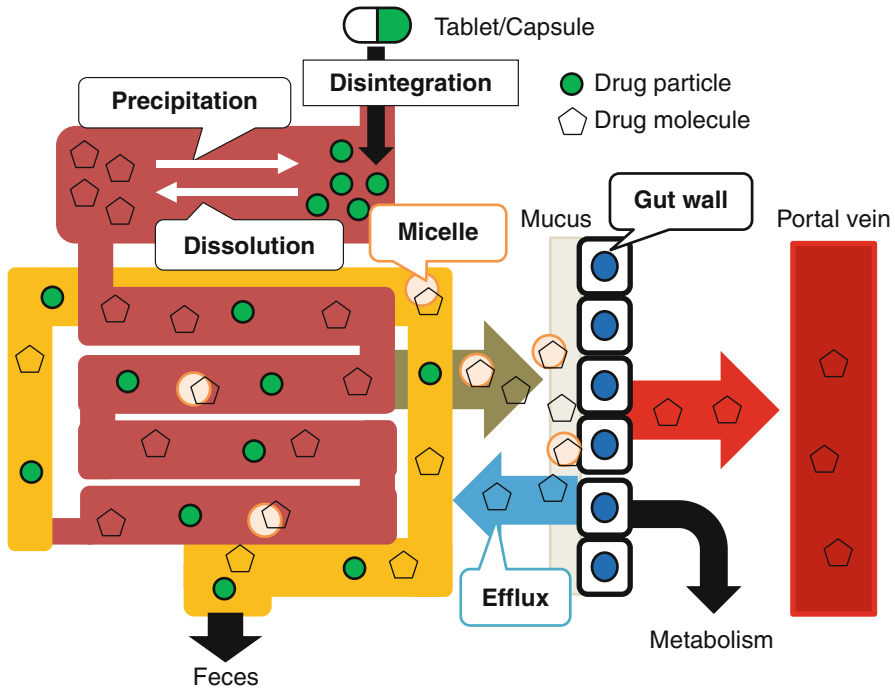


Fig. 4 Tablets or capsules that are orally administered are disintegrated into particles. Depending on the pH condition, drug molecules are either dissolved or precipitated. Drug particles are excreted through the gastrointestinal (GI) tract. Free drugs interact with micelles. Micelle-free drugs can be transported through the gut wall. In the gut wall, drug transporters and drug metabolizing enzymes (DMEs) are expressed. A portion of absorbed drugs are pumped out by transporters or metabolized by the DMEs. Drug molecules that safely cross the gut wall reach the portal vein

the formulated drug. Basic compounds and salts are highly soluble at a lower pH, but they are precipitated again at higher-pH conditions, such as in the intestines. Drug particles that are not fully dissolved act as the seed of nucleation of drug molecules and the amount of precipitated drug molecule is increased. Absorption varies between fasting and fed states. Certain foods inhibit transporters and DMEs, for example, grapefruit juice is an inhibitor of P-gP and CYP3A4, and black pepper is an inhibitor of CYPs 1A, 2B, and 2E1.

Physiological properties are related to the gut wall. Drugs need to cross various barriers to reach the portal vein. The mucus layer is composed of mucus secreted by the mucus membrane and forms a slippery cover on the gut wall. Drug permeation of the gut wall has the following three types of pathways: passive transcellular, active transcellular, and paracellular pathways. The passive transcellular pathway goes through the cell; the active transcellular pathway is permeation using a transporter; and the paracellular pathway is the passing through the tight junction between the cells. During the process of active transportation, drugs interact with DMEs and

transporters. CYPs, UGTs, Esterases, and SULTs are major DMEs in the human gut wall. Regarding transporters, P-gP, MRP2, and BCRP are efflux transporters, and OATP, OCT1, MCT1, ASBT, and PEPT1 are uptake transporters.

In gut physiological models, the anatomy of the gut is described in two ways, the compartmental model and the dispersion model. In the compartmental model, the intestine is a series of compartments that have different permeation properties in each compartment, though in reality, the small intestine cannot be clearly divided according to physiological properties. The dispersion model considers the intestine as a continuous tube. In this model, the velocity of intestinal fluid is taken into consideration. Mass balance is considered in the equation of GI tract transition. The process of flowing in and out and dissolution and precipitation is implemented in differential equations. The details of the physics applied in the absorption simulation are reviewed by Sugano (Sugano 2009).

Tissue Physiological Model

Drugs that reach the portal vein begin interaction with proteins in the blood and tissues in the circulation system. The ionization state and lipophilicity of the drug influences the interaction. Therefore, physical properties of drugs are the most important descriptors for defining the partitioning of the drug between blood and tissue. In describing distribution, protein binding is important because it represents the amount of drug not available for tissue distribution. Acidic drugs bind to albumin, lipophilic neutral drugs bind to lipoproteins, and basic drugs bind to AGP. The unbound fraction of drug interacts with membranes. Poulin and Thiel published models for the tissue/plasma partition coefficient based on the assumption that plasma and tissue is composed of water, lipids, and phospholipids (Poulin and Thiel 2001).

$$K_{T(\text{nonadipose}):P} = \frac{P_{OW} (f_{nl,T} + 0.3 f_{pl,T}) + (f_{w,T} + 0.7 f_{pl,T})}{P_{OW} (f_{nl,P} + 0.3 f_{pl,P}) + (f_{w,P} + 0.7 f_{pl,P})} \cdot \frac{f_{uP}}{f_{uT}} \quad (34)$$

$$K_{T(\text{adipose}):P} = \frac{D_{OW} (f_{nl,T} + 0.3 f_{pl,T}) + (f_{w,T} + 0.7 f_{pl,T})}{D_{OW} (f_{nl,P} + 0.3 f_{pl,P}) + (f_{w,P} + 0.7 f_{pl,P})} \cdot \frac{f_{uP}}{1} \quad (35)$$

Partition coefficients for partitioning between tissue and plasma are defined differently on nonadipose ($K_{T(\text{nonadipose}):P}$) and adipose tissues ($K_{T(\text{adipose}):P}$). P_{OW} is the octanol-water partition coefficient of a drug, and D_{OW} is that of an ionizable drug. f_{pl} , f_{nl} , and f_w are the fractional volume contents of phospholipids, neutral lipid, and water in plasma (P) or in tissue (T), respectively. f_{uP} and f_{uT} are the unbound fractions of drug in plasma and tissue. In the case of adipose tissue, f_{uT} is assumed to be one.

The drug concentration change in tissues that are not eliminating drugs is described as:

$$dC_T/dt = \frac{1}{V_T} \left(Q_{ART} \times C_{ART} - Q_{VEN} \frac{C_T \cdot f_{uT} \cdot R}{f_{uP} K_{T:P}} \right) \quad (36)$$

where V_T is the volume of the tissue, and Q_{ART} and Q_{VEN} are the blood flow rates of the artery and vein, respectively. C_T and C_{ART} are the drug concentrations in the tissue and the artery. $K_{T:P}$ is the partition coefficient for tissue and plasma. R is the blood-plasma ratio.

Liver Physiological Model

After drugs reach the portal vein, they are transported into the liver, which is composed of tissues rich in blood vessels called lobules. These physiological structures of the liver are important because metabolism occurs in a zone-dependent manner. Sinusoids are capillaries in the liver in which drug molecules are transferred through to hepatocytes and metabolized (Fig. 5).

Zone-dependent metabolism cannot be studied through in vitro assay because the architecture of the liver cannot be considered in the assay. For simulating zonation of metabolic processes in a lobule, various simulation models have been applied – multicompartment models, distribution-based models, agent-based models, interconnected parallel tubes, (Lerapetritou et al. 2009), and lattice models (Rezania et al. 2013a, b).

Elimination kinetics occurring in the liver does not always occur by one enzyme alone; sometimes various enzymes work together in biotransformation (Isin and Guengerich 2007). Therefore, various metabolites can be derived from a single drug. CYP enzyme reaction networks are modeled to calculate whole human liver clearance ($CL_{H.liver}$) by considering the clearance of each recombinantly expressed CYP enzyme with a number of metabolic pathways (Rostami-Hodjegan and Tucker 2007).

$$CL_{H.liver} = \left[\sum_{j=1}^n \left(\sum_{i=1}^n a_{ji} \times \frac{V_{\max i} (CYP_i) \times CYP_j}{K_{mi} (CYP_j)} \right) \right] \times MPGL \times \text{Liver weight} \quad (37)$$

Here, i th metabolic pathway for each j th CYP is considered together in the calculation. $V_{\max i}$ and K_{mi} are the maximum rate of metabolism and the Michaelis constant for an individual CYP, respectively. $MPGL$ is the amount of microsomal protein per gram of liver. a_{ji} is a scaling factor to compensate for differences in activity for each enzyme between recombinant systems and hepatic enzymes.

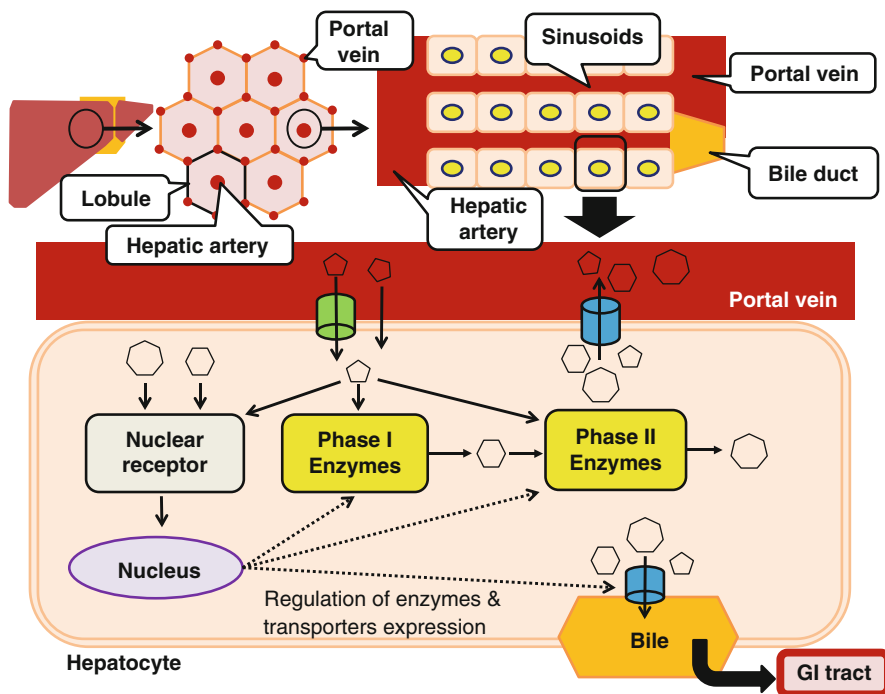


Fig. 5 The structural unit of liver tissue is the lobule. Lobule tissue is richly entangled with arteries and portal veins. Hepatic arteries in lobules are branched out through sinusoids, which are surrounded by hepatocytes. Drug molecules are transferred into hepatocytes while blood is flowing through sinusoids. Drug molecules enter into hepatocytes through passive diffusion or hepatic transporters. Phase I and phase II metabolites are transported by transporters into the portal vein and excreted through the kidney, or into bile, where the metabolites are subsequently excreted through the GI tract. Nuclear receptors (NRs) interact with drug molecules and their phase I and phase II metabolites. NRs regulate gene expression of DMEs and transporters in the nucleus

A single-cell-based simulation has been introduced. In this model, the biochemical network of hepatocytes is modeled; eight models of hepatocytes interact with sinusoids. This model can simulate the influence of the variation in gene expression levels to macroscopic character variations (Ohno et al. 2008).

Hepatosys is a project to develop a virtual liver network (General concept and deliverables of phase II “integration” 2007). This project is comprised of four networks (detoxification, endocytosis, regeneration, and iron regulation of hepatocytes) and two platforms (experiment platform and modeling platform). A mathematical model for each network is examined by genome data from microarrays, proteome and metabolome data, and kinetics of protein interactions. The ultimate goal of the Hepatosys project is to understand drug metabolism in individual patients.

Kidney Physiological Model

Parent drugs and metabolites in the systemic circulation are removed through the kidney. The physiological kidney has been modeled in the Physiome Project (Thomas et al. 2006). This model describes glomerular filtration, tubular transport, tubuloglomerular feedback, renal microcirculation, the renal medulla, and the urine-concentrating mechanism with equations.

In the multicompartiment model, the kidney is assumed to be composed of vascular, interstitial, tubule lumen, and intracellular compartments. Glomerular filtration and urine flow is a parameter in the differential equation for drug concentration. Passive reabsorption and passive permeability are considered between the tubule lumen and interstitial compartments. Clearance by DMEs and transporters is considered in the intracellular compartment (Fig. 6). The Guyton model is an attempt to simulate the physiology of humans. In this model, the physiological model of kidney is implemented. The C++ library of the Guyton model was released (Hernandez et al. 2009; Thomas et al. 2008) and has been actively applied in high blood pressure simulations.

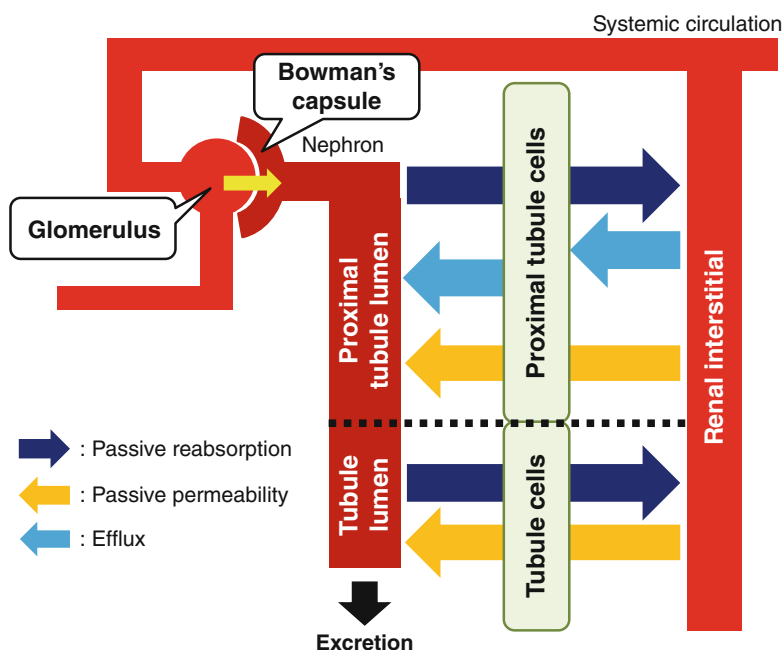


Fig. 6 In the kidney, drugs and their metabolites are eliminated from the body. Drugs and metabolites are filtered at the glomerulus and transported by reabsorption and passive permeation. Reabsorbed drugs and metabolites can be moved into the proximal tubule lumen again by transporters expressed in proximal tubule cells. Drugs and metabolites remaining in the tubule lumen are excreted

Summary

In silico models for ADME predictions were divided into three parts and reviewed. The physical properties of chemicals are important for describing ADME, ADME endpoints, and PBPK models. The number of *in silico* models is growing rapidly. Each model is improved with increased data for model construction and by introducing advanced modeling techniques. These improvements have led to an increase in utilization of *in silico* models in drug discovery and development. Despite the remarkable developments of ADME models, there are several limitations in reproducing experimental results with *in silico* models. These shortcomings are mainly due to: (i) not enough data for the endpoint to be modeled, (ii) lack of full understanding of the biological phenomena of the endpoint, and (iii) the biological complexity of the endpoint is too complicated to be modeled. More research on the molecular level of the interaction between pharmaceuticals and biomolecules is required to improve *in silico* models. Moreover, an understanding on the molecular level of the interactions in ADME could make it possible to simulate individual ADME properties. ADME properties of drugs in individuals are varied since each person manifests with different expression patterns of biomolecules interacting with drugs. This uniqueness in individuals has brought about the concept of personalized medicine. In the future, a deeper understanding upon the mechanisms of ADME in the molecular level may lead to personalized ADME predictions.

Acknowledgments This work is supported by the Industrial Core Technology Development Program (10054749, software development about drug metabolism prediction) and funded by the Ministry of Trade, Industry and Energy (MOTIE).

Bibliography

- Abraham, M. H., Grellier, P. L., Hamerton, I., McGill, R. A., Prior, D. V., & Whiting, G. S. (1988). Solvation of gaseous non-electrolytes. *Faraday Discussions of the Chemical Society*, 85, 107–115.
- Ai, N., Fan, X., & Ekins, S. (2015). In silico methods for predicting drug-drug interactions with cytochrome P-450s, transporters and beyond. *Advanced Drug Delivery Reviews*, 86, 46–60. doi:10.1016/j.addr.2015.03.006.
- Akamatsu, M., Fujiwaka, M., Nakao, K., & Shimizu, R. (2009). In silico prediction of human oral absorption based on QSAR analyses of PAMPA permeability. *Chemistry & Biodiversity*, 6(11), 1845–1866.
- Alkorta, I., & Popelier, P. L. (2015). Linear free-energy relationships between a single gas-phase ab initio equilibrium bond length and experimental pKa values in aqueous solution. *Chemphyschem*, 16(2), 465–469. doi:10.1002/cphc.201402711.
- Andrade, C. H., Silva, D. C., & Braga, R. C. (2014). In silico prediction of drug metabolism by P450. *Current Drug Metabolism*, 15(5), 514–525.
- Bergstrom, C. A., Luthman, K., & Artursson, P. (2004). Accuracy of calculated pH-dependent aqueous drug solubility. *European Journal of Pharmaceutical Sciences*, 22(5), 387–398. doi:10.1016/j.ejps.2004.04.006.
- Campagna-Slater, V., & Schapira, M. (2009). Evaluation of virtual screening as a tool for chemical genetic applications. *Journal of Chemical Information and Modeling*, 49(9), 2082–2091.

- Chen, C. N., Shih, Y. H., Ding, Y. L., & Leong, M. K. (2011). Predicting activation of the promiscuous human pregnane X receptor by pharmacophore ensemble/support vector machine approach. *Chemical Research in Toxicology*, *24*(10), 1765–1778. doi:10.1021/tx200310j.
- Chen, L. L., Yao, J., Yang, J. B., & Yang, J. (2005). Predicting MDCK cell permeation coefficients of organic molecules using membrane-interaction QSAR analysis. *Acta Pharmacologica Sinica*, *26*(11), 1322–1333. doi:10.1111/j.1745-7254.2005.00166.x.
- Cheng, F., Li, W., Liu, G., & Tang, Y. (2013). In silico ADMET prediction: Recent advances, current challenges and future trends. *Current Topics in Medicinal Chemistry*, *13*(11), 1273–1289.
- Choi, I., Kim, S. Y., Kim, H., Kang, N. S., Bae, M. A., Yoo, S. E., Jung, J., & No, K. T. (2009). Classification models for CYP450 3A4 inhibitors and non-inhibitors. *European Journal of Medicinal Chemistry*, *44*(6), 2354–2360. doi:10.1016/j.ejmech.2008.08.013.
- Colmenarejo, G. (2007). In silico prediction of plasma and tissue protein binding. In J. B. Taylor & D. J. Triggle (Eds.), *Comprehensive medicinal chemistry II* (ADME-Tox approaches, pp. 847–866). Amsterdam: Elsevier.
- Costa, P., & Lobo, J. M. S. (2001). Modeling and comparison of dissolution profiles. *European Journal of Pharmaceutical Sciences*, *13*(2), 123–133.
- Cramer, C. J., & Truhlar, D. G. (1995). Continuum solvation models. In O. Tapia & J. Bertrán (Eds.), *Solvent effects and chemical reactivity*. New York: Kluwer.
- Cruciani, G., Carosati, E., Boeck, B. D., Ethirajulu, K., Mackie, C., Howe, T., & Vianello, R. (2006). MetaSite: Understanding metabolism in human cytochromes from the perspective of the chemist. *Journal of Medicinal Chemistry*, *48*(22), 6970–6979.
- Dave, R. A., & Morris, M. E. (2015). Quantitative structure-pharmacokinetic relationships for the prediction of renal clearance in humans. *Drug Metabolism and Disposition*, *43*(1), 73–81. doi:10.1124/dmd.114.059857.
- Dokoumetzidis, A., & Macheras, P. (2006). A century of dissolution research: From Noyes and Whitney to the biopharmaceutics classification system. *International Journal of Pharmaceutics*, *321*(1–2), 1–11. doi:10.1016/j.ijpharm.2006.07.011.
- Dowty, M. E., Messing, D. M., Lai, Y., & Kirkovsky, L. (2011). ADME. In K. Tsaion & S. A. Kates (Eds.), *ADMET for medicinal chemists: A practical guide* (pp. 145–200). Singapore: Wiley.
- Durairaj, C., Shah, J. C., Senapati, S., & Kompella, U. B. (2009). Prediction of vitreal half-life based on drug physicochemical properties: Quantitative structure-pharmacokinetic relationships (QSPKR). *Pharmaceutical Research*, *26*(5), 1236–1260. doi:10.1007/s11095-008-9728-7.
- Edginton, A. N., Theil, F.-P., Schmitt, W., & Willmann, S. (2008). Whole body physiologically-based pharmacokinetic models: Their use in clinical drug development. *Expert Opinion on Drug Metabolism & Toxicology*, *4*(9), 1143–1152.
- Fraczkiewicz, R. (2013). In Silico prediction of ionization. In: Testa B, Waterbeemd Hvd (eds.) *Comprehensive Medicinal Chemistry II*, vol 5. Elsevier. doi:10.1016/b978-0-12-409547-2.02610-x
- Fraczkiewicz, R., Lobell, M., Goller, A. H., Krenz, U., Schoenneis, R., Clark, R. D., & Hillisch, A. (2015). Best of both worlds: Combining pharma data and state of the art modeling technology to improve in Silico pKa prediction. *Journal of Chemical Information and Modeling*, *55*(2), 389–397. doi:10.1021/ci500585w.
- Gasteiger, J., & Hutchings, M. G. (1983). New empirical models of substituent polarisability and their application to stabilisation effects in positively charged species. *Tetrahedron Letters*, *24*(25), 2537–2540.
- General concept and deliverables of Phase II “integration” (2007). Hepatosys project committee.
- Ghafourian, T., Barzegar-Jalali, M., Dastmalchi, S., Khavari-Khorasani, T., Hakimiha, N., & Nokhodchi, A. (2006). QSPR models for the prediction of apparent volume of distribution. *International Journal of Pharmaceutics*, *319*(1–2), 82–97. doi:10.1016/j.ijpharm.2006.03.043.
- Gleeson, M. P., Hersey, A., & Hannongbua, S. (2011). In-silico ADME models: A general assessment of their utility in drug discovery applications. *Current Topics in Medicinal Chemistry*, *11*(4), 358–381.

- Goel, H., Sinha, V. R., Thareja, S., Aggarwal, S., & Kumar, M. (2011). Assessment of biological half life using in silico QSPkR approach: A self organizing molecular field analysis (SOMFA) on a series of antimicrobial quinolone drugs. *International Journal of Pharmaceutics*, 415(1–2), 158–163. doi:10.1016/j.ijpharm.2011.05.065.
- Guidance document on the validation of (Quantitative)Structure-Activity Relationships[(Q)SAR] Models. (2007). vol 69. OECD, OECD Environment Health and Safety publications.
- Hansen, N. T., Kouskoumvekaki, I., Jorgensen, F. S., Brunak, S., & Jonsdottir, S. O. (2006). Prediction of pH-dependent aqueous solubility of druglike molecules. *Journal of Chemical Information and Modeling*, 46(6), 2601–2609.
- Hennemann, M., Friedl, A., Lobell, M., Keldenich, J., Hillisch, A., Clark, T., & Goller, A. H. (2009). CypScore: Quantitative prediction of reactivity toward cytochromes P450 based on semiempirical molecular orbital theory. *ChemMedChem*, 4(4), 657–669. doi:10.1002/cmdc.200800384.
- Hernandez, A. I., Le Rolle, V., Defontaine, A., & Carrault, G. (2009). A multiformalism and multiresolution modelling environment: Application to the cardiovascular system and its regulation. *Philosophical Transactions: A Mathematical, Physical and Engineering Sciences*, 367(1908), 4923–4940. doi:10.1098/rsta.2009.0163.
- Hou, T., Li, Y., Zhang, W., & Wang, J. (2009). Recent developments of in silico predictions of intestinal absorption and oral bioavailability. *Combinatorial Chemistry & High Throughput Screening*, 12(5), 497–506.
- Hou, T., Wang, J., Zhang, W., Wang, W., & Xu, X. (2006). Recent advances in computational prediction of drug absorption and permeability in drug discovery. *Current Medicinal Chemistry*, 13(22), 2653–2667.
- Huang, T. W., Zaretski, J., Bergeron, C., Bennett, K. P., & Breneman, C. M. (2013). DR-predictor: Incorporating flexible docking with specialized electronic reactivity and machine learning techniques to predict CYP-mediated sites of metabolism. *Journal of Chemical Information and Modeling*, 53(12), 3352–3366. doi:10.1021/ci4004688.
- Huang, W., Geng, L., Deng, R., Lu, S., Ma, G., Yu, J., Zhang, J., Liu, W., Hou, T., & Lu, X. (2015). Prediction of Human Clearance Based on Animal Data and Molecular Properties. *Chemical Biology and Drug Design*, 86(5), 990–997. doi:10.1111/cbdd.12567.
- Hughes, T. B., Miller, G. P., & Swamidass, S. J. (2015). Modeling epoxidation of drug-like molecules with a deep machine learning network. *ACS Central Science*, 1(4), 168–180. doi:10.1021/acscentsci.5b00131.
- Insin, E. M., & Guengerich, F. P. (2007). Complex reactions catalyzed by cytochrome P450 enzymes. *Biochimica et Biophysica Acta*, 1770(3), 314–329. doi:10.1016/j.bbagen.2006.07.003.
- Jancova, P., Anzenbacher, P., & Anzenbacherova, E. (2010). Phase II drug metabolizing enzymes. Biomedical papers of the medical faculty of the university palacky, olomouc. *Czechoslovakia*, 154(2), 103–116.
- Jung, J., Kim, N. D., Kim, S. Y., Choi, I., Cho, K.-H., Oh, W. S., Kim, D. N., & No, K. T. (2008). Regioselectivity prediction of CYP1A2-mediated phase I metabolism. *Journal of Chemical Information and Modeling*, 48(5), 1074–1080.
- Kim, D. N., Cho, K.-H., Oh, W. S., Lee, C. J., Lee, S. K., Jung, J., & No, K. T. (2009). EaMEAD: Activation energy prediction of CYP450 mediated metabolism with effective atomic descriptors. *Journal of Chemical Information and Modeling*, 49(7), 1643–1654.
- Kim, M. T., Sedykh, A., Chakravarti, S. K., Saiakhov, R. D., & Zhu, H. (2014a). Critical evaluation of human oral bioavailability for pharmaceutical drugs by using various cheminformatics approaches. *Pharmaceutical Research*, 31(4), 1002–1014. doi:10.1007/s11095-013-1222-1.
- Kim, N., Shin, J.-M., & No, K. T. (2014b). In silico study on the interaction between P-glycoprotein and its inhibitors at the drug binding pocket. *Bulletin of the Korean Chemical Society*, 35(8), 2317–2325. doi:10.5012/bkcs.2014.35.8.2317.
- Kirchmair, J., Goller, A. H., Lang, D., Kunze, J., Testa, B., Wilson, I. D., Glen, R. C., & Schneider, G. (2015). Predicting drug metabolism: Experiment and/or computation? *Nature Reviews Drug Discovery*, 14(6), 387–404. doi:10.1038/nrd4581.

- Kirchmair, J., Williamson, M. J., Afzal, A. M., Tyzack, J. D., Choy, A. P., Howlett, A., Rydberg, P., & Glen, R. C. (2013). FASt METabolizer (FAME): A rapid and accurate predictor of sites of metabolism in multiple species by endogenous enzymes. *Journal of Chemical Information and Modeling*, 53(11), 2896–2907. doi:10.1021/ci400503s.
- Klamt, A., Eckert, F., & Arlt, W. (2010). COSMO-RS: An alternative to simulation for calculating thermodynamic properties of liquid mixtures. *Annual Review of Chemical and Biomolecular Engineering*, 1, 101–122. doi:10.1146/annurev-chembioeng-073009-100903.
- Klepsch, F., Chiba, P., & Ecker, G. F. (2011). Exhaustive sampling of docking poses reveals binding hypotheses for propafenone type inhibitors of P-glycoprotein. *PLoS Computational Biology*, 7(5), e1002036. doi:10.1371/journal.pcbi.1002036.
- Lambrinidis, G., Vallianatou, T., & Tsantili-Kakoulidou, A. (2015). In vitro, in silico and integrated strategies for the estimation of plasma protein binding. A review. *Advanced Drug Delivery Reviews*, 86, 27–45. doi:10.1016/j.addr.2015.03.011.
- Lee, B., & Richards, F. M. (1971). The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology*, 55(3), 379–400.
- Lee, S., Cho, K.-H., Kang, Y.-M., Scheraga, H. A., & No, K. T. (2013). A generalized G-SFED continuum solvation free energy calculation model. *Proceedings of the National Academy of Sciences of the United States of America*, 110(8), E662–E667.
- Lerapetritou, M. G., Georgopoulos, P. G., Roth, C. M., & Androulakis, L. P. (2009). Tissue-level modeling of xenobiotic metabolism in liver: An emerging tool for enabling clinical translational research. *Clinical and Translational Science*, 2(3), 228–237. doi:10.1111/j.1752-8062.2009.00092.x.
- Lewis, D. F., & Ito, Y. (2009). Human P450s involved in drug metabolism and the use of structural modelling for understanding substrate selectivity and binding affinity. *Xenobiotica*, 39(8), 625–635. doi:10.1080/00498250903000255.
- Li, J., Schneebeli, S. T., Bylund, J., Farid, R., & Friesner, R. A. (2011). IDSite: An accurate approach to predict P450-mediated drug metabolism. *Journal of Chemical Theory and Computation*, 7(11), 3829–3845. doi:10.1021/ct200462q.
- Lonsdale, R., Houghton, K. T., Zurek, J., Bathelt, C. M., Foloppe, N., de Groot, M. J., Harvey, J. N., & Mulholland, A. J. (2013). Quantum mechanics/molecular mechanics modeling of regioselectivity of drug metabolism in cytochrome P450 2C9. *Journal of the American Chemical Society*, 135(21), 8001–8015. doi:10.1021/ja402016p.
- Lusci, A., Pollastri, G., & Baldi, P. (2013). Deep architectures and deep learning in chemoinformatics: The prediction of aqueous solubility for drug-like molecules. *Journal of Chemical Information and Modeling*, 53(7), 1563–1575. doi:10.1021/ci400187y.
- Ma, S. L., Joung, J. Y., Lee, S., Cho, K. H., & No, K. T. (2012). PXR ligand classification model with SFED-weighted WHIM and CoMMA descriptors. *SAR and QSAR in Environmental Research*, 23(5–6), 485–504. doi:10.1080/1062936X.2012.665385.
- Mannhold, R., Poda, G. I., Ostermann, C., & Tetko, I. V. (2009). Calculation of molecular lipophilicity: State-of-the-art and comparison of log P methods on more than 96,000 compounds. *Journal of Pharmaceutical Sciences*, 98(3), 861–893. doi:10.1002/jps.21494.
- Marenich, A. V., Cramer, C. J., & Truhlar, D. G. (2013). Generalized Born Solvation Model SM12. *Journal of Chemical Theory and Computation*, 9(1), 609–620. doi:10.1021/ct300900e.
- Matsson, P., & Bergström, CAS. (2015). Computational modeling to predict the functions and impact of drug transporters. *In Silico Pharmacology* 3(1). doi:10.1186/s40203-015-0012-3.
- Mennucci, B. (2012). Polarizable continuum model. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(3), 386–404. doi:10.1002/wcms.1086.
- Mensch, J., Oyarzabal, J., Mackie, C., & Augustijns, P. (2009). In vivo, in vitro and in silico methods for small molecule transfer across the BBB. *Journal of Pharmaceutical Sciences*, 98(12), 4429–4468. doi:10.1002/jps.21745.
- Miley, M. J., Zielinska, A. K., Keenan, J. E., Bratton, S. M., Radomska-Pandya, A., & Redinbo, M. R. (2007). Crystal structure of the cofactor-binding domain of the human phase II drug-metabolism enzyme UDP-glucuronosyltransferase 2B7. *Journal of Molecular Biology*, 369(2), 498–511. doi:10.1016/j.jmb.2007.03.066.

- Miners, J. O., Smith, P. A., Sorich, M. J., McKinnon, R. A., & Mackenzie, P. I. (2004). Predicting human drug glucuronidation parameters: Application of in vitro and in silico modeling approaches. *Annual Review of Pharmacology and Toxicology*, *44*, 1–25. doi:[10.1146/annurev.pharmtox.44.101802.121546](https://doi.org/10.1146/annurev.pharmtox.44.101802.121546).
- Moroy, G., Martiny, V. Y., Vayer, P., Villoutreix, B. O., & Miteva, M. A. (2012). Toward in silico structure-based ADMET prediction in drug discovery. *Drug Discovery Today*, *17*(1–2), 44–55. doi:[10.1016/j.drudis.2011.10.023](https://doi.org/10.1016/j.drudis.2011.10.023).
- Oh, W. S., Kim, D. N., Jung, J., Cho, K.-H., & No, K. T. (2008). New combined model for the prediction of regioselectivity in cytochrome P450-3A4 mediated metabolism. *Journal of Chemical Information and Modeling*, *48*(3), 591–601.
- Ohno, H., Naito, Y., Nakajima, H., & Tomita, M. (2008). Construction of a biological tissue model based on a single-cell model: A computer simulation of metabolic heterogeneity in the liver lobule. *Artificial Life*, *14*(1), 3–28.
- Øie, S., & Tozer, T. N. (1979). Effect of altered plasma protein binding on apparent volume of distribution. *Journal of Pharmaceutical Sciences*, *68*(9), 1203–1205.
- Anzenbacher, P., & Anzenbacherova, E. (2001). Cytochromes P450 and metabolism of xenobiotics. *Cellular and molecular life sciences : CMLS*, *58*(5–6), 737–747.
- Pajeva, I. K., Globisch, C., & Wiese, M. (2009). Comparison of the inward- and outward-open homology models and ligand binding of human P-glycoprotein. *FEBS Journal*, *276*(23), 7016–7026. doi:[10.1111/j.1742-4658.2009.07415.x](https://doi.org/10.1111/j.1742-4658.2009.07415.x).
- Peng, J., Lu, J., Shen, Q., Zheng, M., Luo, X., Zhu, W., Jiang, H., & Chen, K. (2013). In silico site of metabolism prediction for human UGT-catalyzed reactions. *Bioinformatics*, *30*(3), 398–405. doi:[10.1093/bioinformatics/btt681](https://doi.org/10.1093/bioinformatics/btt681).
- Peters, S. A. (2012). *Physiologically-Based Pharmacokinetic (PBPK) modeling and simulation principles, methods, and applications in the pharmaceutical industry*. Hoboken: Wiley.
- Pham The, H., González-Álvarez, I., Bermejo, M., Mangas Sanjuan, V., Centelles, I., Garrigues, T. M., & Cabrera-Pérez, M. Á. (2011). In Silico prediction of Caco-2 cell permeability by a classification QSAR approach. *Molecular Informatics*, *30*(4), 376–385. doi:[10.1002/minf.201000118](https://doi.org/10.1002/minf.201000118).
- Poulin, P., & Theil, F.-p. (2001). Prediction of PK prior to in vivo studies. 1. mechanism-based prediction of volume of distribution. *Journal of Pharmaceutical Sciences*, *91*(1), 129–156.
- Ran, Y., & Yalkowsky, S. H. (2001). Prediction of drug solubility by the general solubility equation (GSE). *Journal of Chemical Information and Modeling*, *41*(2), 354–357.
- Rezania V, Marsh R, Coombe D, & Tuszynski J (2013a). A physiologically-based flow network model for hepatic drug elimination I: Regular lattice lobule model. *Theoretical Biology and Medical Modelling* *10*(52). doi:[10.1186/1742-4682-10-52](https://doi.org/10.1186/1742-4682-10-52).
- Rezania, V., Marsh, R., Coombe, D., & Tuszynski, J (2013b). A physiologically-based flow network model for hepatic drug elimination II: Variable lattice lobule models. *Theoretical Biology and Medical Modelling* *10*(53). doi:[10.1186/1742-4682-10-53](https://doi.org/10.1186/1742-4682-10-53).
- Rostami-Hodjegan, A., & Tucker, G. T. (2007). Simulation and prediction of in vivo drug metabolism in human populations from in vitro data. *Nature Reviews Drug Discovery*, *6*(2), 140–148. doi:[10.1038/nrd2173](https://doi.org/10.1038/nrd2173).
- Rowland, M., Peck, C., & Tucker, G. (2011). Physiologically-based pharmacokinetics in drug development and regulatory science. *Annual Review of Pharmacology and Toxicology*, *51*, 45–73. doi:[10.1146/annurev-pharmtox-010510-100540](https://doi.org/10.1146/annurev-pharmtox-010510-100540).
- Rudik, A. V., Dmitriev, A. V., Lagunin, A. A., Filimonov, D. A., & Poroikov, V. V. (2014). Metabolism site prediction based on xenobiotic structural formulas and PASS prediction algorithm. *Journal of Chemical Information and Modeling*, *54*(2), 498–507. doi:[10.1021/ci400472j](https://doi.org/10.1021/ci400472j).
- Rydberg, P., Gloriam, D. E., Zaretski, J., Breneman, C., & Olsen, L. (2010). SMARTCyp: A 2D method for prediction of cytochrome P450-mediated drug metabolism. *ACS Medicinal Chemistry Letters*, *1*(3), 96–100. doi:[10.1021/ml100016x](https://doi.org/10.1021/ml100016x).

- Skyner, R. E., McDonagh, J. L., Groom, C. R., van Mourik, T., & Mitchell, J. B. (2015). A review of methods for the calculation of solution free energies and the modelling of systems in solution. *PCCP Physical Chemistry Chemical Physics*, 17(9), 6174–6191. doi:10.1039/c5cp00288e.
- Stjernschantz, E., Reinen, J., Meiln, W., George, B. J., Glatt, H., Vermeulen, N. P., & Oostenbrink, C. (2010). Comparison of murine and human estrogen sulfotransferase inhibition in vitro and in silico – implications for differences in activity, subunit dimerization and substrate inhibition. *Molecular and Cellular Endocrinology*, 317(1–2), 127–140. doi:10.1016/j.mce.2009.12.001.
- Sugano, K. (2009). Introduction to computational oral absorption simulation. *Expert Opinion on Drug Metabolism & Toxicology*, 5(3), 259–293.
- Sui, X., Sun, J., Li, H., Wang, Y., Liu, J., Liu, X., Zhang, W., Chen, L., & He, Z. (2009). Prediction of volume of distribution values in human using immobilized artificial membrane partitioning coefficients, the fraction of compound ionized and plasma protein binding data. *European Journal of Medicinal Chemistry*, 44(11), 4455–4460. doi:10.1016/j.ejmech.2009.06.004.
- Svobodova Varekova, R., Geidl, S., Ionescu, C. M., Skrehota, O., Kudera, M., Sehnal, D., Bouchal, T., Abagyan, R., Huber, H. J., & Koca, J. (2011). Predicting pK(a) values of substituted phenols from atomic charges: Comparison of different quantum mechanical methods and charge distribution schemes. *Journal of Chemical Information and Modeling*, 51(8), 1795–1806. doi:10.1021/ci200133w.
- Tao, L., Zhang, P., Qin, C., Chen, S. Y., Zhang, C., Chen, Z., Zhu, F., Yang, S. Y., Wei, Y. Q., & Chen, Y. Z. (2015). Recent progresses in the exploration of machine learning methods as in-silico ADME prediction tools. *Advanced Drug Delivery Reviews*, 86, 83–100. doi:10.1016/j.addr.2015.03.014.
- Thomas, S. R., Baconnier, P., Fontecave, J., Francoise, J. P., Guillaud, F., Hannaert, P., Hernandez, A., Le Rolle, V., Maziere, P., Tahi, F., & White, R. J. (2008). SAPHIR: A physiome core model of body fluid homeostasis and blood pressure regulation. *Philosophical Transactions: A Mathematical, Physical and Engineering Sciences*, 366, 3175–3197. doi:10.1098/rsta.2008.0079.
- Thomas, S. R., Layton, A. T., Layton, H. E., & Moore, L. C. (2006). Kidney modelling: Status and perspectives. *Proceedings of IEEE*, 94(4), 740–752.
- Thompson, J. D., Cramer, C. J., & Truhlar, D. G. (2003). Predicting aqueous solubilities from aqueous free energies of solvation and experimental or calculated vapor pressures of pure substances. *The Journal of Chemical Physics*, 119(3), 1661. doi:10.1063/1.1579474.
- Tomasi, J., Mennucci, B., & Cammi, R. (2005). Quantum mechanical continuum solvation models. *Chemical Reviews*, 105(8), 2999–3093.
- Varma, M. V., Steyn, S. J., Allerton, C., & El-Kattan, A. F. (2015). Predicting clearance mechanism in drug discovery: Extended clearance classification system (ECCS). *Pharmaceutical Research*, 32(12), 3785–3802. doi:10.1007/s11095-015-1749-4.
- Wang, J., & Hou, T. (2015). Advances in computationally modeling human oral bioavailability. *Advanced Drug Delivery Reviews*, 86, 11–16. doi:10.1016/j.addr.2015.01.001.
- Windshugel, B., Jyrkkarinne, J., Poso, A., Honkakoski, P., & Sippl, W. (2005). Molecular dynamics simulations of the human CAR ligand-binding domain: Deciphering the molecular basis for constitutive activity. *Journal of Molecular Modelling*, 11(1), 69–79. doi:10.1007/s00894-004-0227-4.
- Wood, I., & Pickholz, M. (2013). Concentration effects of sumatriptan on the properties of model membranes by molecular dynamics simulations. *European Biophysics Journal*, 42(11–12), 833–841. doi:10.1007/s00249-013-0932-y.
- Xu, X., Zhang, W., Huang, C., Li, Y., Yu, H., Wang, Y., Duan, J., & Ling, Y. (2012). A novel chemometric method for the prediction of human oral bioavailability. *International Journal of Molecular Sciences*, 13(6), 6964–6982. doi:10.3390/ijms13066964.
- Yang, X., Gandhi, Y. A., Duiquan, D. B., & Morris, M. E. (2009). Prediction of biliary excretion in rats and humans using molecular weight and quantitative structure-pharmacokinetic relationships. *The AAPS Journal*, 11(3), 511–525. doi:10.1208/s12248-009-9124-1.

- Yap, C. W., & Chen, Y. Z. (2005). Prediction of cytochrome P450 3A4, 2D6, and 2C9 inhibitors and substrates by using support vector machines. *Journal of Chemical Information and Modeling*, *45*(4), 982–992.
- You, H., Kim, G. E., Na, C. H., Lee, S., Lee, C. J., Cho, K. H., Akiyama, Y., Ishida, T., & No, K. T. (2014). An empirical model for gas phase acidity and basicity estimation. *SAR and QSAR in Environmental Research*, *25*(2), 91–115. doi:[10.1080/1062936X.2013.864997](https://doi.org/10.1080/1062936X.2013.864997).
- You, H., Lee, K., Lee, S., Hwang, S. B., Kim, K. Y., Cho, K. H., & No, K. T. (2015). Computational classification models for predicting the interaction of compounds with hepatic organic ion importers. *Drug Metabolism and Pharmacokinetics*, *30*(5), 347–351. doi:[10.1016/j.dmpk.2015.06.004](https://doi.org/10.1016/j.dmpk.2015.06.004).
- Yu, M. J. (2010). Predicting total clearance in humans from chemical structure. *Journal of Chemical Information and Modeling*, *50*(7), 1284–1295.
- Zhivkova, Z., & Doytchinova, I. (2012). Prediction of steady-state volume of distribution of acidic drugs by quantitative structure-pharmacokinetics relationships. *Journal of Pharmaceutical Sciences*, *101*(3), 1253–1266. doi:[10.1002/jps.22819](https://doi.org/10.1002/jps.22819).
- Zhu, J., Wang, J., Yu, H., Li, Y., & Hou, T. (2011). Recent developments of in silico predictions of oral bioavailability. *Combinatorial Chemistry & High Throughput Screening*, *14*(5), 362–374.

Predictive QSAR Modeling: Methods and Applications in Drug Discovery and Chemical Risk Assessment

57

Alexander Golbraikh, Xiang Simon Wang, Hao Zhu, and Alexander Tropsha

Contents

QSAR Methodology: Summary of Approaches for Model Building and Validation	2304
Data Preparation	2305
The Problem of Outliers	2308
QSAR Model Development	2309
QSAR Methods	2310
Target Functions	2313
Continuous QSAR Models	2313
Target Functions and Validation Criteria for Classification QSAR Models	2314
Target Functions and Validation Criteria for Category QSAR Models	2314
Applicability Domains	2315
Y-randomization	2317
External Validation	2318
“Good Practices” in QSAR Modeling: Examples of Models and Their Application to Virtual Screening and Lead Identification	2319
QSAR-Aided Discovery of Novel Anticonvulsant Compounds	2319
QSAR-Enabled Discovery of Novel Anticancer Agents	2321
QSAR Enabled Discovery of Novel Geranylgeranyltransferase I Inhibitors (GGTIs)	2322
“Good Practices” in QSAR Model Development: Applications to Toxicity Modeling	2323
Quantitative Structure In Vitro–In Vivo Relationship Modeling	2325
Using “Hybrid” Descriptors for QSIIR Modeling of Rodent Carcinogenicity	2327
Using “Hybrid” Descriptors for the QSIIR Modeling of Rodent Acute Toxicity	2327
Collaborative and Consensus Modeling of Aquatic Toxicity	2329

A. Golbraikh (✉) • X.S. Wang • H. Zhu

Laboratory for Molecular Modeling and Carolina Center for Exploratory Cheminformatics Research, Division of Medicinal Chemistry and Natural Products, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA

e-mail: golbraik@email.unc.edu; xiang.wang@Howard.edu; haozhu@email.unc.edu

A. Tropsha

Division of Medicinal Chemistry and Natural Products, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC, USA

e-mail: alex_tropsha@unc.edu

Universal Statistical Figures of Merit for All Models	2330
Consensus QSAR Models of Aquatic Toxicity; Comparison Between Methods and Models	2331
Conclusions: Emerging Chemical/Biological Data and QSAR Research Strategies	2332
Bibliography	2333

Abstract

Quantitative structure–activity relationship (QSAR) modeling is the major cheminformatics approach to exploring and exploiting the dependency of chemical, biological, toxicological, or other types of activities or properties on their molecular features. QSAR modeling has been traditionally used as a lead optimization approach in drug discovery research. However, in recent years QSAR modeling found broader applications in hit and lead discovery by the means of virtual screening as well as in the area of drug-like property prediction, and chemical risk assessment. These developments have been enabled by the improved protocols for model development and most importantly, model validation that focus on developing models with independently validated external prediction power. This chapter reviews the predictive QSAR modeling workflow developed in this laboratory that incorporates rigorous procedures for QSAR model development, validation, and application to virtual screening. It also provides several examples of the workflow application to the identification of experimentally confirmed hit compounds as well as to chemical toxicity modeling. We believe that methods and applications considered in this chapter will be of interest and value to researchers working in the field of computational drug discovery and environmental chemical risk assessment.

QSAR Methodology: Summary of Approaches for Model Building and Validation

In order to find new leads in the process of drug design and discovery, there is a need for efficient and robust computational procedures that can be used to screen chemical databases and virtual libraries against molecules with known activities or properties. For this purpose, quantitative structure–activity relationship (QSAR) analysis is widely used. QSAR modeling provides an effective way for establishing and exploiting the relationship between chemical structures and their biological actions toward the development of novel drug candidates. Theoretically, QSAR analysis is the application of mathematical and statistical methods for the development of models for the prediction of biological activities or properties of compounds. Formally, a QSAR model can be expressed in the following generic format:

$$\text{Predicted Biological Activity} = \text{Function} (\text{Chemical Structure}) \quad (1)$$

A QSAR procedure tries to minimize the error of prediction, for example, in the form of the sum of squares between predicted and observed activities.

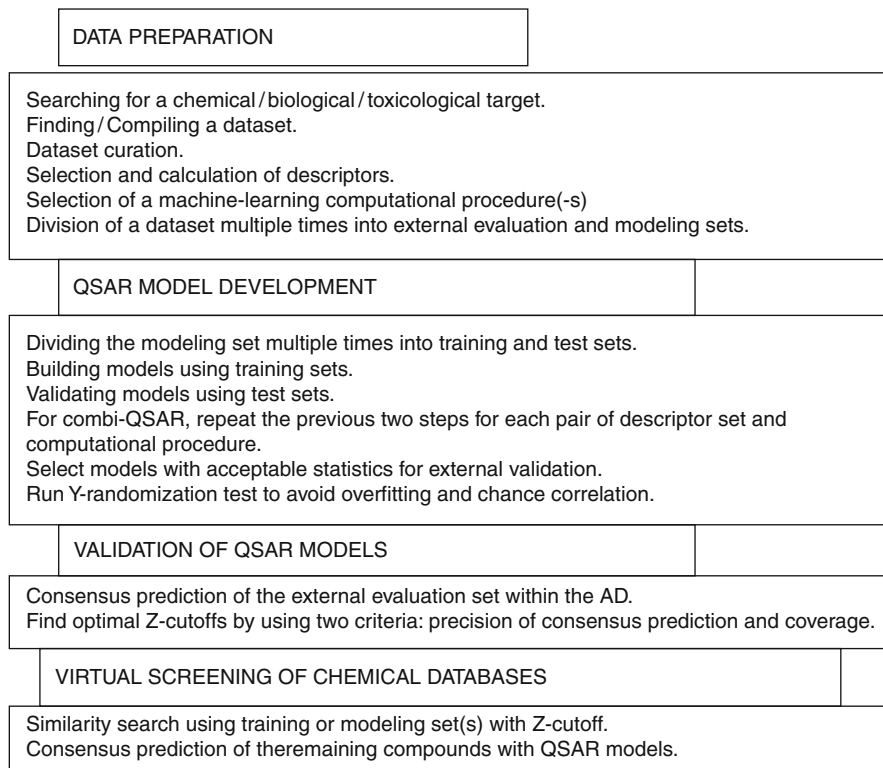


Fig. 1 Major steps of QSAR modeling

The process of QSAR model development can be divided into three parts: data preparation, data analysis, and model validation (Fig. 1). Model validation should include establishment of model applicability domain (AD). Recently, the European Organization for Economic Co-operation and Development (OECD) developed a set of principles for the development and validation of QSAR models, which, in particular, requires “appropriate measures of goodness-of-fit, robustness, and predictivity” (Organisation 2008). The OECD guidance document especially emphasizes that QSAR models should be rigorously validated using external sets of compounds that were not used in the model development.

Data Preparation

The first part of QSAR analysis includes selection of a molecular dataset for QSAR studies, acquiring or calculation of molecular descriptors (quantities characterizing molecular structures), and selection of a QSAR (statistical analysis and correlation) method. Datasets for QSAR studies can be found in research papers or electronic

databases available either publicly (PubChem 2010; BindingDB (Liu et al. 2007); ChEMBL 2010; DSSTox 2008; NIMH Psychoactive Drug Screening (PDSP) 2010) or commercially (e.g., Wombat (Olah et al. 2007) or MDDR 2009); more examples are given in a recent review (Oprea and Tropsha 2006). The dataset should include biological activity values for all compounds (e.g., binding energies to a receptor, or inhibition constants IC_{50} , or in case of toxicity modeling, lethal concentration in water LC_{50} , or lethal dose LD_{50} , etc.) preferably measured in the same lab using the same experimental method. If these experimental data are not available from one lab or one source, and the correlation between measurements made in different labs or by different methods cannot be established, they may not be used directly in QSAR studies. Instead, compounds in the dataset should be given a rank or assigned to categories of activities: for example, a compound can be very active, moderately active, or inactive. In the majority of such cases, binary classification is used, in which a compound is classified as either active or inactive. Another situation may arise, when compounds in the dataset naturally belong to different classes, for example, they are ligands to different receptors. In this case, the types of ligand specificity for a target can be considered as classes of compound activities, and the goal of QSAR analysis becomes to achieve accurate prediction of the target specificity for a new compound.

According to the nature of the activity data, QSAR studies can be divided into continuous (activities, i.e., response variable, takes many different values from within some interval), category (activities are represented by ranks or ordinal numbers), and classification (activities are different types of biological properties which cannot be rank ordered) approaches.

Prior to QSAR modeling, a dataset should be curated, that is, all structures should be verified with respect to their correct representation in the dataset; structures containing atoms, for which there are no parameters for descriptor calculation should be removed; structures consisting of several disconnected parts should be removed; salts should be removed; a problem of isomerism should be addressed; and duplicate structures should be removed. There are different tools available for dataset curation. For example, Molecular Operating Environment (MOE) (2008) includes DatabaseWash tool. It allows changing molecules' names, adding or removing hydrogen atoms, removing salts and heavy atoms, even if they are covalently connected to the rest of the molecule, and changing or generating the tautomers and protomers (cf. the MOE manual for more details). Various database curation tools are included in ChemAxon (2008) as well. If commercial software tools such as MOE are unavailable (notably, ChemAxon software is free to academic investigators), one can use standard UNIX/LINUX tools to perform some of the dataset cleaning tasks (Tropsha and Golbraikh 2010). It is important to have some freely available molecular format converters such as OpenBabel (2010) or MolConverter from ChemAxon (2008). Major procedures for database curation are discussed in our recent paper (Fourches et al. 2010).

After the dataset is selected and curated, the next task is the acquisition or calculation of descriptors. According to an excellent monograph titled Handbook of Molecular Descriptors by Roberto Todeschini and Vivian Consonni (2000)

molecular descriptors can be grouped into zero-dimensional [0D] (sometimes referred to as constitutional descriptors), one-dimensional [1D] (e.g., counts of different molecular groups, physicochemical properties of compounds, etc.), two-dimensional [2D] (invariants of molecular graphs, e.g., connectivity indices, information indices, counts of paths and walks, etc.), three-dimensional [3D], which are based on geometrical spatial properties of molecules [e.g., Comparative Molecular Field Analysis (CoMFA) descriptors (Tripos 2010) which are values of steric and electrostatic fields around aligned molecules, and different CoMFA-like descriptors (Klebe 1998; Kubinyi et al. 1998; Robinson et al. 1999)], and some other descriptors. Some descriptors can be experimental or calculated physicochemical properties of molecules such as molecular weight, molar refraction, energies of HOMO and LUMO, normal boiling point, octanol/water partition coefficient, molecular surface, molecular volume, etc.

Herein, we will not discuss different types of descriptors in detail but mention briefly major descriptor software. Most of descriptors included in the Handbook of Molecular Descriptors (Todeschini and Consonni 2000) can be calculated by the Dragon software (Dragon 2007). Molconn-Z (2007) is another widely used descriptor calculation software which calculates more than 800 descriptors. A relatively small, but diverse set of molecular descriptors can be calculated by the MOE (2008) software. Chirality molecular topological descriptors (CMTDs) developed in our laboratory append 2D descriptors by conformation-independent chirality and ZE-isomerism topological indices (Golbraikh and Tropsha 2003; Golbraikh et al. 2001, 2002). Another group of descriptors frequently used in our laboratory is atom-pair (AP) descriptors (Carhart et al. 1985). Each descriptor is defined as a count of pairs of atoms of certain types being away from each other on a certain topological distance (2D AP descriptors) or a Euclidean distance within certain intervals (3D AP descriptors); chirality AP descriptors can be calculated as well (Kovatcheva et al. 2005).

Many descriptors calculated from the knowledge of 3D structure of molecules (3D descriptors) have been developed and published as well. Although these are inherently more rigorous, one should keep in mind that their calculation is much more time and resource consuming. In many QSAR applications, the calculation of 3D descriptors should be preceded by conformational search and 3D structure alignment. However, even for rigid compounds, it is not generally known whether the alignment corresponds to real positions of molecules in the receptor binding site (Cherkasov 2008). There are different conformational analysis and pharmacophore modeling tools included in molecular modeling packages such as MOE (2008), Sybyl (there are LINUX and MS Windows versions) (Tripos 2010), Discovery Studio (2010), LigandScout (2010), etc. It has been demonstrated that in many cases QSAR models based on 2D descriptors have comparable (or even superior) predictivity than models based on 3D descriptors (Bures and Martin 1998; Golbraikh et al. 2001; Hoffman et al. 1999; Zheng and Tropsha 2000). Thus when 3D QSAR studies are necessary, if possible, 3D alignment of molecules should be preferably obtained by docking studies. VolSurf (Crivori et al. 2000; Cruciani et al. 2000) and GRIND (Pastor et al. 2000) descriptors are examples of alignment-free

3D descriptors. But their calculation still requires extensive conformational analysis of molecules. Both VolSurf and GRIND descriptors are available in Sybyl (VolSurf and Almond modules) (Tripos 2010). Various types of descriptors can be calculated by different modules of Schrodinger software (2010). Virtually, any molecular modeling software package contains sets of its own descriptors and there are many other descriptors not mentioned here that can be found in the specialized literature.

There are sets of descriptors that take values of zero or one depending on the presence or absence of certain predefined molecular features (or fragments) such as oxygen atoms, aromatic rings, rings, double bonds, triple bonds, halogens, and so on. These sets of descriptors are called molecular fingerprints or structural keys. Such descriptors can be represented by bit strings and many are found in popular software packages. For instance, several different sets of such descriptors are included in MOE (2008), Sybyl (Tripos 2010), and others, and examples of their use can be found in the published literature (McGregor and Pallai 1997; Waller 2004). Molecular holograms are similar to fingerprints; however, they use counts of features rather than their presence or absence. For example, holograms are included in the Sybyl HQSAR module (Tripos 2010). There are also more recent approaches when molecular features are not predefined a priori (as fingerprints discussed above) but are identified for each specific dataset. For example, frequent subgraph mining approaches developed independently at the University of North Carolina (Huan et al. 2006) and at the Louis Pasteur University in Strasbourg (Horvath et al. 2007) can find all frequent closed subgraphs (i.e., subgraph descriptors) for given datasets of compounds described as chemical graphs. A large and diverse set of 2D descriptors can be generated by MOLD2 software (Hong et al. 2008) available from FDA. A wide variety of descriptors are included in ADRIANA software (Gasteiger 2006).

Prior to QSAR studies, processing of descriptors is required. It includes: exclusion of descriptors having the same value for all compounds in the dataset as well as duplicate descriptors. To avoid higher influence on QSAR models of descriptors with higher variance, all descriptors are usually normalized (in most cases, range scaling or autoscaling is used). Molecular holograms or AP descriptors do not need to be normalized. Molecular field values around molecules are also not normalized. Preferably, descriptors with low variance and one of the highly correlated pair of descriptors should be excluded as well.

Finally, data for QSAR model development can be represented in a form of a table (see Table 1), in which each compound is a row and each descriptor as well as activity is a column.

The Problem of Outliers

Success of QSAR modeling depends on the appropriate selection of a dataset for QSAR studies. In a recent editorial of the *Journal of Chemical Information and Modeling*, Maggiora (2006) noticed that one of the main deficiencies of many chemical datasets is that they do not fully satisfy the main hypothesis underlying all

Table 1 QSAR table

Compound	Descriptor 1	Descriptor 2	...	Descriptor N	Activity
1	X ₁₁	X ₁₂	...	X _{1N}	Y ₁
2	X ₂₁	X ₂₂	...	X _{2N}	Y ₂
...
M	X _{M1}	X _{M2}	...	X _{MN}	Y _M

QSAR studies: Similar compounds are expected to have similar biological activities or properties. Maggiora defines the “cliffs” in the descriptor space where the properties change so rapidly, that, in fact adding or deleting one small chemical group can lead to a dramatic change in the compound’s property. In other words, small changes of descriptor values can lead to large changes in molecular properties. Generally, in this case there could be not just one outlier, but a subset of compounds properties of which are different from those on the other “side” of the cliff. In other words, cliffs are areas where the main QSAR hypothesis does not hold. So cliff detection remains a major QSAR problem that has not been adequately addressed in most of the reported studies.

There are two types of outliers we must be aware of: leverage (or structural) outliers and activity outliers. In case of activity outliers the problem of “cliffs” should be addressed as well. Recently, different approaches to find activity outliers have been published (Bajorath et al. 2009; Guha and Van Drie 2008a, b; Sisay et al. 2009). We have suggested that Grubb’s (Environmental Protection Agency 1992) and Dixon’s (Fallon et al. 1997) statistical tests can be used to find activity outliers (Tropsha and Golbraikh 2010). Structural outliers can be defined as compounds that are largely dissimilar to all other compounds in the descriptor space. The methods of finding them are similar to finding compounds out of QSAR model applicability domains (Tropsha and Golbraikh 2010) that is discussed below.

QSAR Model Development

The ultimate goal of QSAR analysis is the development of validated models for accurate and precise prediction of biological activities of compounds which could be potential leads in the process of drug discovery. Eventually, predictions should be confirmed by experimental validation. The general QSAR modeling workflow is represented in Fig. 2. Following the data curation step, we start by randomly selecting a fraction of compounds (typically, 10–20%) as an external evaluation set. The Sphere Exclusion protocol implemented in our laboratory (Golbraikh and Tropsha 2002; Golbraikh et al. 2003) is then used to rationally divide the remaining subset of compounds (the modeling set) multiple times into pairs of training and test sets that are used for model development and validation, respectively. We employ multiple QSAR techniques based on the combinatorial exploration of all possible pairs of descriptor sets and various supervised data analysis techniques

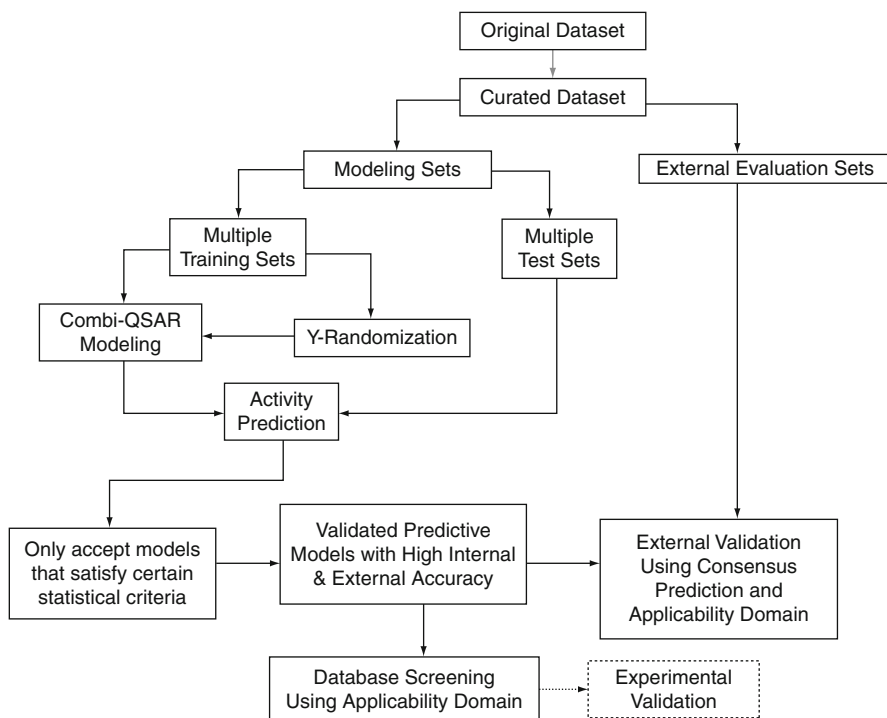


Fig. 2 Predictive QSAR modeling workflow

(combi-QSAR) (Fig. 3) and select models characterized by high accuracy in predicting both training and test sets data. Validated models are finally tested using the external evaluation set. The critical step of the external validation is the use of applicability domains (ADs). If external validation demonstrates the significant predictive power of the models, we employ them for virtual screening of available chemical databases (e.g., ZINC (Irwin and Shoichet 2005)) to identify putative active compounds and work with collaborators who could validate such hits experimentally. The entire approach is described in detail in several recent papers and reviews (Tropsha 2005; Tropsha and Golbraikh 2007).

QSAR Methods

QSAR modeling techniques employ various methods of multidimensional data analysis as well as supervised machine learning used in different areas of research in natural and social sciences such as biological sciences, geography, psychology, medicine, economics, signal processing, speech recognition, forensic studies, etc. Herein, it is impossible to discuss all the methods used in QSAR analysis. Instead, we will name only some of them. All these methods can be classified into linear and

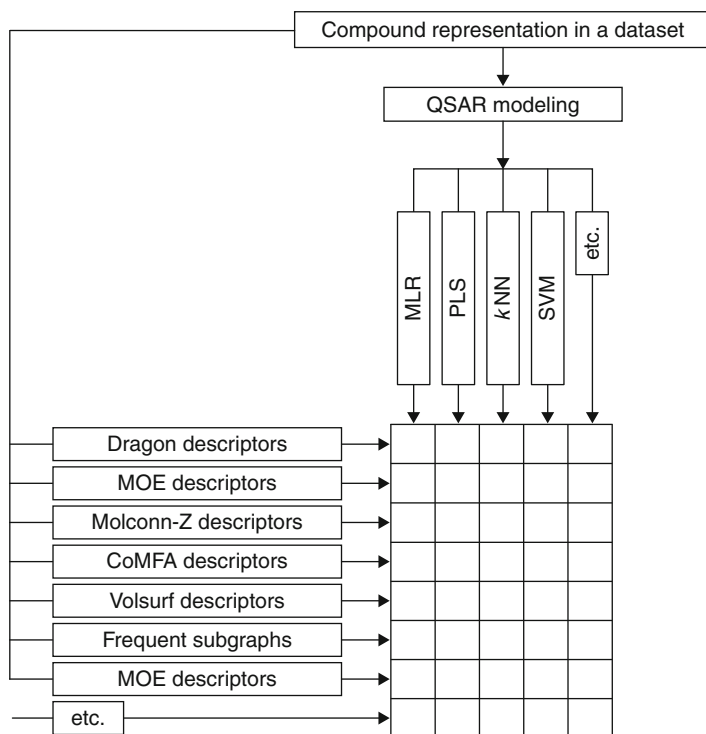


Fig. 3 Combinatorial QSAR modeling

nonlinear approaches. Linear methods include simple and multiple linear regression (MLR), principal component regression (PCR), partial least squares (PLS), etc. The main distinctive characteristic of these methods is the linearity of the function approximating the biological activity (see Eq. 1) of their arguments (which are molecular descriptors). In linear discriminant analysis (LDA), linear combinations of descriptors are built, which define hyperplanes that separate representative points of different classes of compounds in the multidimensional descriptor space.

Nonlinear methods can be based derived from linear or based on more complex approaches that predict compound activities from their descriptors by the means of nonlinear relationships. For example, if nonlinear terms (like squares, products, or logarithms of some descriptors) are added to a linear regression, it becomes nonlinear regression. Many nonlinear methods are derived from linear methods via transforming them by a so-called kernel trick. Calculations are executed in a so-called feature space where linear methods are applied. The advantage of these methods is that there is no need to directly calculate the transformation functions. Examples of such methods include non-linear support vector machines (SVMs) and support vector regression (SVR) methods (Berk 2008; Vapnik 2000), nonlinear discriminant analysis, kernel-PCA, kernel-PLS, etc. In the multidimensional feature

space, SVM builds a soft margin hyperplane, which separates points belonging to two different classes, or more hyperplanes to separate points of larger number of classes. In contrast, SVR builds a hyperplane such that as many points as possible are within the margin. Good SVM tutorial was written by Burges (1998), and SVR tutorial by Smola and Schoelkopf (2004). Other non-linear methods include k-nearest neighbors QSAR, in which the activity of a compound is predicted as a (weighted) average of activities of its nearest neighbors. k-nearest neighbor methods can include stochastic (Zheng and Tropsha 2000) or stepwise variable (descriptor) selection (Ajmani et al. 2006).

Another large group of generally nonlinear methods are artificial neural networks (ANNs) (Neural Networks 1996; Salt et al. 2006; Zupan and Gasteiger 1999). Ensembles of ANNs can make use of bagging and boosting approaches (Agrafiotis et al. 2002). ANNs consist of groups of artificial neurons. In feed-forward back-propagation neural networks (Neural Networks 2010), neurons are organized in input, hidden, and output layers. Input layer neurons receive descriptor values of compounds, which are passed with different weights to the hidden layer neurons. A neuron activation function is then applied at each neuron to the sum of weighted inputs, and the results are passed to the output layer neurons, which calculate predicted activities of compounds. During training process, parameters of neuron functions and weights are adjusted so that the total error of predictions is minimized. There are network architectures with multiple hidden layers.

Recursive partitioning (RP) methods build decision trees in order to precisely assign compounds to their classes. The tree consists of one root node containing all objects (compounds), intermediate (or decision), and leaf (terminal) nodes. A measure of node purity is introduced; for example, it could be the ratio of counts of compounds belonging to majority and minority class in a node. At each node, the procedure tries to partition the data to increase the purity measure, that is, to make the difference between sum of child node purities and parent node purity as higher as possible. Analysis is based on descriptor value distributions between classes at the node. If such a partition at the node is impossible, it becomes a leaf node. Additional criteria may be imposed on the minimum number of compounds in a leaf node, etc. Compounds in each node satisfy certain descriptor criteria. After growing, some leaves are consecutively removed based on the improvement of classification at them (so-called pruning of a tree). Without pruning, the tree could be overfitted. Prediction process consists of moving a query compound up the tree (based on its descriptor values) until it reaches a leaf node. Predicted class of a compound is defined as that of the majority class in this node. There are also RP regression methods which are used, if response variable is continuous. There are several RP algorithms widely used such as Classification and Regression Trees (CART (Berk 2008)), C4.5 (Quinlan 1993), C5.0 (2008), etc.

Random Forest methods (Breiman 2001; Random Forests 2001) construct ensembles of trees based on multiple random selections of subsets of descriptors and bootstrapping of compounds. The compounds not selected in a particular bootstrapping are considered as a so-called out of bag set, and used as the test

set. The trees are not pruned. Best trees in the forest are chosen for consensus prediction of external compounds. The method can include bagging (Berk 2008; Breiman 1996) and boosting (Berk 2008; Breiman 1998) approaches.

Target Functions

Based on the nature of the response variable, QSAR approaches can be grouped into classification, category, or continuous QSAR (vide infra). Classes are different from categories in a sense that the former cannot be ordered in any scientifically meaningful way, while the latter can be rank ordered.

Continuous QSAR Models

We suggested that the following validation criteria should be used for continuous QSAR models (Tropsha and Golbraikh 2010): (1) leave-one-out (LOO) cross-validated q^2 (which is also used as the target function, that is, it is optimized by the QSAR modeling procedure) (2) square of the correlation coefficient $R(R^2)$ between the predicted and observed activities of the test set; (3) coefficients of determination (predicted versus observed activities (R_0^2), and observed versus predicted activities ($R'_0{}^2$) for the test set) for regressions through the origin; (4) slopes k and k' of regression lines through the origin (predicted versus observed activities, and observed versus predicted activities for the test set). In our studies, we consider models acceptable, if they have (1) $q^2 > 0.5$; (2) $R^2 > 0.6$; (3) $(R^2 - R_0^2)/R^2 < 0.1$ and $0.85 \leq k \leq 1.15$ or $(R^2 - R'_0{}^2)/R^2 < 0.1$ and $0.85 \leq k' \leq 1.15$; (4) $|(R_0^2 - R'_0{}^2)| < 0.3$. Sometimes, stricter criteria are used (Tropsha and Golbraikh 2010).

In some papers, other criteria are used. For example, sometimes standard error of prediction is used instead of (or together with) R^2 . Standard error of prediction itself makes no sense until we compare it with the standard deviation for activities of the test set, which brings us back to the correlation coefficients. If used, mean absolute error (MAE) should be compared with the mean absolute deviation from the mean. Sometimes, F-ratio is calculated, which is the variance explained by the model divided by the unexplained variance. It is believed that the higher is the F-ratio, the better is the model. We suppose that when F-ratio is used, it must be always accompanied by the corresponding p-value.

Frequently, especially for linear models such as developed with multiple linear regression (MLR) or partial least squares (PLS) the adjusted R^2 is used:

$$R_{adj}^2 = 1 - (1 - R^2) \frac{n - 1}{n - c - 1}, \quad (2)$$

where n is the number of compounds in the dataset, and c is the number of variables (descriptors or principal components) included in the regression equation. It should

be recognized that $R_{adj}^2 \leq R^2$. The higher the number of explanatory variables c is, the lower R_{adj}^2 is. R_{adj}^2 is particularly important for linear QSAR models developed with variable selection. R_{adj}^2 is not a good criterion for variable selection k NN QSAR models, since contrary to regression methods, in the k NN algorithm descriptors are just selected or not selected, that is, their weights are either zero or one. As a result, much larger set of descriptors is selected by the k NN procedure than, for example, by stepwise regression.

Target Functions and Validation Criteria for Classification QSAR Models

We consider a classification QSAR model predictive, if the prediction accuracy characterized by the correct classification rate (CCR) for each class is sufficiently large:

$$CCR_{\text{class}} = \frac{N_{\text{class}}^{\text{corr}}}{N_{\text{class}}^{\text{total}}} \quad (3)$$

and the p-value for each CCR_{class} value is not higher than a predefined threshold (in case of two classes, the CCR_{class} threshold should not be lower than 0.65–0.70, and generally, for any number of classes, p-value should not be higher than 0.05 for each class).

For the classification QSAR with K classes, we shall use the following criterion

$$CCR = \frac{1}{K} \sum_{i=1}^K CCR_i = \frac{1}{K} \sum_{i=1}^K \frac{N_k^{\text{corr}}}{N_k^{\text{total}}} \quad (4)$$

along with the correct classification rate for each class (see Eq. 2). Criterion 4 is correct for both balanced and imbalanced (biased) datasets (i.e., when the number of compounds of each class is different). For imbalanced datasets, formula $N(\text{corr})/N(\text{total})$, where $N(\text{corr})$ and $N(\text{total})$ are the number of compounds predicted correctly and the total number of compounds in the dataset) is incorrect. QSAR procedure should maximize the CCR value calculated according to Eq. 4, and at the same time it should be penalized by too high differences between CCR values for different classes.

Target Functions and Validation Criteria for Category QSAR Models

Category QSAR with more than two classes should use target functions and validation criteria other than those used in classification QSAR. These target functions and validation criteria should consider errors as differences between

predicted and observed categories, or increasing functions of these differences. The total error of prediction over all compounds is the sum of all errors of predictions for individual compounds. Let n_{ij} be the number of compounds of category i assigned by a model to category j ($i, j = 1, \dots, K$). Then the total error is calculated as follows:

$$E = \sum_{i=1}^K \sum_{j=1}^K n_{ij} f(|i - j|). \quad (5)$$

where $f(|i - j|)$ is the increasing function of errors. In case of biased datasets, it would be important to normalize the errors for compounds of category i on the number of compounds in this category:

$$E = \sum_{i=1}^K \frac{1}{N_i} \sum_{j=1}^K n_{ij} f(|i - j|). \quad (6)$$

where N_i is the number of compounds of category i . QSAR procedure should minimize the total error of prediction calculated with 5 or 6. In practice, the accuracy can be defined as $A = 1 - E/E_{\text{exp}}$, where E_{exp} is the expected total error. Thus, QSAR procedure should maximize the target function A penalized by too high differences between CCR values for different classes.

More detailed consideration of target functions and validation criteria as well as different aspects of cost-sensitive learning, weighting, penalties, as well as threshold moving in QSAR studies are discussed in our recent review (Tropsha and Golbraikh 2010). General aspects of cost-sensitive learning are discussed by Elkan (The Foundations 2001) and Chen et al. (2004). Oversampling of the minority class, that is, inclusion of compounds of the minority class in the dataset more than once, is considered by Yen and Lee (2006), and Kubat and Matwin (1997). The opposite approach, called undersampling, that is, removing part of the majority class from the dataset, is considered by Japkowicz (2000). Using moving threshold for dividing compounds into active and inactive classes when continuous property values are available but one desires to use classification modeling approaches is considered by Zhou and Liu (2006). In QSAR studies, threshold is usually moved toward the larger class, which is easier to predict correctly.

Applicability Domains

Here we are approaching an extremely important problem of QSAR studies: model applicability domain (AD). Formally, a QSAR model can predict the target property for any compound for which chemical descriptors can be calculated. However, if a compound is highly dissimilar from all compounds of the modeling set, reliable prediction of its activity is unlikely to be realized. A concept of AD was

developed and used to avoid such an unjustified extrapolation in activity prediction. Applicability domains are one of the areas of intensive research. Different methods of defining AD exist. Among others, the following definitions are considered by Jaworska and colleagues (2005, 2008).

AD is defined as a hyperparallelepiped in the descriptor space in which representative points are distributed (Netzeva et al. 2006; Nikolova-Jeliazkova and Jaworska 2005; Saliner et al. 2006). Dimensionality of the hyperparallelepiped is equal to the number of descriptors, and the size of each dimension is defined by the minimum and maximum values of the corresponding descriptor or it stretches beyond these limits to some extent up to predefined thresholds.

AD is defined as a convex hull of points in the multidimensional descriptor space (Fechner et al. 2008).

The drawbacks of these definitions are as follows. Generally, the representative points are distributed not in the entire hyperparallelepiped or convex hull, but only in a small part of it. Another drawback is that structural outliers in the dataset can enormously increase the size of the hyperparallelepiped, and the area around the outlier will contain no other points. Consequently, for many compounds within the hyperparallelepiped or convex hull, prediction will be unreliable. Besides, if the number of linearly independent descriptors exceeds the number of compounds, the convex hull is not unique.

Leverage for a compound is defined as the corresponding diagonal element of the hat matrix (Afantitis et al. 2006). A compound is defined as outside of the AD, if its leverage L is higher than $3 K/N$, where K is the number of descriptors and N is the number of compounds. The drawbacks of the leverage-based AD are as follows. (a) for each external compound, it is necessary to recalculate leverage; (b) if there are cavities in the representative point distribution area, a query compound the representative point of which is in this area will be considered to be within the AD, while in fact it is far from all other compounds (Tropsha and Golbraikh 2010).

In our studies, the AD is defined as the Euclidean distance threshold DT between a query compound and its closest k -nearest neighbors of the training set. It is calculated as follows:

$$DT = \bar{y} + Z\sigma \quad (7)$$

Here, \bar{y} is the average Euclidean distance between each compound and its k -nearest neighbors in the training set k is optimized in the course of QSAR modeling, and the distances are calculated using descriptors selected by the optimized (model only), σ is the standard deviation of these Euclidean distances, and Z is an arbitrary cutoff parameter defined by a user (de Cerqueira et al. 2006; Hsieh et al. 2008; Kovatcheva et al. 2005; Zhang et al. 2008). We set the default value of this parameter Z at 0.5, which formally places the allowed distance threshold at the mean plus one-half of the standard deviation. We also define the AD in the entire descriptor space. In this case, the same Eq. 7 is used, $k = 1$, $Z = 0.5$, and Euclidean distances are calculated using all descriptors. Thus, if the distance of the external compound from its nearest neighbor in the training set within either the entire descriptor space or

the selected descriptor space exceeds these thresholds, the prediction is not made. We have also investigated changes of predictive power by changing the values of Z-cutoff. We have found that in general, starting from some Z-cutoff value, predictive power decreases while Z-cutoff value increases (Zhu et al. 2009), as expected. Instead of Euclidean distances, other distances and similarity measures can be used.

The predicted activity of a query compound by an ensemble of QSAR models is calculated as the average over all predicted values. In binary QSAR modeling, each model will predict the compound category as either 0 (inactive) or 1 (active); however, different models used in an ensemble may yield inconsistent predictions. Consequently, the averaged predicted activity value for an external compound resulting from the use of an ensemble of models may fall anywhere within the [0;1] range. For classification and category QSAR, the average predicted value is rounded to the closest integer (which is a class or category number); in the case of imbalanced datasets, rounding can be done using the moving threshold (*vide supra*). Predicted average classes or categories (before rounding) that are closer to the nearest integers are considered more reliable since such value indicates higher concordance between different models. For example, before rounding, one compound has the predicted value of 0.2, but the other has 0.4. Hence, both compounds are predicted to belong to class 0 but the prediction for the first compound is considered more reliable. Using these prediction values, additional constraint on the AD can be defined by a threshold of the absolute difference between the predicted and the rounded predicted activity. There are several other definitions of AD (Jaworska and Nikolova-Jeliazkova 2008; Tetko et al. 2006) based on probability density distributions, distances to models, etc.

Y-randomization

To establish model robustness, Y-randomization (randomization of the response variable) test should be used. This test consists of repeating all the calculations with scrambled activities of the training set. Ideally, calculations should be repeated at least five (better, more) times. The goal of this procedure is to establish whether models built with real activities of the training set have good statistics not due to overfitting or chance correlation. If predictive power for the training or the test set of all models built with randomized activities of the training set is significantly lower than that of models built with real activities of the training set, the latter ones are considered reliable. Using different parameters of the model development procedure, multiple QSAR models are built which have acceptable statistics. Suppose, the number of these models is m . Y-randomization test can also give n models with acceptable statistics. For acceptance of models developed with real activities of the training set, the condition $n \ll m$ should be satisfied. In (Kovatcheva et al. 2005) and (de Cerqueira et al. 2006), we have introduced the measure of robustness $R = 1 - n/m$. If $R > 0.9$, the models are considered robust and their high predictive accuracy cannot be explained by the chance correlation or overfitting. Y-randomization test is particularly important for small datasets.

Unfortunately, in many publications on QSAR studies, Y-randomization test is not carried out but all QSAR practitioners must be strongly encouraged to use this simple procedure.

External Validation

Our previous experience suggests that the consensus prediction, which is the average of predicted activities over all predictive models, always provides the most stable results (Zhang et al. 2008; Zhu et al. 2008), and thus naturally avoids the need for (the best) model selection based on the statistics for the training and test sets. The consensus prediction of biological activity for an external compound on the basis of several QSAR models is more reliable and provides better justification for the experimental exploration of hits.

External evaluation set compounds are predicted by models that have passed all validation criteria described above. Each compound is predicted by models for which the compound is within the AD. Actually, each external compound should be within the AD of the training set within the entire descriptor space as well (*vide supra*). A useful parameter for consensus prediction is the minimum number (or percentage) of models for which a compound is within the AD; it is defined by the user. If the compound is found within the AD of a lower number of models, it is considered to be outside of the AD. Prediction value is the average of predictions by all models. If a compound is predicted by more than one model, standard deviation of all predictions by these models is also calculated. For classification and category QSAR, the average prediction value is rounded to the closest integer (which is a class or category number); in case of imbalanced datasets, rounding can be done using the moving threshold.

Predicted average classes or categories (before rounding), which are closer to the nearest integers are considered more reliable (Zhang et al. 2008). Using these prediction values, AD can be defined by a threshold of the absolute difference between predicted and rounded predicted activity. For classification and category QSAR, the same prediction accuracy criteria are used as for the training and test sets. The situation is more complex for the continuous QSAR. In this case, if the range of activities of the external evaluation set is comparable to that for the modeling set, criteria (1)–(4) are used (see section “[Target Functions](#)”). Sometimes, however, the external evaluation set may have a much smaller range of activities than the modeling set, so it could be impossible to obtain sufficiently large R^2 value (and other acceptable statistical characteristics) for it. In this case, we recommend using the mean absolute error (MAE) or the standard error of prediction (SEP) as discussed in one of our previous publications (Tropsha and Golbraikh 2010).

We have used consensus prediction in many studies (de Cerqueira et al. 2006; Kovatcheva et al. 2005; Shen et al. 2004; Votano et al. 2004; Zhang et al. 2007, 2008; Zhu et al. 2008) and have shown that in most cases it gives better prediction and coverage than most of the individual predictive models. Thus, we recommend using

consensus prediction for virtual screening of chemical databases and combinatorial libraries for finding new lead compounds for drug discovery.

“Good Practices” in QSAR Modeling: Examples of Models and Their Application to Virtual Screening and Lead Identification

As discussed above, our experience in QSAR model development and validation has led us to establishing a complex but straightforward workflow summarized in Fig. 2. The last critical component of this workflow is the use of models to identify tentative active hits that should be validated in experimental laboratories, and we strongly encourage every computational scientist to use this ultimate model validation strategy. We note that this approach shifts the emphasis from ensuring good (best) statistics for the model that fits known experimental data toward generating testable hypotheses about purported bioactive compounds. Thus, the output of the modeling has exactly same format as the input, that is, chemical structures and (predicted) activities making model interpretation and utilization completely seamless for medicinal chemists. In our recent studies, we have been fortunate to recruit experimental collaborators who have validated computational hits identified through our modeling of anticonvulsants (Shen et al. 2004), HIV-1 reverse transcriptase inhibitors (Medina-Franco et al. 2005), D1 antagonists (Oloff et al. 2005), antitumor compounds (Zhang et al. 2007), beta-lactamase inhibitors (Hsieh et al. 2008), geranylgeranyltransferase inhibitors (Peterson et al. 2009), and others. The discovery of novel bioactive chemical entities is the primary goal of computational drug discovery, and the development of validated and predictive QSAR models is critical to achieve this goal. We note that such studies could only be done if there is sufficient data available for a series of tested compounds such that robust validated models could be developed using the workflow described in Fig. 2. We present several examples of these studies below to illustrate the use of QSAR models as virtual screening tools for lead identification.

QSAR-Aided Discovery of Novel Anticonvulsant Compounds

We have applied kNN (Zheng and Tropsha 2000) and simulated annealing – partial least squares (SA-PLS) (Cho et al. 1998) QSAR approaches to a dataset of 48 chemically diverse functionalized amino acids (FAAs) with anticonvulsant activity that were synthesized previously, and successful QSAR models of FAA anticonvulsants have been developed (Shen et al. 2002). Both methods utilized multiple descriptors such as molecular connectivity indices or atom-pair descriptors, which are derived from two-dimensional molecular topology. QSAR models with high internal accuracy were generated, with leave-one-out cross-validated $R^2(q^2)$ values ranging between 0.6 and 0.8. The q^2 values for the actual dataset were significantly higher than those obtained for the same dataset with randomly shuffled

activity values, indicating that models were statistically significant. The original dataset was further divided into several training and test sets, and highly predictive models providing q^2 values for the training sets greater than 0.5 and R^2 values for the test sets greater than 0.6.

In the second phase of modeling, we have applied the validated QSAR models to mining available chemical databases for new lead FAA anticonvulsant agents. Two databases have been explored: the National Cancer Institute (nci 2007) and Maybridge (2005) databases, including (at the time of that study) 237,771 and 55,273 chemical structures, respectively. Database mining was performed independently using ten individual QSAR models that have been extensively validated using several criteria of robustness and accuracy. Each individual model selected some number of hits as a result of independent database mining, and the consensus hits (i.e., those selected by all models) were further explored experimentally for their anticonvulsant activity. As a result of computational screening of the NCI database, 27 compounds were selected as potential anticonvulsant agents and submitted to our experimental collaborators. Of these 27 compounds, our collaborators chose two for synthesis and evaluation; their choice was based on the ease of synthesis and the fact that these two compounds had structural features that would not be expected to be found in active compounds based on prior experience. Several additional compounds, which were close analogs of these two were either taken from the literature or designed in our collaborator's laboratory. In total, seven compounds were resynthesized and sent to the NIH for the Maximum Electroshock (MES) test (a standard test for the anticonvulsant activity, which was used for the training set compounds as well). The biological results indicated that upon initial and secondary screening, five out of seven compounds tested showed anticonvulsant activity with ED_{50} less than 100 mg/kg, which is considered promising. Interestingly, all seven compounds were also found to be very active in the same tests performed on rats (a complete set of experimental data on rats for the training set were not available, and therefore no QSAR models for rats were built).

Mining of the Maybridge database yielded two additional promising compounds that were synthesized and sent to the NIH for the MES anticonvulsant test. One of the compounds showed moderate anticonvulsant activity of ED_{50} between 30 and 100 mg/kg (in mice), while the other was found to be a very potent anticonvulsant agent with ED_{50} of 18 mg/kg in mice (ip). In summary, both compounds were found to be very active in both mice and rats. Figure 4 shows chemical structures of experimentally confirmed hits that were identified by using validated QSAR models for virtual screening as applied to the anticonvulsant dataset. It is important to note that none of the compounds identified in external databases as potent anticonvulsants and validated experimentally belong to the same class of FAA molecules as the training set. This observation was very stimulating because it underscored the power of our methodology to identify potent anticonvulsants of novel chemical classes as compared to the training set compounds, which is one of the most important goals of virtual screening.

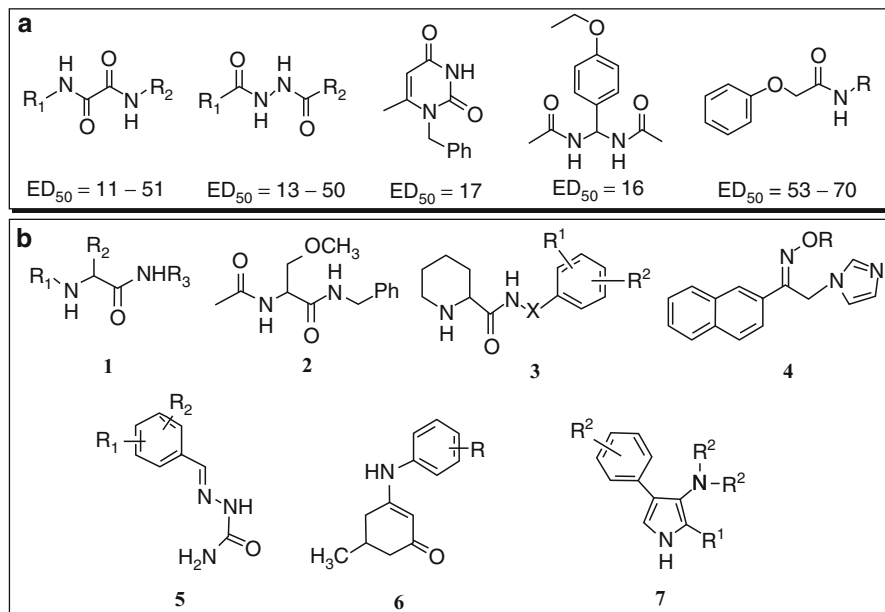


Fig. 4 Uniqueness of scaffolds for QSAR-based experimentally confirmed virtual screening hits (a) as compared to training set compounds; (b) for the anticonvulsant dataset

QSAR-Enabled Discovery of Novel Anticancer Agents

A combined approach of validated QSAR modeling and virtual screening was successfully applied to the discovery of novel tylophorine derivatives as anticancer agents (Zhang et al. 2007). QSAR models have been initially developed for 52 chemically diverse phenanthrene-based tylophorine derivatives (PBTs) with known experimental EC_{50} using chemical topological descriptors (calculated with the Molconn-Z program) and variable selection k-nearest neighbor (kNN) method. Several validation protocols have been applied to achieve robust QSAR models. The original dataset was divided into multiple training and test sets, and the models were considered acceptable only if the leave-one-out cross-validated $R^2(q^2)$ values were greater than 0.5 for the training sets and the correlation coefficient R^2 values were greater than 0.6 for the test sets. Furthermore, the q^2 values for the actual dataset were shown to be significantly higher than those obtained for the same dataset with randomized target properties (Y-randomization test), indicating that models were statistically significant. Ten best models were then employed to mine a commercially available ChemDiv Database (ca. 500 K compounds) resulting in 34 consensus hits with moderate to high predicted activities. Ten structurally diverse hits were experimentally tested and eight were confirmed active with the highest

experimental EC_{50} of 1.8 μM implying an exceptionally high hit rate (80 %). The same ten models were further applied to predict EC_{50} for four new PBTs, and the correlation coefficient (R^2) between the experimental and predicted EC_{50} for these compounds plus eight active consensus hits was shown to be as high as 0.57.

QSAR Enabled Discovery of Novel Geranylgeranyltransferase I Inhibitors (GGTIs)

The proper functioning of proteins often relies on posttranslational modification of the polypeptide leading to changes in chemical characteristics. Found at the extreme carboxyl terminus of the protein, one posttranslational “program” utilized for over 140 proteins is the so-called CaaX box, where “C” is a cysteine, “aa” is any aliphatic dipeptide, and “X” is the terminal residue that directs which of two prenyl groups is added (Cox and Der 2002; Zhang and Casey 1996). Protein geranylgeranyltransferase type I (GGTase-I) transfers the 20-carbon geranylgeranyl group to proteins including critical signaling molecules from many classes, for example, the Ras superfamily (including K-Ras, Rho, Rap, Cdc42, and Rac), several G-protein gamma subunits, protein kinases (rhodopsin kinase, phosphorylase kinase, and GRK7), and protein phosphatases (Casey and Seabra 1996; Sebti and Hamilton 2000). Several GGTIs have been developed that inhibit C20 lipid modification of GGTase-I substrates. GGTIs have been primarily developed for use as cancer therapeutics, particularly in cancers that have high levels, or activating mutations of geranylgeranylated proteins (Sebti and Hamilton 2000; Winter-Vann and Casey 2005).

The pharmacological data for 48 GGTIs reported in (Peterson et al. 2009) were generated as part of an iterative drug discovery program that led to GGTI-DU40 (Peterson et al. 2006). The structure of GGTI-DU40 can be discussed in the context of the CaaL peptide framework. There is a free amide group, a spacer domain relating to the dialiphatic motif, and critical sulfur as found in the requisite cysteine residue of GGTase-I’s substrates. Four additional GGTIs included in the data set were peptidomimetics as well. Importantly, the modeling set included compounds with different (chemical scaffolds), which in theory (and as we have established in our study, in practice) should have enabled the identification of chemically diverse hits from virtual screening.

Three different modeling techniques have been used to model GGTIs following our general combi-QSAR strategy (Fig. 3); the specific workflow as applied to the GGTI dataset is shown in Fig. 5. As the first step of our QSAR-based virtual screening, the preliminary filtering of the 9.5 million compounds in our screening library yielded 79 initial hits. This was done by using the global applicability domain of all 48 GGTIs in the modeling set. After consensus predictions by 104 validated kNN models, their predicted activities (pIC_{50}) were found ranging from 4.51 to 5.96. Only 47 hits, including two pairs of stereoisomers, showed high predicted activity ($pIC_{50} > 5.50$) as well as high model coverage and were designated as the final hits. Concurrently, two additional QSAR models were employed to reevaluate

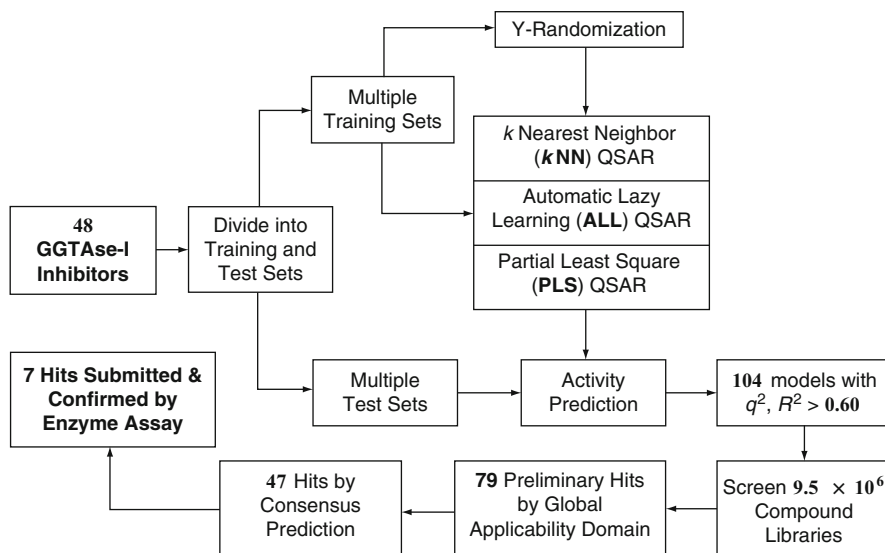


Fig. 5 The predictive QSAR modeling workflow illustrated for GGTIs Using purified recombinant GGTase-I as an enzyme source and GGpp and Ras-CVLL as substrates, seven hit compounds were tested in vitro as a matter of the experimental validation. The selection was based on high predicted activity, availability, and structural uniqueness. All tested compounds showed inhibition of GGTase-I with the pIC_{50} ranging from 3.63 to 5.44 (cf. Fig. 6)

those 79 hits in order to identify the consensus hits among all three methods. In the end, seven compounds were prioritized for experimental validation based on high predicted activity, uniqueness of structure, and availability.

The unexpected result was to identify several predicted actives that did not have a common ring feature in their structure. In fact, seven highly ranked hits had no apparent relationship with any of the training set molecules. They had furan, triazole, tetrazole, and pyridine cores in their scaffolds while all non-peptidomimetic compounds of the training set were based on a pyrazole core. Therefore, the seven hit compounds appeared to be the structurally novel hits. Figure 6b shows chemical structures of the three representative confirmed hits with novel scaffolds highlighted. This study reconfirmed the observation that we already emphasized earlier with anticonvulsant compounds that contrary to the common belief, QSAR-based virtual screening is capable of identifying experimentally confirmed hit compounds with novel scaffolds.

“Good Practices” in QSAR Model Development: Applications to Toxicity Modeling

Many compounds entering clinical studies do not survive as a good pharmacological lead to become a marketed drug. Chemical toxicity and safety have been regarded

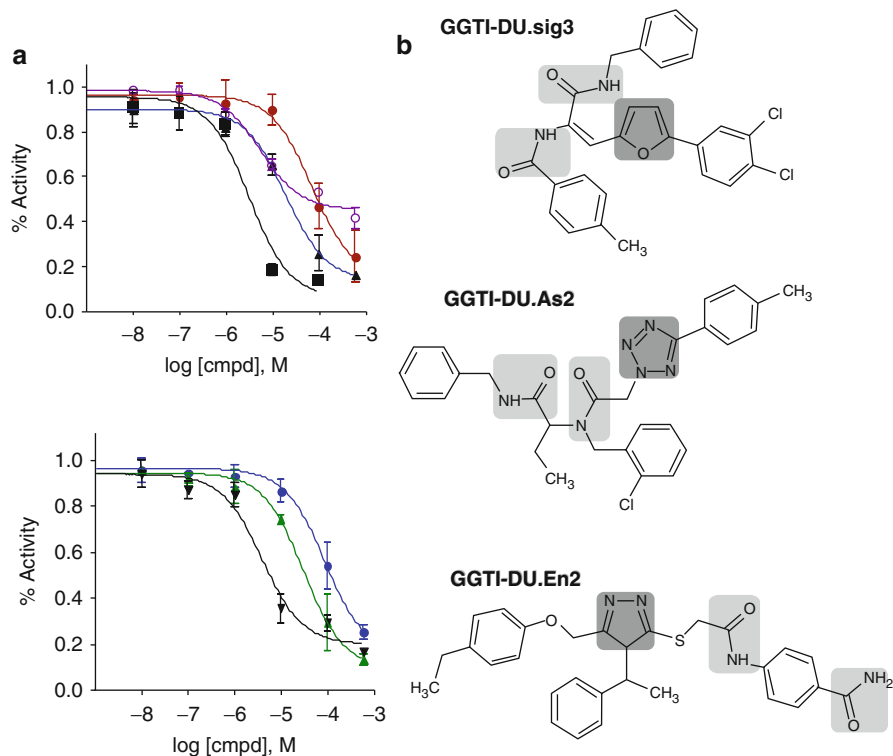


Fig. 6 Experimental validations of computational GGTI hits using GGTase-I in vitro activity assay. **(a)** Inhibition curves; **(b)** Chemical structures of three representative confirmed hits; the novel scaffolds in the structures are highlighted

as the major reason for attrition in the past decades (Kola and Landis 2004). However, evaluation of chemical toxicity and safety in vivo at the early stage of drug discovery process is expensive and time consuming. To replace the traditional animal toxicity testing and to understand the relevant toxicological mechanisms, many in vitro toxicity screens and computational toxicity models have been developed and implemented by academic institutes and pharmaceutical companies (Cheeseman 2005; Dash et al. 2009; Dix et al. 2007; Inglese et al. 2006; Park et al. 2009; Riley and Kenna 2004; Valerio 2009; Yang et al. 2009). In the past 15 years, innovative technologies that enable rapid synthesis and high throughput screening of large libraries of compounds have been adopted for toxicity studies. As a result, there has been a huge increase in the number of compounds and the associated testing data in different in vitro screens. With this data, it becomes feasible to reveal the relationship between the high throughput in vitro toxicity testing results and the low throughput in vivo low dose toxicity evaluation for the same set of compounds. Understanding these relationships could help us delineate

the mechanisms underlying animal toxicity of chemicals as well as potentially improve our ability to predict chemical toxicity using short-term bioassays.

The unique advantage of using a computational toxicity model in risk analysis is that a chemical could be evaluated for its toxicity potential even before it is synthesized. The computational toxicity tools based on QSAR models have been used to assist in predictive toxicological profiling of pharmaceutical substances for understanding drug safety liabilities (Durham and Pearl 2001; Jacobson-Kram and Contrera 2007; Muster et al. 2008; Valerio 2009), supporting regulatory decision making on chemical safety and risk of toxicity (Bailey et al. 2005), and are effectively enhancing an already rigorous US regulatory safety review of pharmaceutical substances (Valerio 2008). Predictive QSAR models of chemical toxicity are beginning to be used to evaluate compounds' safety in the pharmaceutical industry and environmental agencies (Durham and Pearl 2001; Snyder 2009). However, it has been reported that most QSAR models do not work well for evaluating *in vivo* toxicity, especially for external compounds (Zvinavashe et al. 2008, 2009). Several reviews were published recently that challenge the feasibility and reliability of QSAR models of chemical toxicity (Johnson 2008; Stouch et al. 2003). At the same time, experimental data resulting from short-term high throughput screening assays are emerging prompting the development of novel modeling approaches that can combine short-term assay data and conventional chemical descriptors of molecules to develop enhanced QSAR models of animal toxicity. We briefly review these emerging approaches and applications below.

Quantitative Structure In Vitro–In Vivo Relationship Modeling

To stress a broad appeal of the conventional QSAR approach, it should be made clear that from the statistical viewpoint QSAR modeling is a special case of general statistical data mining and data modeling where the data is formatted to represent objects described by multiple descriptors and the robust correlation between descriptors and a target property (e.g., chemical toxicity *in vivo*) is sought. In previous computational toxicology studies, additional physicochemical properties, such as water partition coefficient (logP) (Klopman et al. 2003), water solubility (Stoner et al. 2004), and melting point (Mayer and Reichenberg 2006) were used successfully to augment computed chemical descriptors and improve the predictive power of QSAR models. These studies suggest that using experimental results as descriptors in QSAR modeling could prove beneficial. The already available and rapidly growing HTS data for large and diverse chemical libraries makes it possible to extend the scope of the conventional QSAR in toxicity studies by using *in vitro* testing results as additional biological descriptors. Therefore, in some of the most recent toxicology studies, the relationships between various *in vitro* and *in vivo* toxicity testing results were generated (Forsby and Blaauboer 2007; Piersma et al. 2008; Schirmer et al. 2008; Sjoström et al. 2008). Based on these reports, we proposed a new modeling workflow called Quantitative Structure In vitro–In vivo Relationship (QSIIRQuantitative structure in vitro–in vivo relationship (QSIIR)

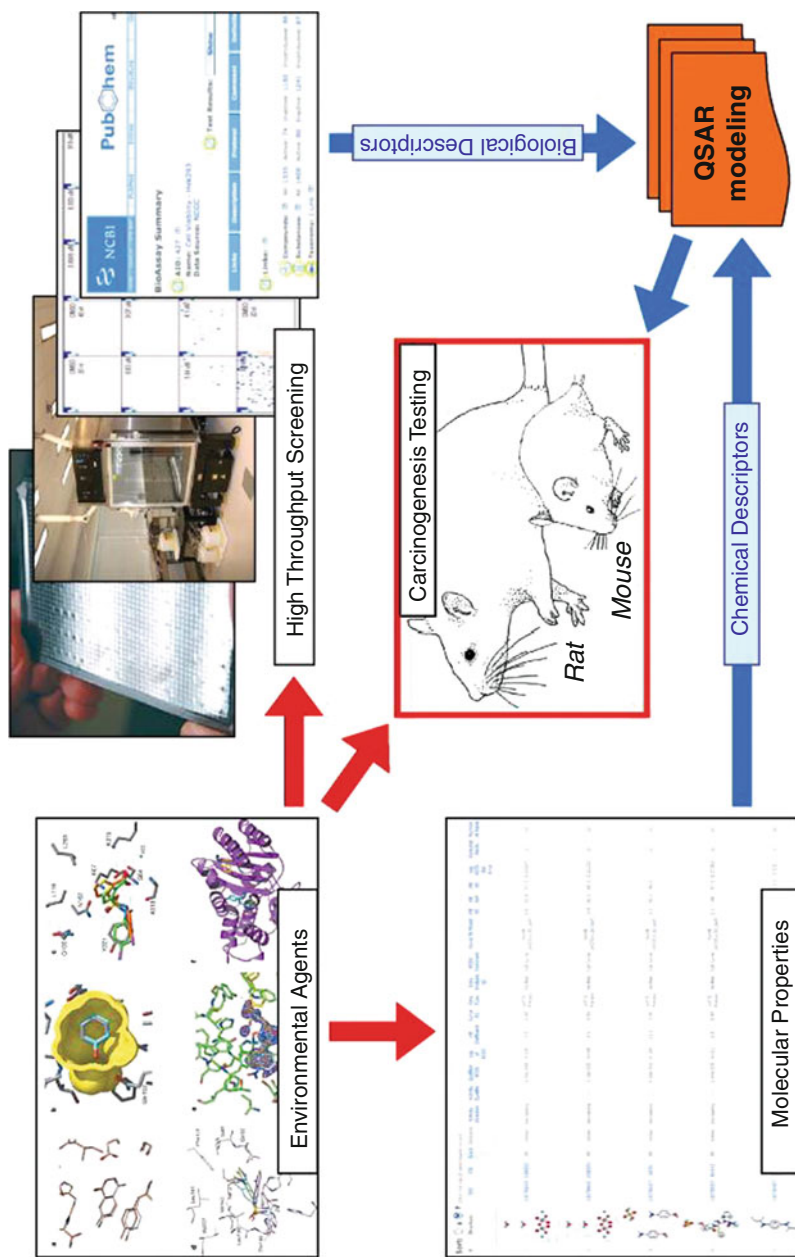


Fig. 7 Combining chemical and biological profiles as descriptors in QSAR modeling of chemical carcinogenicity

modeling) and used it in animal toxicity modeling studies (Zhu et al. 2008, 2009). The target properties of QSIIR modeling were still biological activities, such as different toxicity end points, but the content and interpretation of “descriptors” and the resulting models is different. This focus on the prediction of the same target property from different (chemical, biological, and genomic) characteristics of environmental agents affords an opportunity to most fully explore the source-to-outcome continuum of the modern experimental toxicology using cheminformatics approaches. Figure 7– provides visual illustration of the integrated QSIIR approach to in vivo toxicity modeling.

Using “Hybrid” Descriptors for QSIIR Modeling of Rodent Carcinogenicity

To explore efficient approaches for rapid evaluation of chemical toxicity and human health risk of environmental compounds, the National Toxicology Program (NTP), in collaboration with the National Center for Chemical Genomics (NCGC) has initiated an HTS Project (Inglese et al. 2006; Thomas et al. 2009). The first batch of HTS results for a set of 1,408 compounds tested in six human cell lines was released via PubChem. We have explored this data in terms of their utility for predicting adverse health effects of the environmental agents (Zhu et al. 2008). Initially, the classification k-nearest neighbor (kNN) QSAR modeling method was applied to the HTS data only for the curated dataset of 384 compounds. The resulting models had prediction accuracies for training, test (containing 275 compounds together), and external validation (109 compounds) sets as high as 89 %, 71 %, and 74 %, respectively. We then asked if HTS results could be of value in predicting rodent carcinogenicities. We identified 383 compounds for which data were available from both the Berkeley Carcinogenic Potency Database and NTP-HTS studies. We found that compounds classified by HTS as “actives” in at least one cell line were likely to be rodent carcinogens (sensitivity 77 %); however, HTS “inactives” were far less informative (specificity 46 %). Using chemical descriptors only, kNN QSAR modeling resulted in the overall external prediction accuracy of 62 % for rodent carcinogenicity. Importantly, the prediction accuracy of the model was significantly improved (to 73 %) when chemical descriptors were augmented by the HTS data, which were regarded as biological descriptors (Fig. 8). Thus, our studies suggested, for the first time, that combining HTS profiles with conventional chemical descriptors could considerably improve the predictive power of computational approaches in chemical toxicology.

Using “Hybrid” Descriptors for the QSIIR Modeling of Rodent Acute Toxicity

We used the cell viability qHTS data from NCGC as mentioned in the above section for the same 1,408 compounds but in 13 cell lines (Xia et al. 2008).

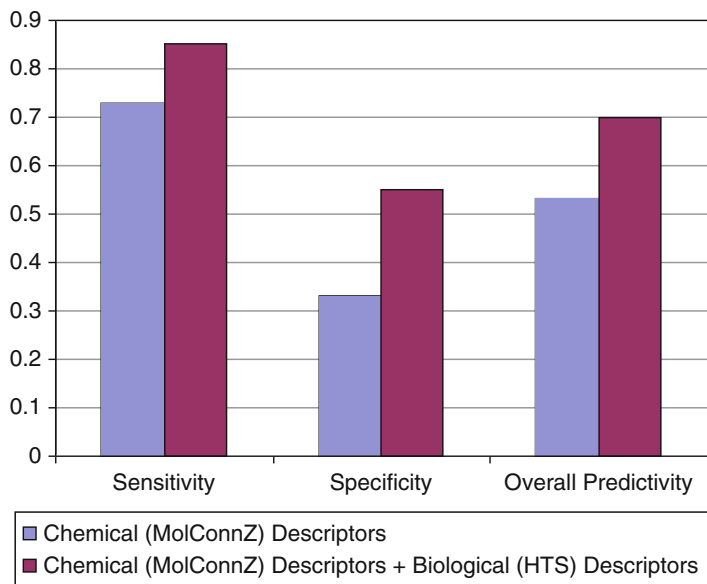


Fig. 8 Comparison of the prediction power of QSTR models of chemical carcinogenicity for the independent validation set using conventional versus hybrid descriptors

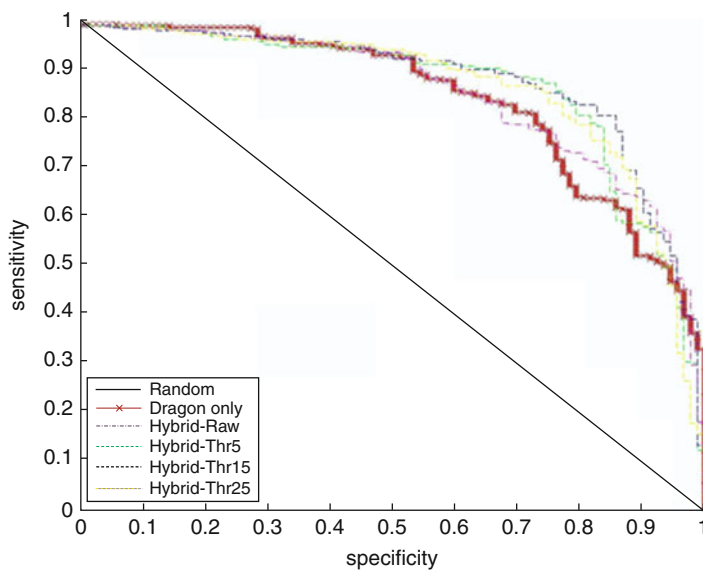


Fig. 9 Acute toxicity modeling. The ROC curves for conventional QSAR model (bold line) and different hybrid models for the same external compounds.

Besides the carcinogenicity, we asked if HTS results could be of value in predicting rodent acute toxicity (Sedykh et al. [in press](#)). For this purpose, we have identified 690 of these compounds, for which rodent acute toxicity data (i.e., toxic or nontoxic) was also available. The classification kNN QSAR modeling method was applied to these compounds using either chemical descriptors alone or as a combination of chemical and qHTS biological (hybrid) descriptors as compound features. The external prediction accuracy of models built with chemical descriptors only was 76 %. In contrast, the prediction accuracy was significantly improved to 85 % when using hybrid descriptors. The receiver operating characteristic (ROC) curves of conventional QSAR models and different hybrid models are shown in Fig. 9. The sensitivities and specificities of hybrid models are clearly better than for conventional QSAR model for predicting the same external compounds. Furthermore, the prediction coverage increased from 76 % when using chemical descriptors only to 93 % when qHTS biological descriptors were also included. Our studies suggest that combining HTS profiles, especially the dose–response qHTS results, with conventional chemical descriptors could considerably improve the predictive power of computational approaches for rodent acute toxicity assessment.

Collaborative and Consensus Modeling of Aquatic Toxicity

We discuss below the results of a recent important study of aquatic toxicity (Zhu et al. 2008). In our opinion, this particular study may serve as a useful example to illustrate the complexity and power of modern QSAR modeling approaches and highlight the importance of collaborative and consensual model development.

The combinational QSAR modeling approach has been applied to a diverse series of organic compounds tested for aquatic toxicity in *Tetrahymena pyriformis* in the same laboratory over nearly a decade (Aptula et al. 2005; Netzeva and Schultz 2005; Schultz 1999; Schultz and Netzeva 2004; Schultz et al. 2001, 2002, 2003, 2005a, 2005b). The unique aspect of this research was that it was conducted in collaboration between six academic groups specializing in cheminformatics and computational toxicology. The common goals for our virtual collaboratory were to explore the relative strengths of various QSAR approaches in their ability to develop robust and externally predictive models of this particular toxicity end point. We have endeavored to develop the most statistically robust, validated, and externally predictive QSAR models of aquatic toxicity. The members of our collaboratory included scientists from the University of North Carolina at Chapel Hill in the United States (UNC); University of Louis Pasteur (ULP) in France; University of Insubria (UI) in Italy; University of Kalmar (UK) in Sweden; Virtual Computational Chemistry Laboratory (VCCLAB) in Germany; and the University of British Columbia (UBC) in Canada. Each group relied on its own QSAR modeling approaches to develop toxicity models using the same modeling set, and we agreed to evaluate the realistic model performance using the same external validation set(s).

The *T. pyriformis* toxicity dataset used in this study was compiled from several publications of the Schultz group as well as from data available at the Tetratox database Web site of (<http://www.vet.utk.edu/TETRATOX/>). After deleting duplicates as well as several compounds with conflicting test results and correcting several chemical structures in the original data sources, our final dataset included 983 unique compounds. The dataset was randomly divided into two parts: (1) the modeling set of 644 compounds; (2) the validation set including 339 compounds. The former set was used for model development by each participating group and the latter set was used to estimate the external prediction power of each model as a universal metric of model performance. In addition, when this project was already well underway, a new dataset had become available from the most recent publication by the Schultz group (Schultz et al. 2007). It provided us with an additional external set to evaluate the predictive power and reliability of all QSAR models. Among compounds reported in (Schultz et al. 2007) 110 were unique, that is, not present among the original set of 983 compounds; thus, these 110 compounds formed the second independent validation set for our study.

Universal Statistical Figures of Merit for All Models

Different groups have employed different techniques and (sometimes) different statistical parameters to evaluate the performance of models developed independently for the modeling set (described below). To harmonize the results of this study, the same standard parameters were chosen to describe each model's performance as applied to the modeling and external test set predictions. Thus, we have employed Q_{abs}^2 (squared leave-one-out cross-validation correlation coefficient) for the modeling set, R_{abs}^2 (frequently described as coefficient of determination) for the external validations sets, and MAE (mean absolute error) for the linear correlation between predicted (Y_{pred}) and experimental (Y_{exp}) data (here, $Y = pIGC_{50}$); these parameters are defined as follows:

$$Q_{abs}^2 = 1 - \frac{\sum_Y (Y_{exp} - Y_{LOO})^2}{\sum_Y (Y_{exp} - \langle Y \rangle_{exp})^2} \quad (8)$$

$$R_{abs}^2 = 1 - \frac{\sum_Y (Y_{exp} - Y_{pred})^2}{\sum_Y (Y_{exp} - \langle Y \rangle_{exp})^2} \quad (9)$$

$$MAE = \frac{\sum_Y |Y - Y_{pred}|}{n} \quad (10)$$

Many other statistical characteristics can be used to evaluate model performance; however, we restricted ourselves to these three parameters that provide minimal but sufficient information concerning any model's ability to reproduce both the trends

in experimental data for the test sets as well as mean accuracy of predicting all experimental values. The models were considered acceptable if R_{abs}^2 exceeded 0.5.

Consensus QSAR Models of Aquatic Toxicity; Comparison Between Methods and Models

The objective of this study from methodological prospective was to explore the suitability of different QSAR modeling tools for the analysis of a dataset with an important toxicological end point. Typically, such datasets are analyzed with one (or several) modeling techniques, with a great emphasis on the (high value of) statistical parameters of the training set models. In this study, we went well beyond the modeling studies reported in the original publications in several respects. First, we have compiled all reported data on chemical toxicity against *T. pyriformis* in a single large dataset and attempted to develop global QSAR models for the entire set. Second, we have employed multiple QSAR modeling techniques thanks to the engagement of six collaborating groups. Third, we have focused on defining model performance criteria not only using training set data but most importantly using external validation sets that were not used in model development in any way (unlike any common cross-validation procedure) (Gramatica 2007). This focus afforded us the opportunity to evaluate and compare all models using simple and objective universal criteria of external predictive accuracy, which in our opinion is the most important single figure of merit for a QSAR model that is of practical significance for experimental toxicologists. Fourth, we have explored the significance of applicability domains and the power of consensus modeling in maximizing the accuracy of external predictivity of our models.

We believe that results of our analysis lend a strong support for our strategy. Indeed, all models performed quite well for the training set with even the lowest Q_{abs}^2 among them as high as 0.72. However, there was much greater variation between these models when looking at their (universal and objective) performance criteria as applied to the validation sets.

Of 15 QSAR approaches used in this study, nine implemented method-specific applicability domains. Models that did not define the AD showed a reduced predictive accuracy for the validation set II even though they yielded reasonable results for the validation set I. On average, the use of applicability domains improved the performance of individual models although the improvement came at the expense of the lower chemistry space coverage.

For the most part all models succeeded in achieving reasonable accuracy of external prediction especially when using the AD. It then appeared natural to bring all models together to explore the power of consensus prediction. Thus, the consensus model was constructed by averaging all available predicted values taking into account the applicability domain of each individual model. In this case, we could use only 9 of 15 models that had the AD defined. Since each model had its unique way of defining the AD, each external compound could be found within

the AD of anywhere between one and nine models so for averaging we only used models covering the compound. The advantage of this data treatment is that the overall coverage of the prediction is still high because it was rare to have an external compound outside of the ADs of all available models. The results showed that the prediction accuracy for both the modeling set and the validation sets was the best compared to any individual model. The same observation could be made for the correlation coefficient R_{abs}^2 . The coverage of this consensus model II was 100 % for all three data sets. This observation suggests that consensus models afford both high space coverage and high accuracy of prediction

In summary, this study presents an example of a fruitful international collaboration between researchers that use different techniques and approaches but share general principles of QSAR model development and validation. Significantly, we did not make any assumptions about the purported mechanisms of aquatic toxicity yet were able to develop statistically significant models for all experimentally tested compounds. In this regard it is relevant to cite an opinion expressed in an earlier publication by T. Schultz that “models that accurately predict acute toxicity without first identifying toxic mechanisms are highly desirable” (Schultz 1999). However, the most significant single result of our studies is the demonstrated superior performance of the consensus modeling approach when all models are used concurrently and predictions from individual models are averaged. We have shown that both the predictive accuracy and coverage of the final consensus QSAR models were superior as compared to these parameters for individual models. The consensus models appeared robust in terms of being insensitive to both incorporating individual models with low prediction accuracy and the inclusion or exclusion of the AD. Another important result of this study is the power of addressing complex problems in QSAR modeling by forming a virtual collaboratory of independent research groups leading to the formulation and empirical testing of best modeling practices. This latter endeavor is especially critical in light of the growing interest of regulatory agencies to developing most reliable and predictive models for environmental risk assessment (Yang et al. 2006) and placing such models in the public domain.

Conclusions: Emerging Chemical/Biological Data and QSAR Research Strategies

In the past 15 years, innovative technologies that enable rapid synthesis and high throughput screening of large libraries of compounds have been adopted in almost all major pharmaceutical and biotech companies. As a result, there has been a huge increase in the number of compounds available on a routine basis to quickly screen for novel drug candidates against new targets or pathways. In contrast, such technologies have rarely become available to the academic research community, thus limiting its ability to conduct large-scale chemical genetics or chemical genomics research. The NIH Molecular Libraries Roadmap Initiative has changed this situation by forming the national Molecular Library Screening

Centers Network (MLSCN) (Austin et al. 2004) with the results of screening assays made publicly available via PubChem (2010). These efforts have already led to the unprecedented growth of available databases of biologically tested compounds [cf. our recent review where we list about 20 available databases of compounds with known bioactivity (Oprea and Tropsha 2006)].

This growth creates new challenges for QSAR modeling such as developing novel approaches for the analysis and visualization of large databases of screening data, novel biologically relevant chemical diversity or similarity measures, and novel tools for virtual screening of compound libraries to ensure high expected hit rates. Application studies discussed in this chapter have established that QSAR models could be used successfully as virtual screening tools to discover compounds with the desired biological activity in chemical databases or virtual libraries (Hsieh et al. 2008; Oloff et al. 2005; Shen et al. 2004; Tropsha 2005; Tropsha and Zheng 2001; Zhang et al. 2007). The discovery of novel bioactive chemical entities is the primary goal of computational drug discovery, and the development of validated and predictive QSAR models is critical to achieve this goal. Due to the significant recent increase in publicly available datasets of biologically active compounds and the critical need to improve the hit rate of experimental compound screening there is a strong need in developing widely accessible and reliable computational QSAR modeling techniques and specific end-point predictors.

Acknowledgments The studies described in this chapter were supported in parts by the NIH research grants R01GM066940 and R21GM076059 and EPA grants EPA (RD832720 and RD833825).

Bibliography

- Afantitis, A., Melagraki, G., Sarimveis, H., Koutentis, P. A., Markopoulos, J., & Igglessi-Markopoulou, O. (2006). A novel QSAR model for predicting induction of apoptosis by 4-aryl-4H-chromenes. *Bioorganic & Medicinal Chemistry*, 14, 6686.
- Agrafiotis, D. K., Cedeno, W., & Lobanov, V. S. (2002). On the use of neural network ensembles in QSAR and QSPR. *Journal of Chemical Information and Computer Science*, 42, 903.
- Ajmani, S., Jadhav, K., & Kulkarni, S. A. (2006). Three-dimensional QSAR using the k-nearest neighbor method and its interpretation. *Journal of Chemical Information and Modeling*, 46, 24.
- Aptula, A. O., Roberts, D. W., Cronin, M. T. D., & Schultz, T. W. (2005). Chemistry-toxicity relationships for the effects of Di- and trihydroxybenzenes to *Tetrahymena pyriformis*. *Chemical Research in Toxicology*, 18, 844.
- Austin, C. P., Brady, L. S., Insel, T. R., & Collins, F. S. (2004). NIH molecular libraries initiative. *Science*, 306, 1138.
- Bailey, A. B., Chanderbhan, R., Collazo-Braier, N., Cheeseman, M. A., & Twaroski, M. L. (2005). The use of structure-activity relationship analysis in the food contact notification program. *Regulatory Toxicology and Pharmacology*, 42, 225.
- Bajorath, J., Peltason, L., Wawer, M., Guha, R., Lajiness, M. S., & Van Drie, J. H. (2009). Navigating structure-activity landscapes. *Drug Discovery Today*, 14, 698.
- Berk, R. A. (2008). *Classification and Regression Trees (CART). Statistical learning from a regression perspective*. New York: Springer.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24, 123.
- Breiman, L. (1998). Arcing classifiers. *The Annals of Statistics*, 26, 801.

- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5.
- Bures, M. G., & Martin, Y. C. (1998). Computational methods in molecular diversity and combinatorial chemistry. *Current Opinion in Chemical Biology*, 2, 376.
- Burges, J. C. (1998). Tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 121.
- C5.0.(2008).
- Carhart, R. E., Smith, D. H., & Venkataraghavan, R. (1985). Atom pairs as molecular features in structure-activity studies: Definition and applications. *Journal of Chemical Information and Computer Science*, 25, 64.
- Casey, P. J., & Seabra, M. C. (1996). Protein prenyltransferases. *The Journal of Biological Chemistry*, 271, 5289.
- Cheeseman, M. A. (2005). Thresholds as a unifying theme in regulatory toxicology. *Food Additives & Contaminants*, 22, 900.
- ChemAxon. (2008). <http://www.chemaxon.com>.
- ChEMBL Database. (2010). <http://www.ebi.ac.uk/chembl/db/>.
- Chen, C., Liaw, A., & Breiman, L. (2004). *Using random forest to learn imbalanced data* (Vol. 666). Berkeley: Department of Statistics, University of California.
- Cherkasov, A. (2008). An updated steroid benchmark set and its application in the discovery of novel nanomolar ligands of sex hormone-binding globulin. *Journal of Medicinal Chemistry*, 51, 2047.
- Cho, S. J., Zheng, W., & Tropsha, A. (1998). Rational combinatorial library design. 2. Rational design of targeted combinatorial peptide libraries using chemical similarity probe and the inverse QSAR approaches. *Journal of Chemical Information and Computer Science*, 38, 259.
- Cox, A. D., & Der, C. J. (2002). Farnesyltransferase inhibitors: Promises and realities. *Current Opinion in Pharmacology*, 2, 388.
- Crivori, P., Cruciani, G., Carrupt, P. A., & Testa, B. (2000). Predicting blood-brain barrier permeation from three-dimensional molecular structure. *Journal of Medicinal Chemistry*, 43, 2204.
- Cruciani, G., Pastor, M., & Guba, W. (2000). VolSurf: A new tool for the pharmacokinetic optimization of lead compounds1. *European Journal of Pharmaceutical Sciences*, 11(Suppl 2), S29–S39.
- Dash, A., Inman, W., Hoffmaster, K., Sevidal, S., Kelly, J., Obach, R. S., et al. (2009). Liver tissue engineering in the evaluation of drug safety. *Expert Opinion on Drug Metabolism & Toxicology*, 5, 1159.
- de Cerqueira, L. P., Golbraikh, A., Oloff, S., Xiao, Y., & Tropsha, A. (2006). Combinatorial QSAR modeling of P-Glycoprotein substrates. *Journal of Chemical Information and Modeling*, 46, 1245.
- Discovery Studio. (2010).
- Dix, D. J., Houck, K. A., Martin, M. T., Richard, A. M., Setzer, R. W., & Kavlock, R. J. (2007). The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicological Sciences*, 95, 5.
- Dragon. (2007). http://www.talete.mi.it/help/dragon_help/index.html?IntroducingDRAGON
- DSSTox. (2008). <http://www.epa.gov/nheerl/dsstox/About.html>.
- Durham, S. K., & Pearl, G. M. (2001). Computational methods to predict drug safety liabilities. *Current Opinion in Drug Discovery & Development*, 4, 110.
- Environmental Protection Agency. (1992). *Statistical training course for ground-water monitoring data analysis EPA/530-R-93-003*. Washington: Office of Solid Waste.
- Fallon, A., Spada, C., & Gallagher, D. (1997). Detection and Accommodation of Outliers in Normally Distributed Data Sets. <http://ewr.cee.vt.edu/environmental/teach/smprimer/outlier/outlier.html>. Accessed 25 April 2005.
- Fechner, N., Hinselmann, G., Schmiedl, C., & Zell, A. (2008). Estimating the applicability domain of kernel-based QSPR models using classical descriptor vectors. *Chemistry Central Journal*, 2(Suppl.1), P2.

- Forsby, A., & Blaauboer, B. (2007). Integration of in vitro neurotoxicity data with biokinetic modelling for the estimation of in vivo neurotoxicity. *Human & Experimental Toxicology*, 26, 333.
- Fourches, D., Muratov, E., & Tropsha, A. (2010). Trust, but verify: On the importance of chemical structure curation in cheminformatics and QSAR modeling research. *Journal of Chemical Information and Modeling*, 50, 1189–1204.
- Gasteiger, J. (2006). Of molecules and humans. *Journal of Medicinal Chemistry*, 49, 6429.
- Golbraikh, A., & Tropsha, A. (2002). Predictive QSAR modeling based on diversity sampling of experimental datasets for the training and test set selection. *Journal of Computer-Aided Molecular Design*, 16, 357.
- Golbraikh, A., & Tropsha, A. (2003). QSAR modeling using chirality descriptors derived from molecular topology. *Journal of Chemical Information and Computer Science*, 43, 144.
- Golbraikh, A., Bonchev, D., & Tropsha, A. (2001). Novel chirality descriptors derived from molecular topology. *Journal of Chemical Information and Computer Science*, 41, 147.
- Golbraikh, A., Bonchev, D., & Tropsha, A. (2002). Novel ZE-isomerism descriptors derived from molecular topology and their application to QSAR analysis. *Journal of Chemical Information and Computer Science*, 42, 769.
- Golbraikh, A., Shen, M., Xiao, Z., Xiao, Y. D., Lee, K. H., & Tropsha, A. (2003). Rational selection of training and test sets for the development of validated QSAR models. *Journal of Computer-Aided Molecular Design*, 17, 241.
- Gramatica, P. (2007). Principles of QSAR models validation: Internal and external. *QSAR & Combinatorial Science*, 26, 694.
- Guha, R., & Van Drie, J. H. (2008a). Structure–activity landscape index: Identifying and quantifying activity cliffs. *Journal of Chemical Information and Modeling*, 48, 646.
- Guha, R., & Van Drie, J. H. (2008b). Assessing how well a modeling protocol captures a structure-activity landscape. *Journal of Chemical Information and Modeling*, 48, 1716.
- Hoffman, B., Cho, S. J., Zheng, W., Wyrick, S., Nichols, D. E., Mailman, R. B., et al. (1999). Quantitative structure-activity relationship modeling of dopamine D(1) antagonists using comparative molecular field analysis, genetic algorithms-partial least-squares, and K nearest neighbor methods. *Journal of Medicinal Chemistry*, 42, 3217.
- Hong, H., Xie, Q., Ge, W., Qian, F., Fang, H., Shi, L., et al. (2008). Mold(2), molecular descriptors from 2D structures for cheminformatics and toxicoinformatics. *Journal of Chemical Information and Modeling*, 48, 1337.
- Horvath, D., Bonachera, F., Solov'ev, V., Gaudin, C., & Varnek, A. (2007). Stochastic versus stepwise strategies for quantitative structure-activity relationship generation—how much effort may the mining for successful QSAR models take? *Journal of Chemical Information and Modeling*, 47, 927.
- Hsieh, J. H., Wang, X. S., Teotico, D., Golbraikh, A., & Tropsha, A. (2008). Differentiation of AmpC beta-lactamase binders vs. decoys using classification kNN QSAR modeling and application of the QSAR classifier to virtual screening. *Journal of Computer-Aided Molecular Design*, 22, 593.
- Huan, J., Bandyopadhyay, D., Prins, J., Snoeyink, J., Tropsha, A., & Wang, W. (2006). Distance-based identification of structure motifs in proteins using constrained frequent subgraph mining. *Computational Systems Bioinformatics Conference*, 227.
- Inglese, J., Auld, D. S., Jadhav, A., Johnson, R. L., Simeonov, A., Yasgar, A., et al. (2006). Quantitative high-throughput screening: A titration-based approach that efficiently identifies biological activities in large chemical libraries. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 11473.
- Irwin, J. J., & Shoichet, B. K. (2005). ZINC—a free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling*, 45, 177.
- Jacobson-Kram, D., & Contrera, J. F. (2007). Genetic toxicity assessment: Employing the best science for human safety evaluation. Part I: Early screening for potential human mutagens. *Toxicological Sciences*, 96, 16.

- Japkowicz, N. (2000). *Learning from imbalanced datasets: A comparison of various strategies*. AAAI Workshop. Menlo Park: AAAI Press.
- Jaworska, J., & Nikolova-Jeliazkova, N. (2008). *Review of methods to assess a QSAR Applicability Domain*. http://ambit.acad.bg/nina/publications/2004/AppDomain_sar04.ppt
- Jaworska, J., Nikolova-Jeliazkova, N., & Aldenberg, T. (2005). QSAR applicability domain estimation by projection of the training set descriptor space: A review. *Alternatives to Laboratory Animals*, 33, 445.
- Johnson, S. R. (2008). The trouble with QSAR (or how I learned to stop worrying and embrace fallacy). *Journal of Chemical Information and Modeling*, 48, 25.
- Klebe, G. (1998). Comparative molecular similarity indices: CoMSI. In H. Kubinyi, G. Folkers, & Y. Martin (Eds.), *3D QSAR in drug design* (pp. 87–104). Great Britain: Kluwer.
- Klopman, G., Zhu, H., Ecker, G., & Chiba, P. (2003). MCASE study of the multidrug resistance reversal activity of propafenone analogs. *Journal of Computer-Aided Molecular Design*, 17, 291.
- Kola, I., & Landis, J. (2004). Can the pharmaceutical industry reduce attrition rates? *Nature Reviews Drug Discovery*, 3, 711.
- Kovatcheva, A., Golbraikh, A., Oloff, S., Feng, J., Zheng, W., & Tropsha, A. (2005). QSAR modeling of datasets with enantioselective compounds using chirality sensitive molecular descriptors. *SAR and QSAR in Environmental Research*, 16, 93.
- Kubat, M., & Matwin, S. (1997). *Addressing the curse of imbalanced training sets: One sided selection*. San Francisco: Morgan Kaufmann.
- Kubinyi, H., Hamprecht, F. A., & Mietzner, T. (1998). Three-dimensional quantitative similarity-activity relationships (3D QSiAR) from SEAL similarity matrices. *Journal of Medicinal Chemistry*, 41, 2553.
- LigandScout. (2010).
- Liu, T., Lin, Y., Wen, X., Jorissen, R. N., & Gilson, M. K. (2007). BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Research*, 35, D198–D201.
- Maggiora, G. M. (2006). On outliers and activity cliffs—why QSAR often disappoints. *Journal of Medicinal Chemistry*, 46, 1535.
- Maybridge. (2005). <http://www.daylight.com/products/databases/Maybridge.html>
- Mayer, P., & Reichenberg, F. (2006). Can highly hydrophobic organic substances cause aquatic baseline toxicity and can they contribute to mixture toxicity? *Environmental Toxicology & Chemistry*, 25, 2639.
- McGregor, M. J., & Pallai, P. V. (1997). Clustering of large databases of compounds: Using the MDL “Keys” as structural descriptors. *Journal of Chemical Information and Computer Science*, 37, 443.
- MDDR. SYMYX technologies. (2009). http://www.mdl.com/products/knowledge/drug_data_report/index.jsp
- Medina-Franco, J. L., Golbraikh, A., Oloff, S., Castillo, R., & Tropsha, A. (2005). Quantitative structure-activity relationship analysis of pyridinone HIV-1 reverse transcriptase inhibitors using the k nearest neighbor method and QSAR-based database mining. *Journal of Computer-Aided Molecular Design*, 19, 229.
- Molconn-Z. (2007). <http://www.edusoft-lc.com/>
- Molecular Operating Environment (MOE). (2008). <http://www.chemcomp.com/>
- Muster, W., Breidenbach, A., Fischer, H., Kirchner, S., Muller, L., & Pahler, A. (2008). Computational toxicology in drug development. *Drug Discovery Today*, 13, 303.
- nci. (2007). http://dtp.nci.nih.gov/docs/3d_database/structural_information/smiles_strings.html.
- Netzeva, T. I., & Schultz, T. W. (2005). QSARs for the aquatic toxicity of aromatic aldehydes from Tetrahymena data. *Chemosphere*, 61, 1632.
- Netzeva, T. I., Gallegos, S. A., & Worth, A. P. (2006). Comparison of the applicability domain of a quantitative structure-activity relationship for estrogenicity with a large chemical inventory. *Environmental Toxicology & Chemistry*, 25, 1223.
- Neural Networks. (1996). *Neural networks in QSAR and drug design*. San Diego: Academic.

- Neural Networks. (2010). <http://www.learnartificialneuralnetworks.com/>.
- Nikolova-Jeliazkova, N., & Jaworska, J. (2005). An approach to determining applicability domains for QSAR group contribution models: An analysis of SRC KOWWIN. *Alternatives to Laboratory Animals*, 33, 461.
- Olah, M., Rad, R., Ostopovici, L., Bora, A., Hadaruga, N., Hadaruga, D., et al. (2007). WOMBAT and WOMBAT-PK: Bioactivity databases for lead and drug discovery. In S. L. Schreiber, T. M. Kapoor, & G. Weiss (Eds.), *Chemical biology: From small molecules to systems biology and drug design* (pp. 760–786). Weinheim: Wiley-VCH.
- Oloff, S., Mailman, R. B., & Tropsha, A. (2005). Application of validated QSAR models of D1 dopaminergic antagonists for database mining. *Journal of Medicinal Chemistry*, 48, 7322.
- (2010). OpenBabel: The OpenSource Chemistry Toolbox. [Openbabel.org](http://openbabel.org). 2-1-2010.
- Oprea, T., & Tropsha, A. (2006). Target, chemical and bioactivity databases – integration is key. *Drug Discovery Today*, 3, 357–365.
- Organisation for Economic and Co-operation Development. (2008). OECD Quantitative Structure-Activity Relationships [(Q)SARs] Project. http://www.oecd.org/document/23/0,3343,en_2649_34365_33957015_1_1_1_1,00.html.
- Park, M. V., Lankveld, D. P., Van, L. H., & de Jong, W. H. (2009). The status of in vitro toxicity studies in the risk assessment of nanomaterials. *Nanomedicine (London, England)*, 4, 669.
- Pastor, M., Cruciani, G., McLay, I., Pickett, S., & Clementi, S. (2000). GRIND-INdependent descriptors (GRIND): A novel class of alignment-independent three-dimensional molecular descriptors. *Journal of Medicinal Chemistry*, 43, 3233.
- PDSP. (2010). PDSP. <http://pdsp.med.unc.edu>.
- Peterson, Y. K., Kelly, P., Weinbaum, C. A., & Casey, P. J. (2006). A novel protein geranylgeranyltransferase-I inhibitor with high potency, selectivity, and cellular activity. *The Journal of Biological Chemistry*, 281, 12445.
- Peterson, Y. K., Wang, X. S., Casey, P. J., & Tropsha, A. (2009). Discovery of geranylgeranyltransferase-I inhibitors with novel scaffolds by the means of quantitative structure-activity relationship modeling, virtual screening, and experimental validation. *Journal of Medicinal Chemistry*, 52, 4210.
- Piersma, A. H., Janer, G., Wolterink, G., Bessems, J. G., Hakkert, B. C., & Slob, W. (2008). Quantitative extrapolation of in vitro whole embryo culture embryotoxicity data to developmental toxicity in vivo using the benchmark dose approach. *Toxicological Sciences*, 101, 91.
- PubChem. (2010). <http://pubchem.ncbi.nlm.nih.gov/>.
- Quinlan, J. R. (1993). *C4.5: Programs for machine learning*. San Mateo: Morgan Kaufmann.
- Random Forests. (2001).
- Riley, R. J., & Kenna, J. G. (2004). Cellular models for ADMET predictions and evaluation of drug-drug interactions. *Current Opinion in Drug Discovery & Development*, 7, 86.
- Robinson, D. D., Winn, P. J., Lyne, P. D., & Richards, W. G. (1999). Self-organizing molecular field analysis: A tool for structure-activity studies. *Journal of Medicinal Chemistry*, 42, 573.
- Saliner, A. G., Netzeva, T. I., & Worth, A. P. (2006). Prediction of estrogenicity: Validation of a classification model. *SAR and QSAR in Environmental Research*, 17, 195.
- Salt, D. V., Yildiz, N., Livingstone, D. J., & Tinsley, C. J. (2006). The use of artificial neural networks in QSAR. *Pesticide Science*, 36, 161.
- Schirmer, K., Tanneberger, K., Kramer, N. I., Volker, D., Scholz, S., Hafner, C., et al. (2008). Developing a list of reference chemicals for testing alternatives to whole fish toxicity tests. *Aquatic Toxicology*, 90, 128.
- Schrodinger Software. (2010).
- Schultz, T. W. (1999). Structure-toxicity relationships for benzenes evaluated with *Tetrahymena pyriformis*. *Chemical Research in Toxicology*, 12, 1262.
- Schultz, T. W., & Netzeva, T. I. (2004). Development and evaluation of QSARs for ecotoxic endpoints: The benzene response-surface model for *Tetrahymena* toxicity. In M. T. D. Cronin & D. J. Livingstone (Eds.), *Modeling environmental fate and toxicity* (pp. 265–284). Boca Raton: CRC Press.

- Schultz, T. W., Sinks, G. D., & Miller, L. A. (2001). Population growth impairment of sulfur-containing compounds to *Tetrahymena pyriformis*. *Environmental Toxicology*, *16*, 543.
- Schultz, T. W., Cronin, M. T., Netzeva, T. I., & Aptula, A. O. (2002). Structure-toxicity relationships for aliphatic chemicals evaluated with *Tetrahymena pyriformis*. *Chemical Research in Toxicology*, *15*, 1602.
- Schultz, T. W., Netzeva, T. I., & Cronin, M. T. (2003). Selection of data sets for QSARs: Analyses of *Tetrahymena* toxicity from aromatic compounds. *SAR and QSAR in Environmental Research*, *14*, 59.
- Schultz, T. W., Netzeva, T. I., Roberts, D. W., & Cronin, M. T. (2005a). Structure-toxicity relationships for the effects to *Tetrahymena pyriformis* of aliphatic, carbonyl-containing, alpha, beta-unsaturated chemicals. *Chemical Research in Toxicology*, *18*, 330.
- Schultz, T. W., Yarbrough, J. W., & Woldemeskel, M. (2005b). Toxicity to *Tetrahymena* and abiotic thiol reactivity of aromatic isothiocyanates. *Cell Biology and Toxicology*, *21*, 181.
- Schultz, T. W., Hewitt, M., Netzeva, T. I., & Cronin, M. T. D. (2007). Assessing applicability domains of toxicological QSARs: Definition, confidence in predicted values, and the role of mechanisms of action. *QSAR & Combinatorial Science*, *26*, 238.
- Sebti, S. M., & Hamilton, A. D. (2000). Farnesyltransferase and geranylgeranyltransferase I inhibitors in cancer therapy: Important mechanistic and bench to bedside issues. *Expert Opinion on Investigational Drugs*, *9*, 2767.
- Sedykh, A., Zhu, H., Tang, H., Zhang, L., Rusyn, I., Richard, A., et al. The use of dose-response qHTS data as biological descriptors improves the prediction accuracy of QSAR models of acute rat toxicity. Environmental Health Perspect, In press.
- Shen, M., LeTiran, A., Xiao, Y., Golbraikh, A., Kohn, H., & Tropsha, A. (2002). Quantitative structure-activity relationship analysis of functionalized amino acid anticonvulsant agents using k nearest neighbor and simulated annealing PLS methods. *Journal of Medicinal Chemistry*, *45*, 2811.
- Shen, M., Beguin, C., Golbraikh, A., Stables, J. P., Kohn, H., & Tropsha, A. (2004). Application of predictive QSAR models to database mining: Identification and experimental validation of novel anticonvulsant compounds. *Journal of Medicinal Chemistry*, *47*, 2356.
- Sisay, M. T., Peltason, L., & Bajorath, J. (2009). Structural interpretation of activity cliffs revealed by systematic analysis of structure-activity relationships in analog series. *Journal of Chemical Information and Modeling*, *49*, 2179.
- Sjostrom, M., Kolman, A., Clemedson, C., & Clothier, R. (2008). Estimation of human blood LC50 values for use in modeling of in vitro-in vivo data of the ACuteTox project. *Toxicology In Vitro*, *22*, 1405.
- Smola, A. J., & Schoelkopf, B. A. (2004). Tutorial on support vector regression. Tuebingen: Max Planck Society - eDocument Server (Germany).
- Snyder, R. D. (2009). An update on the genotoxicity and carcinogenicity of marketed pharmaceuticals with reference to in silico predictivity. *Environmental and Molecular Mutagenesis*, *50*, 435.
- Stoner, C. L., Gifford, E., Stankovic, C., Lepsy, C. S., Brodfuehrer, J., Prasad, J. V. N. V., et al. (2004). Implementation of an ADME enabling selection and visualization tool for drug discovery. *Journal of Pharmaceutical Sciences*, *93*, 1131.
- Stouch, T. R., Kenyon, J. R., Johnson, S. R., Chen, X. Q., Doweiko, A., & Li, Y. (2003). In silico ADME/Tox: why models fail. *Journal of Computer-Aided Molecular Design*, *17*, 83.
- Tetko, I. V., Bruneau, P., Mewes, H. W., Rohrer, D. C., & Poda, G. I. (2006). Can we estimate the accuracy of ADME-Tox predictions? *Drug Discovery Today*, *11*, 700.
- The Foundations of Cost-sensitive Learning. (2001).
- Thomas, C. J., Auld, D. S., Huang, R., Huang, W., Jadhav, A., Johnson, R. L., et al. (2009). The pilot phase of the NIH chemical genomics center. *Current Topics in Medicinal Chemistry*, *9*, 1181.
- Todeschini, R., & Consonni, V. (2000). *Handbook of molecular descriptors*. Weinheim: Wiley-VCH.
- Tripos. (2010). Sybyl-X 1.0

- Tropsha, A. (2005). Application of predictive QSAR models to database mining. In T. Oprea (Ed.), *Cheminformatics in drug discovery* (pp. 437–455). Wiley-VCH.
- Tropsha, A., & Golbraikh, A. (2007). Predictive QSAR modeling workflow, model applicability domains, and virtual screening. *Current Pharmaceutical Design, 13*, 3494.
- Tropsha, A., & Golbraikh, A. (2010a). Predictive quantitative structure–activity relationships modeling: Development and validation of QSAR models. In J.-L. Faulon & A. Bender (Eds.), *Handbook of chemoinformatics algorithms*. The Netherlands: Leiden University, Chapman and Hall/CRC.
- Tropsha, A., & Golbraikh, A. (2010b). Predictive quantitative structure–activity relationships modeling. Data Preparation and the General Modeling Workflow. In J.-L. Faulon & A. Bender (Eds.), *Handbook of chemoinformatics algorithms* (pp. 175–214). The Netherlands: Leiden University, Chapman and Hall/CRC.
- Tropsha, A., & Zheng, W. (2001). Identification of the descriptor pharmacophores using variable selection QSAR: Applications to database mining. *Current Pharmaceutical Design, 7*, 599.
- Valerio, L., Jr. (2008). Tools for evidence-based toxicology: Computational-based strategies as a viable modality for decision support in chemical safety evaluation and risk assessment. *Human & Experimental Toxicology, 27*, 757.
- Valerio, L. G., Jr. (2009). In silico toxicology for the pharmaceutical sciences. *Toxicology and Applied Pharmacology, 241*, 356.
- Vapnik, V. (2000). *Nature of statistical learning theory*. New York: Springer.
- Votano, J. R., Parham, M., Hall, L. H., Kier, L. B., Oloff, S., Tropsha, A., et al. (2004). Three new consensus QSAR models for the prediction of Ames genotoxicity. *Mutagenesis, 19*, 365.
- Waller, C. L. (2004). A comparative QSAR study using CoMFA, HQSAR, and FRED/SKEYS paradigms for estrogen receptor binding affinities of structurally diverse compounds. *Journal of Chemical Information and Computer Science, 44*, 758.
- Winter-Vann, A. M., & Casey, P. J. (2005). Post-prenylation-processing enzymes as new targets in oncogenesis. *Nature Reviews Cancer, 5*, 405.
- Xia, M., Huang, R., Witt, K. L., Southall, N., Fostel, J., Cho, M. H., et al. (2008). Compound cytotoxicity profiling using quantitative high-throughput screening. *Environmental Health Perspectives, 116*, 284.
- Yang, C., Richard, A. M., & Cross, K. P. (2006). The art of data mining the minefields of toxicity databases to link chemistry to biology. *Current Computer-Aided Drug Design, 2*, 135.
- Yang, C., Valerio, L. G., Jr., & Arvidson, K. B. (2009). Computational toxicology approaches at the US food and drug administration. *Alternatives to Laboratory Animals, 37*, 523.
- Yen, S.-J., & Lee, Y.-S. (2006). Under-sampling approaches for improving prediction of the minority class in an imbalanced dataset. *Lecture Notes in Control and Information Sciences, 344*, 731.
- Zhang, F. L., & Casey, P. J. (1996). Protein prenylation: Molecular mechanisms and functional consequences. *Annual Review of Biochemistry, 65*, 241.
- Zhang, S., Wei, L., Bastow, K., Zheng, W., Brossi, A., Lee, K. H., et al. (2007). Antitumor agents 252. Application of validated QSAR models to database mining: Discovery of novel tylophorine derivatives as potential anticancer agents. *Journal of Computer-Aided Molecular Design, 21*, 97.
- Zhang, L., Zhu, H., Oprea, T. I., Golbraikh, A., & Tropsha, A. (2008). QSAR modeling of the blood–brain barrier permeability for diverse organic compounds. *Pharmaceutical Research, 25*, 1902.
- Zheng, W., & Tropsha, A. (2000). Novel variable selection quantitative structure–property relationship approach based on the k-nearest-neighbor principle. *Journal of Chemical Information and Computer Science, 40*, 185.
- Zhou, Z. H., & Liu, X.-Y. (2006). Training cost-sensitive neural networks with methods addressing the class imbalance problem. *IEEE Transactions on Knowledge and Data Engineering, 18*, 63.
- Zhu, H., Rusyn, I., Richard, A. M., & Tropsha, A. (2008a). Use of cell viability assay data improves the prediction accuracy of conventional quantitative structure–activity relationship models of animal carcinogenicity. *Environmental Health Perspectives, 116*, 506.

- Zhu, H., Tropsha, A., Fourches, D., Varnek, A., Papa, E., Gramatica, P., et al. (2008b). Combinatorial QSAR modeling of chemical toxicants tested against *tetrahymena pyriformis*. *Journal of Chemical Information and Modeling*, *48*, 766.
- Zhu, H., Ye, L., Golbraikh, A., & Tropsha, A. (2009). QSAR studies of chemical aquatic acute toxicity using k Nearest Neighbor (kNN) Methodology
- Zhu, H., Ye, L., Richard, A., Golbraikh, A., Wright, F. A., Rusyn, I., et al. (2009a). A novel two-step hierarchical quantitative structure-activity relationship modeling work flow for predicting acute toxicity of chemicals in rodents. *Environmental Health Perspectives*, *117*, 1257.
- Zupan, J., & Gasteiger, J. (1999). *Neural networks in chemistry and drug design*. Weinheim: Wiley-VCH.
- Zvinavashe, E., Murk, A. J., & Rietjens, I. M. (2008). Promises and pitfalls of quantitative structure-activity relationship approaches for predicting metabolism and toxicity. *Chemical Research in Toxicology*, *18*, 844.
- Zvinavashe, E., Murk, A. J., & Rietjens, I. M. (2009). On the number of EINECS compounds that can be covered by (Q)SAR models for acute toxicity. *Toxicology Letters*, *184*, 67.

F.P. Maguna, N.B. Okulik, and Eduardo A. Castro

Contents

Introduction	2342
General Overview of Methods Used in QSAR of Antimicrobial Compounds	2343
Antimicrobial Activity Determination	2343
Quantitative Structure–Activity Relationship (QSAR)	2343
Generation of Molecular Descriptors from Structure	2344
Selection of Relevant Molecular Descriptors	2344
Mapping the Descriptors to Activity	2345
Descriptors Used in Antimicrobial Activity Studies	2345
QSAR Studies of Antimicrobial Compounds	2346
Coumarins	2346
Benzamides	2347
Cinnamic Acid	2347
Flavanones	2348
Phenolic Compounds	2348
Furan Derivatives	2349
mt-QSAR Studies	2350
Conclusion	2351
Bibliography	2352

Abstract

A thorough antimicrobial review of an increasing number of reports reveals a broad spectrum of research activity in the development practices that are used to treat a variety of diseases. The quantitative relationship between chemical struc-

F.P. Maguna (✉) • N.B. Okulik
Facultad de Agroindustrias, Universidad Nacional del Nordeste, Chaco, Argentina

E.A. Castro
INIFTA, Divisin Química Terica, Departamento de Química, Facultad de Ciencias Exactas,
La Plata, Buenos Aires, Argentina
e-mail: castro@quimica.unlp.edu.ar; eacast@gmail.com

ture and biological activity has received considerable attention in recent years because it allows one to predict theoretically bioactivity without an inordinate amount of experimental time and effort. In this chapter we collect and discuss critically published results concerning the QSAR research on antimicrobial compounds. Finally, we present an updated perspective about the future trends in this area.

Introduction

Antimicrobial drugs are drugs designed to kill, or prevent the growth of microorganisms (bacteria, fungi, and viruses). The development of antimicrobial agents for clinical use has brought unquestionable benefit to individuals and society. Infectious diseases that were formerly often fatal became curable (Lerner 1998). However, antibiotics ultimately have lost their original effectiveness as they have been used over time and resistant strains of bacteria have been developed and there is thus an urgent need to identify novel, active chemotypes as leads for better drug development (Cragg et al. 1997; Hall 2004; Mcdermott et al. 2003).

The versatility of bacterial populations to adapt to environmental toxicity and associated facilities for transferring genetic material show that the antibiotic resistance a biological phenomenon is inevitable and will continue to be a chronic medical condition. The appropriate employment of antimicrobials in use and the continued development of new ones are vital to protect the health of men and animals from pathogens (Mcdermott et al. 2003).

Natural products could play a crucial role in meeting this demand (Cragg and Newman 2001). The use of natural products for the treatment of human diseases is a practice that has been used for a long time. It is estimated that 40 % of medicines available for the current treatment of diseases have been developed from natural sources (Cragg and Newman 2001; Shu 1998). The knowledge generated as a result of the use of natural product derivatives of the higher plants, microorganisms, and toxins animals was essential in the discovery of new drugs for modern medical treatment.

In medicine, natural products provide high number of drugs useful in case of complex chemical synthesis and can also serve as basic compound models of synthetic drugs or suitable modifications that reduce its toxicity (Demain 1999).

The quantitative relationship between chemical structure and biological activity has received considerable attention in recent years because it allows one to predict theoretically bioactivity without an inordinate amount of experimental time and demanding efforts.

Computational methods for predicting compounds with specific pharmacodynamic, pharmacokinetic, or toxicological properties are useful for facilitating drug discovery and drug safety evaluation. The quantitative structure-activity relationship (QSAR) and quantitative structure–property relationship (QSPR) methods are the most successfully used statistical learning methods for predicting compounds with specific properties.

More recently, other statistical learning methods such as neural networks and support vector machines have been explored for predicting compounds of higher structural diversity than those covered by QSAR and QSPR approaches. These methods have shown promising potential in a number of studies. Many attempts have been made to elucidate the QSAR of antimicrobials by using different physicochemical parameters.

This chapter is intended to review some of the strategies and current progresses in using statistical learning methods for predicting compounds with specific desired properties in antimicrobial activity.

General Overview of Methods Used in QSAR of Antimicrobial Compounds

Antimicrobial Activity Determination

As microbiological methods incorporate viable test microorganisms, predictability of the outcome is not always clear and subject to many environmental influences that may impact on a given response. Hence, it is of the utmost importance that parameters such as plant collection, validation of laboratory equipment, chemical analysis, and various intricacies of antimicrobial investigations be carefully defined.

Although several methods are used to assess antimicrobial activity, the two most common methods used are disk diffusion and minimum inhibitory concentration (MIC) assays. Generally disk diffusion studies are the method of choice due to their simplicity and capacity to analyze a large number of test samples. Even though disk diffusion methodology is a quick simple means of screening for antimicrobial activity, it is a qualitative assay because it shows if there is antimicrobial activity or not. The MIC measurement to determine antimicrobial activity is a quantitative method based on the principle of contact of a test organism to a series of dilutions of test substance. Assays involving MIC methodology are widely used and an accepted criterion for measuring the susceptibility of organisms to inhibitors (Lambert and Pearson 2000).

Quantitative Structure–Activity Relationship (QSAR)

Quantitative structure–activity relationship (QSAR) analysis is in constant development since the works of Hansch (Hansch and Fujita 1964) in early 1960s. The QSAR methodology focuses on finding a model, which allows one for correlating the activity to structure within a family of compounds. QSAR studies can reduce the costly failures of drug candidates in clinical trials by filtering the combinatorial libraries. Virtual filtering can eliminate compounds with predicted toxic or poor pharmacokinetic properties (Hodgson 2001; van de Waterbeemd and Gifford 2003) early in the pipeline.

Considering activity optimization, building target-specific structure-activity models based on identified hits can guide high throughput screening (HTS), a recent technological improvement in drug discovery pipeline, by rapidly screening the library for most promising candidates (Dudek et al. 2006). Such focused screening can reduce the number of experiments and allows for use of more complex and low throughput assays (Bajorath 2002). Interpretation of created models gives insight into the chemical space in proximity of the hit compound. Feedback loops of high-throughput and virtual screening, resulting in sequential screening approach (Lewis 2005), allow therefore for more rational progress toward high quality lead compounds. Later in the drug discovery pipeline, accurate QSAR models constructed on the basis of the lead series can assist in optimizing the lead (Prado-Prado et al. 2008).

Main stages of a QSAR study can be divided into three groups, i.e., extracting descriptors from molecular structure, choosing those informative in the context of the analyzed activity, and, finally, using the values of the descriptors as independent variables to define a mapping that correlates them with the activity in question (Dudek et al. 2006).

Generation of Molecular Descriptors from Structure

Compounds are defined by their structure, encoded as a set of atoms and covalent bonds between them. However, the structure cannot be directly used for creating structure-activity mappings because it does not usually contain in an explicit form the information that relates to activity. Molecular descriptors accentuate different chemical properties implicit in the structure of the molecule and so those properties may correlate more directly with the activity. Such properties range from physicochemical and quantum-chemical to geometrical and topological features.

On the other hand, most methods of statistical data analysis employed to predict the activity require as input numerical vectors of features of uniform length for all molecules. Chemical structures of compounds are diverse in size and nature and as such do not fit into this model directly. Molecular descriptors convert the structure to the form of well-defined sets of numerical values.

Selection of Relevant Molecular Descriptors

Although it is possible to generate hundreds or thousands of different molecular descriptors, only some of them are significantly correlated with the activity. Furthermore, many of the descriptors are intercorrelated. This has negative effects on several aspects of QSAR analysis. Some statistical methods require that the number of compounds is significantly greater than the number of descriptors and using large descriptor sets would require large datasets. Other methods, while capable of handling datasets with large descriptors to compounds ratios, nonetheless

suffer from loss of accuracy. Large number of descriptors also affect interpretability of the final model. To tackle these problems, a wide range of methods for automated narrowing of the set of descriptors to the most informative ones is used in QSAR analysis.

Mapping the Descriptors to Activity

Once the relevant molecular descriptors are computed and selected, the final task of creating a function between their values and the analyzed activity can be carried out. The value quantifying the activity is expressed as a function of the values of the descriptors. The most accurate mapping function from some wide family of functions is usually fitted based on the information available in the training set, i.e., compounds for which the activity is known. A wide range of mapping function families can be used, including linear or nonlinear ones, and many methods for carrying out the training to obtain the optimal function can be employed.

Descriptors Used in Antimicrobial Activity Studies

The QSAR techniques involve correlating the logarithm of the reciprocal molar concentration of a bioactive compounds required for a specific biological response such as ED_{50} (dose of a drug that is pharmacologically effective for 50 % of the population exposed to the drug or a 50 % response in a biological system that is exposed to the drug) or LD_{50} (Lethal Dose 50 %, dose that kills half population tested) values with linear free energy constants such as the Hammett constant σ , a measure of aromatic substituent electronic effects; Taft polar constant σ^* , a measure of substituent polar effects; the logarithm of the 1-octanol/water partition coefficient ($\log P$), a measure of hydrophobic-hydrophilic effects; and the Taft steric constant E_s , a measure of substituent steric effects, etc.

A very important role may be played by computer aided drug design techniques based on multi-target quantitative structure–activity relationships (mt-QSAR) studies. It means that they are models connecting the structure of drugs with the biological activity against different targets (microbial species in the case of antimicrobial drugs) (Prado-Prado et al. 2008, 2009). This kind of study may also help in a multi objective optimization (MOOP) of desired properties or activity of drugs against different targets; see for instance the recent works carried out by Cruz-Monteagudo in the topic (Cruz-Monteagudo et al. 2008a, b). In principle, up to date there are over 1,500 molecular descriptors that may be generalized and used to solve the former problem (González et al. 2004; Kubinyi 1990; Marrero-Ponce et al. 2005; Todeschini and Consonni 2002).

Many of these indices are known as topological indices (TIs) or simply invariants of a molecular graph, whose vertices are atoms weighed with physicochemical properties (mass, polarity, electro negativity, or charge) (Estrada and Molina 2001).

QSAR Studies of Antimicrobial Compounds

Coumarins

As a result of the continuously growing interest for antimicrobial activity of coumarins (Al-Haiza et al. 2003; Althaus et al. 1988; Galm et al. 2004; Hoult and Paya 1996; Kulkarni and Patil 1981, 1983; Laurin et al. 1999; Reusser and Dolak 1986; Rodighiero and Antonello 1958; Vieira et al. 2001), a few coumarin antibiotics became candidates for human and veterinary medicine applications. The most important representative is the 3-aminocoumarin derivative novobiocin, the antibiotic that has been relatively recently approved for medical use in the USA for SA infection treatment (Schmutz et al. 2003). Besides novobiocin, other coumarin derivatives like eskuletin, umbelliferon, and related compounds possess antibacterial properties as well (Heinrich et al. 2004). Antifungal activity has been attributed to some of the coumarin derivatives, including coumarin (1,2-benzopyranone) itself (Mares 1987). Antimicrobial activity of 3-nitrocoumarins and related compounds against *Candida albicans* and *Staphylococcus aureus* has been shown recently (Tisi et al. 2001).

Debeljak et al. (2007) studied 3-nitrocoumarins and related compounds, in order to explore their activity and molecular properties that determine their antimicrobial effects. QSAR models involved most of the 64 descriptors extracted from semiempirical and density functional theory (DFT) founded calculations have been proposed. For the study, literature data containing results of microbiological activity screening of 33 coumarin derivatives against selected clinical isolates of *Candida albicans* and *Staphylococcus aureus* were selected. Candidate molecules were checked by cross-validated models, and selected derivatives were synthesized. Their antimicrobial activities were compared to antimicrobial activities of the representative derivatives from the original set in terms of minimal inhibitory concentration (MIC). High ranking of descriptors consistent with the degree of hydrolytic instability of selected compounds is common to models of antimicrobial activity against both microorganisms. However, descriptor ranking indicates different antimicrobial mechanisms of action of chosen coumarin derivatives against selected microbial species.

A set of 68 coumarins and coumarin derivatives with already reported antifungal activities was selected and eleven attributes were generated in order to represent a relationship between the physicochemical properties and their biological activities. The descriptors were used to perform artificial neural network (ANN) and to build a model for predicting effectiveness of the new ones. With good correlation between the experimental and the predicted MIC values pertaining to all the coumarins, the study paves the way for further researches about antifungal activity of coumarins, and offers a powerful tool in modeling and prediction of their bioactivities (Soltani et al. 2009).

A series of new coumarin derivatives has been synthesized and the *in vitro* antimicrobial activity against thirteen strains of bacteria and three fungal/yeast strains were screened. They were shown to possess a wide range of activities from almost completely inactive compounds to medium active ones (Dekić et al. 2010).

Benzamides

Benzamides are an important class of compounds that show various types of biological activities (Itaru et al. 1973; Mrozik et al. 1969). Oxiclozanide was reported as an antihelmintic agent effective against *Fasciola hepatica* for the treatment of liver fluke infection (Mrozik et al. 1969). The synthesis of some *N*-(hydroxyphenyl) benzamides and phenylacetamides as possible metabolites of antimicrobial active benzoxazoles has been reported (Sener et al. 2000).

A series of substituted benzamides were synthesized by Kumar et al. (2007). The synthesized compounds were evaluated for in vitro antibacterial activity against Gram-positive *Staphylococcus aureus*, *Bacillus subtilis* and Gram-negative *Escherichia coli*, and in vitro antifungal activity against *Aspergillus ficcum* and *Aspergillus parasiticus*. In antibacterial assay minimum inhibitory concentrations (MIC) were determined, and the antifungal activity against the fungal species was determined by serial dilution method. The structural characteristics governing antibacterial activities of substituted benzamides were studied using QSAR methodology. The results obtained indicate that the benzamides are effective against the microbial species tested, and that the *N*-(2-hydroxyphenyl)-3-methoxy-2-nitrobenzamide and *N*-(2-hydroxy-6-carbomethoxyphenyl)-2-benzyloxybenzamide are the most effective ones. A general trend showed that the presence of electron-withdrawing groups (NO₂, Cl) leads to an increase in the activity in comparison to the presence of electron releasing group. The results showed that the antimicrobial activity could be modeled using the topological descriptors, molecular connectivity indices (${}^2\chi^v$ and ${}^2\chi$), and Kiers shape index ($\kappa\alpha_1$). The low residual activity and high cross-validated r^2 values (r^2_{cv}) observed indicated the predictive ability of the developed QSAR models.

Cinnamic Acid

Cinnamic acid plays an important role for the antimicrobial activity (Ahluwalia et al. 1986; Christine et al. 1984; Cremllyn et al. 1984, 1986). Derivatives of cinnamic acid, displaying a broad spectrum of biological activity and low toxicity, are of interest for the purposes of creating new effective drugs based on them (Simonyan 1993). There are several studies on cinnamic acid derivatives with antibacterial and antifungal activity (Ahmed et al. 1995; Lee and Ahn 1998; Ovale et al. 1996; Ramanan and Rao 1987; Srivastava et al. 1999; Tawata et al. 1996). Potential antimicrobial activity of sorbic, cinnamic, and ricinoleic acid derivatives was reported some time ago (Narasimhan et al. 2003). The synthesis and the correlation between physicochemical properties and biological activity were carried out. Later, the evaluation of the in vitro antimicrobial activity of over 30 compounds belonging to a series of esters, substituted derivatives, and amides of cinnamic acid was performed by Narasimhan et al. and the investigation of the relationship between their physicochemical properties and microbiological effects have been widely discussed (Narasimhan et al. 2004). Quantitative structure–activity relationship investigation with multiple linear regression analysis was applied to find a correlation between

different calculated physicochemical parameters of the compounds and biological activity. All the compounds showed a good antibacterial activity against Gram-negative *Escherichia coli* than Gram-positive *Staphylococcus aureus* and *Bacillus subtilis*. The chemical structure of each compound was described by three groups of parameters: steric, electronic and hydrophobic which were selected due to their encouraging effect in describing antimicrobial activity (Narasimhan et al. 2003). The quantitative models relating the structural features of cinnamic acid derivatives and their antimicrobial activity showed that Gram-negative *Escherichia coli* and *Candida albicans* (fungus) were the most sensitive microorganisms. The regression equations obtained for the bacterial species showed the importance of constitutional parameter unsaturation index (U_i), global topological charge index (JGT), and the lipophilic parameter log P in contribution to antibacterial activity.

Flavonones

Among the drugs approved between 1983 and 1994 by either the United States Food and Drug Administration (FDA) or comparable entities in other countries, drugs of natural origin predominated (78 %) in the area of antibacterials (Cragg and Newman 2001). A large number of homoisoflavanones have been isolated from several hyacinthaceous genera including *Eucomis* L'Hér., *Merwillia* Speta, *Ledebouria* Roth, *Veltheimia* Gled. and *Drimiopsis* Lindl. and Paxton (Pohl et al. 2000). Homoisoflavanones belong to a small homogeneous group of naturally occurring oxygen heterocycles which, within the Hyacinthaceae, are largely but not exclusively restricted to the subfamily Hyacinthoideae. The few reports on the biological activity of homoisoflavanones describe anti-inflammatory, antibacterial, antihistaminic, antimutagenic, and angioprotective properties, and potent phosphodiesterase inhibition (Amschler et al. 1996; Della Loggia et al. 1989; Heller and Tamm 1981).

Du Toit et al. (2007) determined the antibacterial activity of thirteen homoisoflavanones isolated from six Hyacinthaceae species against *Staphylococcus aureus*. They also developed a set of physicochemical parameters that would describe antibacterial activity for these and future compounds. Stepwise multiple linear regression analysis of the data yielded a statistically significant two-component model ($R^2 = 0.81$, $p < 0.003$).

Phenolic Compounds

Natural and synthetic phenolic compounds were evaluated against oral bacteria. Thus, many antimicrobial agents have been developed for the inhibition of halitosis bacteria and thus for the treatment of oral malodor (Giertsen 2004; Greenstein et al. 1997; Hayashi et al. 2007; Loesche 1979). Antibacterial compounds such as chlorhexidine, cetylpyridinium chloride, triclosan, and chlorine dioxide have been

tested either alone or in different combinations. However, most compounds have been known to induce undesired side effects (Rule et al. 2005).

Greenberg et al. performed a systematic evaluation of various phenolic compounds to develop a quantitative structure-active relationship (QSAR) (Greenberg et al. 2007). They observed that a number of phenolic compounds in natural botanic extracts and flavors demonstrated antimicrobial activity. Among them, eugenol, magnolol, honokiol, thymol, and xanthorrhizol showed strong activity against oral bacteria. Magnolia bark extract, a traditional Chinese medicine isolated from the stem bark of *Magnolia officinalis*, consists primarily of magnolol and honokiol, the two phenolic isomers, and has a strong germ-kill activity against oral bacteria.

In the QSAR approach applied for a range of about 20 phenolic compounds the lipophilicity and steric effects were found to be two key factors in determining germ-kill activity (Greenberg et al. 2008). The optimum lipophilicity, measured by the logarithm of the octanol/water partition coefficient, or log P, was found to be 5.5 for *Fusobacterium nucleatum*, a Gram-negative type of oral bacteria that causes bad breath. The optimum log P was found to be 7.9 for *Streptococcus mutans*, a Gram-positive type of oral bacteria that causes tooth decay.

The steric effect of substituents ortho to the phenolic group was found to be critical in reducing antibacterial activity despite having increased lipid solubility approaching the optimum lipophilicity value. The antibacterial activity of phenolic compounds is likely exerted by multiple functions, primarily comes from its capability to act as a nonionic surface-active agent therefore disrupting the lipid–protein interface.

Furan Derivatives

Furan derivatives, both obtained from synthetic and natural sources, have much interest due to the wide range of pharmaceutical applications they have shown (Hofnung et al. 2002; Khan et al. 2005; Kupchan et al. 1971; Shevchenko 1999). A series of synthetic nitrofuranyl amides showed good in vitro inhibitory activity against *Mycobacterium tuberculosis* (Tangallapally et al. 2006; Tomlin 1994) especially 5-nitro-furan-2-carboxylic acid *N*-[4-(4-benzylpiperazin-1-yl)-benzyl]-5-nitro-furan-2-carboxamide and 2-methyl-*N*-phenylfuran-3-carboxamide.

Preparation of furan-3-carboxylic acid and derivatives and their assessment against a panel of microorganisms including yeast-like fungi, bacteria, and algae were reported (Zanatta et al. 2004). Because some of the furan-3-carboxamides exhibited significant in vitro antimicrobial activity, the synthesis and characterization of an extended and planned series of new furan-3-carboxamides was carried out (Zanatta et al. 2007). The obtained furan-3-carboxamides were assessed against a panel of microorganisms including yeast, filamentous fungi, bacteria, and algae. Preliminary antimicrobial activity assays of some of the furan-3-carboxamides exhibited significant in vitro antimicrobial activity. QSAR investigation was applied to find a correlation between the different physicochemical parameters of the compounds studied and their biological activity. Yeasts showed a negative cor-

relation with the indicator variables I_R , but *Saccharomyces cerevisiae* showed a better correlation with the steric and polarity descriptors, Gram-positive and Gram-negative bacteria correlate with the increase of the steric volume and the polarizability parameter and negatively correlated with hardness, and filamentous fungi correlate with the increase of the steric volume and the polarity of groups.

An interesting QSAR study of diacyl-hydrazine derivatives containing furan rings was conducted and compared with the DFT method and AM1-MOPAC method (Zhang et al. 2010). The DFT-optimized conformations and ESP-fitting charges of the target compounds were also used for 3D-QSAR analysis, including CoMFA and CoMSIA. The QSAR results were consistent with the 3D-QSAR results, indicating that the electrostatic and hydrophobic properties of the target compounds were significant to the biological activity.

mt-QSAR Studies

One limitation of almost QSAR models is that they predict the biological activity of drugs against only one species of fungi, virus, bacteria or parasite species. Consequently, the development of multitasking QSAR models (mt-QSAR) to predict drugs activity against different species of antimicrobial agents is of vital importance. These mt-QSARs offer also a good opportunity to construct drug–drug Complex Networks (CNs) that can be used to explore large and complex drug-viral species databases. In very large CNs it is possible to use the Giant Component (GC) as a representative sub-set of nodes (drugs) and but the drug–drug similarity function selected may strongly determines the final network obtained. Several mt-QSAR models were reported to predict the antimicrobial activity against different fungi (González-Díaz et al. 2006), bacteria (Prado-Prado et al. 2007), parasite (Prado-Prado et al. 2008), and virus species (Prado-Prado et al. 2009) and to calculate the parameters for RNAs of both parasites and hosts (González-Díaz et al. 2011).

For example, the most important limitation of antifungal QSAR models is that they predict the biological activity of drugs against only one fungal species due the fact that most of the up-to-date reported molecular descriptors encode only information about the molecular structure. Prado-Prado et al. calculated, within a unifying framework, the probabilities of antifungal action of drugs against many different species based on spectral moment's analysis (Prado-Prado et al. 2009). They calculated new multi-target spectral moments to fit a QSAR model that predicts the antifungal activity of more than 280 drugs against 90 fungi species. Linear discriminant analysis was used to classify drugs into two classes as active or non-active against the different tested fungal species. Moreover, it was developed one single unified equation explaining the antifungal activity of structurally heterogeneous series of compounds against as many fungus species as possible (González-Díaz and Prado-Prado 2008; González-Díaz et al. 2006).

In fact, other mt-QSAR approaches, with demonstrated usefulness, have been introduced recently in medicinal chemistry (González-Díaz et al. 2006; Marrero-

Ponce et al. 2004; Molina et al. 2004). A Markov Model encoding molecular backbones information was introduced in the method named the MARCH-INSIDE, MARKovian CHEmicals IN SILico DESign (González-Díaz et al.). This method allows one to introduce matrix invariants such as stochastic entropies, potentials, and spectral moments for the study of molecular properties (González-Díaz et al. 2005; Ramos de Armas et al. 2004) and they have been largely used for small molecule mt-QSAR problems including the design of flucicidal, anticancer, and antihypertensive drugs (Prado-Prado et al. 2007). Applications to macromolecules have been restricted to the field of RNA without applications to proteins (González-Díaz and Uriarte 2005; González-Díaz et al. 2003, 2007; Saiz-Urrea et al. 2005).

The QSAR models based on different MARCH-INSIDE indices may be very useful to optimize important aspects such as activity, toxicity, or pharmacokinetics using one single model in many bioorganic and medicinal chemistry problems such as estimation of anticocccidial activity, modeling the interaction between drugs and HIV-packaging-region RNA, and predicting proteins and virus activity (González-Díaz et al. 2004, 2006). In recent studies, the MARCH-INSIDE method has been extended to encompass molecular environment, interesting information in addition to molecular structure data (Cruz-Monteagudo et al. 2007).

Multiple applications of MARCH-INSIDE to classic QSAR, macromolecular QSAR, and specially mt-QSAR were discussed (González-Díaz et al. 2007, 2008a, b; Mahiwal et al. 2010).

Finally, with the mt-QSAR methodology it is possible to predict the biological activity of drugs in more general situations than with the traditional QSAR models, whose greatest limitation is predicting the biological activity of drugs against only one microbial species. Then, this methodology improves models and allows one to predict biological activity of many organic compounds against a very large diversity of pathogens microorganisms.

Conclusion

During the last decade an increasing number of reports describe the antimicrobial activity of several compounds. A review of the literature reveals the existence of a broad-spectrum of research activity in the development practices that are used to treat a variety of diseases and significant current progresses in using statistical learning methods for predicting compounds of specific property in antimicrobial activity.

The quantitative relationship between chemical structure and biological activity allows one to predict theoretically bioactivity without resorting to an inordinate amount of time and effort making the experimental determinations. Many attempts have been performed to elucidate the QSAR of antimicrobials by using different physicochemical parameters.

In this review we have discussed published results concerning the QSAR research on antimicrobial properties of synthetic and natural compounds. Antifungal and antimicrobial activity of some coumarin derivatives, various types of biological

activities of benzamides, several studies on cinnamic acid derivatives with antibacterial and antifungal activity, reports on the biological activity of homoisoflavanones and phenolic compounds, as well as pharmaceutical applications of furan derivatives have been shown.

The actual capabilities of multitasking QSAR methodology and mt-QSAR models to predict drugs activity against different species of antimicrobial agents have also been revised.

It was been demonstrated that with the mt-QSAR methodology it is possible to predict the biological activity of drugs in more general situations than with the traditional QSAR models, predicting biological activity of many organic compounds against a very large diversity of pathogens microorganisms.

Bibliography

- Ahluwalia, V. K., Kaila, N., & Bala, S. (1986). *Indian Journal of Chemistry Section B*, 25, 663.
- Ahmed, S. P., Ahmad, M., Kazmi, S. U., Shahid, M., & Ahmed, S. I. (1995). Studies on the antibacterial and HPFC activity of a phenylpropanoid glycoside. *Pakistan Journal of Pharmacology*, 12(2), 31.
- Al-Haiza, M. A., Mostafa, M. S., & El-Kady, M. Y. (2003). Synthesis and biological evaluation of some new coumarin derivatives. *Molecules*, 8, 275.
- Althaus, I. W., Dolak, L., & Reusser, F. (1988). Coumarins as inhibitors of bacterial DNA gyrase. *The Journal of Antibiotics*, 41, 373.
- Amschler, G., Frahm, A. W., Hatzelmann, A., Kilian, U., Muller-Doblies, D., & Muller-Doblies, U. (1996). Spirocyclic nortriterpenes from the bulbs of *Veltheimia viridifolia*. *Planta Medica*, 62, 534.
- Bajorath, J. (2002). Integration of virtual and high-throughput screening. *Nature Reviews Drug Discovery*, 1, 882.
- Christine, S. V., Rohan, K. G., & Ian, B. R. (1984). The effect of food preservatives on pH homeostasis in *Escherichia coli*. *Journal of General Microbiology*, 130, 2845.
- Cragg, G. M., & Newman, D. J. (2001). Natural product drug discovery in the next millennium. *Pharmaceutical Biology*, 39, 8.
- Cragg, G. M., Newman, D. J., & Snader, K. M. (1997). Natural products in drug discovery and development. *Journal of Natural Products*, 60, 52.
- Cremlyn, R. J., Thandi, K., & Wilson, R. (1984). *Indian Journal of Chemistry Section B*, 23, 94.
- Cremlyn, R. J., Obioran, O., & Singh, G. (1986). *Indian Journal of Chemistry Section B*, 25, 559.
- Cruz-Monteagudo, M., González-Díaz, H., Agüero-Chapin, G., Santana, L., Borges, F., Domínguez, R. E., Podda, G., & Uriarte, E. (2007). Computational chemistry development of a unified free energy Markov model for the distribution of 1,300 chemicals to 38 different environmental or biological systems. *Journal of Computational Chemistry*, 28, 1909.
- Cruz-Monteagudo, M., Borges, F., Cordeiro, M. N., Cagide Fajin, J. L., Morell, C., & Ruiz, R. M. (2008a). Desirability-based methods of multiobjective optimization and ranking for global QSAR studies. Filtering safe and potent drug candidates from combinatorial libraries. *Journal of Combinatorial Chemistry*, 10, 897.
- Cruz-Monteagudo, M., Borges, F., & Cordeiro, M. N. (2008b). Desirability-based multiobjective optimization for global QSAR studies: Application to the design of novel NSAIDs with improved analgesic, antiinflammatory, and ulcerogenic profiles. *Journal of Computational Chemistry*, 29, 2445.
- Debeljak, Ž., Šykrbo, A., Jasprica, I., Mornar, A., Plecko, V., Banjanac, M., & Medic-Šaric, M. (2007). QSAR study of antimicrobial activity of some 3-Nitrocoumarins and related compounds. *Journal of Chemical Information and Modeling*, 7, 918.

- Dekić, B., Dekić, V., Radulović, N., Vukićević, R., & Palić, R. (2010). Synthesis of new antimicrobial 4-aminosubstituted 3-nitrocoumarins. *Chemical Papers*, 64(3), 354–359.
- Della Loggia, R., Del Negro, P., Tubaro, A., Barone, G., & Parrilli, M. (1989). Homoisoflavanones as anti-inflammatory principles. *Planta Medica*, 55, 587.
- Demain, A. L. (1999). Metabolites, primary and secondary. In M. C. Flickinger & S. C. Drew (Eds.), *Encyclopedia of bioprocess technology: Fermentation*. New York: Wiley.
- Du Toit, K., Elgorashi, E. E., Malan, S. F., Mulholland, D. A., Drewes, S. E., & Van Staden, J. (2007). Antibacterial activity and QSAR of homoisoflavanones isolated from six Hyacinthaceae species. *South African Journal of Botany*, 73, 236.
- Dudek, A. Z., Arodzb, T., & Gálvez, J. (2006). Computational methods in developing quantitative structure-activity relationships (QSAR): A review. *Combinatorial Chemistry & High Throughput Screening*, 9, 213.
- Estrada, E., & Molina, E. (2001). 3D connectivity indices in QSPR/QSAR studies. *Journal of Chemical Information and Computer Sciences*, 41, 791.
- Galm, U., Dessoy, M. A., Schmidt, J., Wessjohann, L. A., & Heide, L. (2004). In vitro and in vivo production of new aminocoumarins by a combined biochemical, genetic and synthetic approach. *Chemistry & Biology*, 11, 173.
- Giertsen, E. (2004). Effects of mouthrinses with triclosan, zinc ions, copolymer, and sodium lauryl sulphate combined with fluoride on acid formation by dental plaque in vivo. *Caries Research*, 38, 430.
- González, M. P., Morales, A. H., & González-Díaz, H. (2004). A TOPS-MODE approach to predict permeability coefficients. *Polymer*, 45, 2073.
- González-Díaz, H., & Prado-Prado, F. (2008). Unified QSAR and network-based computational chemistry approach to antimicrobials, part 1: Multispecies activity models for antifungals. *Journal of Computational Chemistry*, 29, 656.
- González-Díaz, H., & Uriarte, E. (2005). Proteins QSAR with Markov average electrostatic potentials. *Bioorganic & Medicinal Chemistry Letters*, 15, 5088.
- González-Díaz, H., Gia, O., Uriarte, E., Hernández, I., Ramos, R., & Chaviano, M. (2003). Markovian chemicals “in silico” design (MARCH-INSIDE), a promising approach for computer-aided molecular design I: Discovery of anticancer compounds. *Journal of Molecular Modeling*, 9, 395.
- González-Díaz, H., Molina, R., & Uriarte, E. (2004). Markov entropy backbone electrostatic descriptors for predicting proteins biological activity. *Bioorganic & Medicinal Chemistry Letters*, 14, 4691.
- González-Díaz, H., Agüero, G., Cabrera, M. A., Molina, R., Santana, L., Uriarte, E., Delogu, G., & Castañedo, N. (2005). Unified Markov thermodynamics based on stochastic forms to classify drugs considering molecular structure, partition system, and biological species: Distribution of the antimicrobial G1 on rat tissues. *Bioorganic & Medicinal Chemistry Letters*, 15, 551.
- González-Díaz, H., Perez-Bello, A., Uriarte, E., & González-Díaz, Y. (2006a). QSAR study for mycobacterial promoters with low sequence homology. *Bioorganic & Medicinal Chemistry Letters*, 16, 547.
- González-Díaz, H., Prado-Prado, F. J., Santana, L., & Uriarte, E. (2006b). Unify QSAR approach to antimicrobials. Part 1: Predicting antifungal activity against different species. *Bioorganic & Medicinal Chemistry*, 14, 5973.
- González-Díaz, H., Sanchez-Gonzalez, A., & González-Díaz, Y. (2006c). 3D-QSAR study for DNA cleavage proteins with a potential anti-tumor ATCUN-like motif. *Journal of Inorganic Biochemistry*, 100, 1290.
- González-Díaz, H., Molina-Ruiz, V., & Hernández, I. (2007a). MARCH-INSIDE v3.0 (MARKov CHains INvariants for SIMulation & DESign), Cuba.
- González-Díaz, H., Saiz-Urrea, L., Molina, R., Santana, L., & Uriarte, E. (2007a). A model for the recognition of protein kinases based on the entropy of 3D van der Waals interactions. *Journal of Proteome Research*, 6, 904.

- González-Díaz, H., Vilar, S., Santana, L., & Uriarte, E. (2007b). Medicinal chemistry and bioinformatics—current trends in drugs discovery with networks topological indices. *Current Topics in Medicinal Chemistry*, 7, 1015.
- González-Díaz, H., González-Díaz, Y., Santana, L., Ubeira, F. M., & Uriarte, E. (2008a). Proteomics, networks and connectivity indices. *Proteomics*, 8, 750.
- González-Díaz, H., Prado-Prado, F., & Ubeira, F. M. (2008b). Predicting antimicrobial drugs and targets with the MARCH-INSIDE approach. *Current Topics in Medicinal Chemistry*, 8, 1676.
- González-Díaz, H., Vilar, S., & Pérez-Montoto, L. G. (2011). 6. Predicting parasite-host networks with Markov Entropy measures for secondary structures of RNA phylogenetic biomarkers. In H. González-Díaz, F. J. Prado-Prado, & X. García-Mera (Eds.), *Complex network entropy: From molecules to biology, parasitology, technology, social, legal, and neurosciences*. Kerala: Transworld Research Network.
- Greenberg, M., Urnezis, P., & Tian, M. (2007). Compressed mints and chewing gum containing magnolia bark are effective against bacteria that are response for oral malodor. *Journal of Agricultural and Food Chemistry*, 55, 9465.
- Greenberg, M., Dodds, M., & Tian, M. (2008). Naturally occurring phenolic antibacterial compounds show effectiveness against oral bacteria by a quantitative structure-activity relationship study. *Journal of Agricultural and Food Chemistry*, 56, 11151.
- Greenstein, R. B.-N., Goldberg, S., Marku-Cohen, S., Sterer, N., & Rosenberg, M. (1997). Reduction of oral malodor by oxidizing lozenges. *Journal of Periodontology*, 68, 1176.
- Hall, B. G. (2004). Predicting the evolution of antibiotic resistance genes. *Nature Reviews Microbiology*, 2, 430.
- Hansch, C., & Fujita, T. (1964). ρ - σ - π analysis. A method for the correlation of biological activity and chemical structure. *Journal of the American Chemical Society*, 86, 1616.
- Hayashi, Y., Ohara, N., Ganno, T., Yamaguchi, K., Ishizaki, T., Nakamura, T., & Sato, M. (2007). Chewing chitosan-containing gum effectively inhibits the growth of cariogenic bacteria. *Archives of Oral Biology*, 52, 290.
- Heinrich, M., Barnes, J., Gibbons, S., & Williamson, E. M. (2004). *Fundamentals of pharmacology and phytotherapy*. Edinburgh: Churchill Livingstone.
- Heller, W., & Tamm, C. (1981). Homoioflavanones and biogenetically related compounds. *Fortschritte der Chemie Organischer Naturstoffe*, 40, 106.
- Hodgson, J. (2001). ADMET-turning chemicals into drugs. *Nature Biotechnology*, 19, 722.
- Hofnung, M., Quillardet, V. M., & Touati, E. (2002). Genotoxicity of 2-nitro-7-methoxynaphtho[2,1-b]furan (R7000): A case study with some considerations on nitrofurantoin and nifuroxazide. *Research in Microbiology*, 153, 427.
- Hoult, J. R. S., & Paya, M. (1996). Pharmacological and biochemical actions of simple coumarins: Natural products with therapeutic potential. *General Pharmacology*, 27, 713.
- INIFTA, Divisin Química Terica, Departamento de Química, Facultad de Ciencias Exactas, UNLP Diag. 113 y 64, Suc. 4, C.C. 16, 1900 La Plata.
- Itaru, S., Seizo, M., & Takashi, K. (1973). Japan Patent, 73, 37, p. 819.
- Khan, M. W., Alam, M. J., Rashid, M. A., & Chowdhury, R. (2005). A new structural alternative in benzo[b]furans for antimicrobial activity. *Bioorganic & Medicinal Chemistry*, 13, 4796.
- Kubinyi, H. J. (1990). Quantitative structure-activity relationships (QSAR) and molecular modeling in cancer research. *Journal of Cancer Research and Clinical Oncology*, 116, 529.
- Kulkarni, M. V., & Patil, V. D. (1981). Studies on coumarins I. *Archives of Pharmacology*, 314, 708.
- Kulkarni, M. V., & Patil, V. D. (1983). Studies on coumarins II. *Archives of Pharmacology*, 316, 15.
- Kumar, A., Narasimhanb, B., & Kumar, D. (2007). Synthesis, antimicrobial, and QSAR studies of substituted benzamides. *Bioorganic & Medicinal Chemistry*, 15, 4113.
- Kupchan, S. M., Eakin, M. A., & Thomas, A. M. (1971). Tumor inhibitors. 69. Structure-cytotoxicity relationships among the sesquiterpene lactones. *Journal of Medicinal Chemistry*, 111, 1147.

- Lambert, R. J. W., & Pearson, J. (2000). Susceptibility testing: Accurate and reproducible minimum inhibitory concentration (MIC) and non-inhibitory concentration (NIC) values. *Journal of Applied Microbiology*, 88, 784.
- Laurin, P., Ferroud, D., Klich, M., Dupuis-Hamelin, C., Mauvais, P., Lassaigne, P., Bonnefoy, A., & Musicki, B. (1999). Synthesis and in vitro evaluation of novel highly potent coumarin inhibitors of gyrase B. *Bioorganic & Medicinal Chemistry Letters*, 9, 2079.
- Lee, H. S., & Ahn, Y. J. (1998). Growth-inhibiting effects of *Cinnamomum cassia* bark-derived materials on human intestinal bacteria. *Journal of Agricultural and Food Chemistry*, 46, 8.
- Lerner, S. A. (1998). Clinical impact of antibiotic resistance. In B. P. Rosen & S. Mobashery (Eds.), *Resolving the antibiotic paradox. Progress in understanding drug resistance and development of new antibiotics*. New York: Plenum Publishers.
- Lewis, R. A. (2005). A general method for exploiting QSAR models in lead optimization. *Journal of Medicinal Chemistry*, 48, 1638.
- Loesche, W. (1979). Clinical and microbiological aspects of chemotherapeutic agents used according to specific plaque hypothesis. *Journal of Dental Research*, 58, 2404.
- Mahiwal, K., Kumar, P., & Narasimhan, B. (2010). Synthesis, antimicrobial evaluation, ot-QSAR and mt-QSAR studies of 2-amino benzoic acid derivatives. *Medicinal Chemistry Research*, <https://springerlink3.metapress.com/content/v8rr537t3g832716/resource-secured/?target=fulltext.pdf&sid=hgvns0moaos2mq3vea5ali2b&sh=www.springerlink.com>
- Mares, D. (1987). Antimicrobial activity of protoanemonin, a lactone from Ranunculaceous Plants. *Mycopathologia*, 98, 133.
- Marrero-Ponce, Y., Díaz, H. G., Zaldívar, F., Torrens, V. R., & Castro, E. A. (2004). 3D-Chiral quadratic indices of the “molecular pseudograph’s atom adjacency matrix” and their application to central chirality codification: Classification of ACE inhibitors and prediction of σ -receptor antagonist activities. *Bioorganic & Medicinal Chemistry*, 12, 5331.
- Marrero-Ponce, Y., Marrero, R. M., Torrens, F., Martínez, Y., Bernal, M. G., & Zaldívar, V. R. (2005). Non-stochastic and stochastic linear indices of the molecular pseudograph’s atom-adjacency matrix: A novel approach for computational in silico screening and “rational” selection of new lead antibacterial agents. *Journal of Molecular Modeling*, 12(3), 255–271, (Online), 1.
- Mcdermott, P. F., Walker, R. D., & White, D. G. (2003). Antimicrobials: Modes of action and mechanisms of resistance. *International Journal of Toxicology*, 22(2), 135.
- Molina, E., González-Díaz, H., González, M. P., Rodríguez, E., & Uriarte, E. (2004). Designing antibacterial compounds through a topological substructural approach. *Journal of Chemical Information and Computer Science*, 44, 515.
- Mrozik, H., Jones, H., Frieddman, J., Schwartzkopf, G., Schardt, R. A., Patchett, A. A., Hoff, D. R., Yaktis, J. J., Riek, R. F., Ostlind, D. A., Plishker, G. A., Butler, R. W., Cuckler, A. C., & Campbell, W. C. (1969). A new agent for the treatment of liver fluke infection (fascioliasis). *Experientia*, 25, 883.
- Narasimhan, B., Kothawade, U. R., Pharande, D. S., Mourya, V. K., & Dhake, A. S. (2003). Syntheses and QSAR studies of sorbic, cinnamic and ricinoleic acid derivatives as potential antibacterial agents. *Indian Journal of Chemistry*, 42(B), 2828.
- Narasimhan, B., Belsare, D., Pharande, D., Mourya, V., & Dhake, A. (2004). Esters, amides and substituted derivatives of cinnamic acid: Synthesis, antimicrobial activity and QSAR investigations. *European Journal of Medicinal Chemistry*, 39, 827.
- Ovale, J. M., Landman, D., Zaman, M. M., Burney, S., & Sathe, S. S. (1996). In vitro activity of *Cinnamomum zeylanicum*. *American Journal of Chinese Medicine*, 24, 103.
- Pohl, T. S., Crouch, N. R., & Mulholland, D. A. (2000). Southern African Hyacinthaceae: Chemistry, bioactivity and ethnobotany. *Current Organic Chemistry*, 4, 1287.
- Prado-Prado, F. J., González-Díaz, H., Santana, L., & Uriarte, E. (2007). Unified QSAR approach to antimicrobials. Part 2: Predicting activity against more than 90 different species in order to halt antibacterial resistance. *Bioorganic & Medicinal Chemistry*, 15, 897.
- Prado-Prado, F. J., González-Díaz, H., de la Vega, O. M., Ubeira, F. M., & Chou, K. C. (2008). Unified QSAR approach to antimicrobials. Part 3: First multi-tasking QSAR model for input-

- coded prediction, structural back-projection, and complex networks clustering of antiprotozoal compounds. *Bioorganic & Medicinal Chemistry*, 16, 5871.
- Prado-Prado, F. J., Borges, F., Perez-Montoto, L. G., & González-Díaz, H. (2009a). Multi-target spectral moment: QSAR for antifungal drugs vs. different fungi species. *European Journal of Medicinal Chemistry*, 44, 4051.
- Prado-Prado, F. J., de la Vega, O. M., Uriarte, E., Ubeira, F. M., Chou, K. C., & González-Díaz, H. (2009b). Unified QSAR approach to antimicrobials. 4. Multi-target QSAR modeling and comparative multi-distance study of the giant components of antiviral drug-drug complex networks. *Bioorganic & Medicinal Chemistry*, 17, 569.
- Ramanan, P. N., & Rao, M. N. (1987). Antimicrobial activity of cinnamic acid derivatives. *Indian Journal of Experimental Biology*, 25, 42.
- Ramos de Armas, R., González-Díaz, H., Molina, R., Pérez González, M., & Uriarte, E. (2004). Stochastic-based descriptors studying peptides biological properties: modeling the bitter tasting threshold of dipeptides. *Bioorganic & Medicinal Chemistry*, 12, 4815.
- Reusser, F., & Dolak, L. A. (1986). Novenamamine is the active moiety in novobiocine. *The Journal of Antibiotics*, 39, 272.
- Rodighiero, G., & Antonello, C. (1958). Sintesi di Alcuni Derivati 3-Ammino Cumarinici e Prime Notizie Sulle Loro Proprieta' Antibatteriche. *Bollettino Chimico Farmaceutico*, 97, 592.
- Rule, K. L., Ebbett, V. R., & Vikes, P. J. (2005). Formation of chloroform and chlorinated organics by free-chlorine-mediated oxidation of triclosan. *Environmental Science and Technology*, 39, 3176.
- Saiz-Urra, L., González-Díaz, H., & Uriarte, E. (2005). Proteins Markovian 3D-QSAR with spherically-truncated average electrostatic potentials. *Bioorganic & Medicinal Chemistry*, 13, 3641.
- Schmutz, E., Mühlenweg, A., Li, S., & Heide, L. (2003). Resistance genes of aminocoumarin producers: Two type II topoisomerase genes confer resistance against coumermycin A1 and clorobiocin. *Antimicrobial Agents and Chemotherapy*, 47, 869.
- Sener, E. A., Bingol, K. K., Oren, I., Arpacı, O. T., Yalcin, I., & Altanlar, N. (2000). Synthesis and microbiological activity of some *N*-(*o*-hydroxyphenyl)benzamides and phenylacetamides as the possible metabolites of antimicrobial active benzoxazoles: Part II. *Il Farmaco*, 55, 469.
- Shevchenko, N. E. (1999). Synthesis of 3-substituted furylethylamines. *Chemistry of Heterocyclic Compounds*, 35, 164.
- Shu, Y. Z. (1998). Recent natural products based drug development: A pharmaceutical industry perspective. *Journal of Natural Products*, 61, 1053.
- Simonyan, A. V. (1993). Activity of cinnamic acid derivatives and new methods for their synthesis (review). *Pharmaceutical Chemistry Journal*, 27(2), 92.
- Soltani, S., Dianat, S., & Sardari, S. (2009). Forward modeling of the coumarin antifungals; SPR/SAR based perspective avicenna. *Journal of Medical Biotechnology*, 1(2), 95.
- Srivastava, A., Shukla, Y. N., & Kumar, S. (1999). Phytotoxic and antimicrobial constituents of *Argyrea speciosa* and *Oenothera biennis*. *Journal of Ethnopharmacology*, 67, 241.
- Tangallapally, R. P., Lee, R. E. B., Lenaerts, A. J. M., & Lee, R. E. (2006). Synthesis of new and potent analogues of anti-tuberculosis agent 5-nitro-furan-2-carboxylic acid 4-(4-benzylpiperazin-1-yl)-benzylamide with improved bioavailability. *Bioorganic & Medicinal Chemistry Letters*, 16, 2584.
- Tawata, S., Taira, S., Kobamoto, N., Zhu, J., Ishihara, M., & Toyama, S. (1996). Synthesis and antifungal activity of cinnamic and esters. *Bioscience Biotechnology & Biochemistry*, 60, 909.
- Tisi, R., Coccetti, P., Banfi, S., & Martegani, E. (2001). 3-Nitrocoumarin is an efficient inhibitor of budding yeast phospholipase-C. *Cell Biochemistry and Function*, 19, 229.
- Todeschini, R., & Consonni, V. (2002). *Handbook of molecular descriptors*. Bicocca: Wiley-VCH.
- Tomlin, C. (1994). *The pesticide manual* (10th ed.). Cambridge: British Crop Protection Council and Royal Society of Chemistry.
- van de Waterbeemd, H., & Gifford, E. (2003). ADMET in silico modelling: Towards prediction paradise? *Nature Reviews Drug Discovery*, 2, 192.

- Vieira, P. C., Mafezoli, J., Pupo, M. T., Fernandes, J. B., da Silva, M. F. G. F., de Albuquerque, S., & Pavao, F. (2001). Strategies for the isolation and identification of trypanocidal compounds from the Rutales. *Pure and Applied Chemistry*, *73*, 617.
- Zanatta, N., Faoro, D., Silva, S. C., Bonacorso, H. G., & Martins, M. A. P. (2004). Convenient synthesis of furan-3-carboxylic acid and derivatives. *Tetrahedron Letters*, *45*, 5689.
- Zanatta, N., Alves, S. H., Coelho, H. S., Borchhardt, D. M., Machado, P., Flores, K. M., da Silva, F. M., Spader, T. B., Santurio, J. M., Bonacorso, H. G., & Martins, M. A. P. (2007). Synthesis, antimicrobial activity, and QSAR studies of furan-3-carboxamides. *Bioorganic & Medicinal Chemistry*, *15*, 1947.
- Zhang, L., Cui, Z., Yin, B., Yang, G., Ling, Y., & Yang, X. (2010). QSAR and 3D-QSAR studies of the diacyl-hydrazine derivatives containing furan rings based on the density functional theory. *Science China: Chemistry*, *53*(6), 1322.

Index

A

- Abasic sites, 1902
- Ab initio methods, 650, 1600
- Ab initio molecular dynamics (AIMD),
 - 343–344, 1902, 1923
 - simulation, 365–368
- Absorption, distribution, metabolism,
 - excretion (ADME) prediction models, 2266–2268, 2275
 - absorption, 2278–2279
 - distribution, 2279–2282
 - excretion, 2287–2288
 - metabolism, 2282–2287
 - for oral administration of drugs, 2267
 - physicochemical property, 2268–2274
- Absorption spectra
 - coumarin in solution, 1981–1983
 - one-photon, 778
 - two-photon, 778
- Acetylcholine, 2053
- Acetylcholinesterase, 2053
- Active orbitals, 165
- Additivity, 24
- ADFT-GIAO methodology, 820–828
- Adiabatic approximation, 182–188,
 - 202–208
- Adiabatic electron affinity (AEA), 688,
 - 1756, 1899
- Adiabatic local density approximation (ALDA), 782
- Adjacency matrix, 2071
- Adjustable parameter, 25
- Adsorption, hydrogen, 1504
- Advanced concepts in electronic structure (ACES II), 873–874
- Advanced concepts in electronic structure (ACES III), 874
- Ag clusters, 1366
- Agglomerative hierarchical clustering,
 - 2102, 2103
- Al clusters, 1366
- Algebraic approximation, 5
- AMBER, 51
- γ aminobutyric acid, 2053
- Aminopyrimidine, 1712–1715
- AMOEBAs, 45
- Anaerobic glycolysis, 1559
- Androgen receptor (AR) antagonists, 2246
- Angle bending interactions, 361
- Anharmonic approximation, 197–198
- Anionic systems, electronic states of,
 - 686–689
- ANO-RCC, 919
- Anthocyanidins, 2057
- Antimicrobial compounds
 - benzamides, 2347
 - cinnamic acid, 2347–2348
 - coumarins, 2346
 - flavanones, 2348
 - furan derivatives, 2349–2350
 - phenolic compounds, 2348–2349
- Antisymmetrizer, 111
- Apocarotenoid oxygenase (ACO), 1581
- Applicability domain (AD), 2068, 2234
 - binary, 2317
 - definitions, 2316
 - leverage-based, 2316
- Approximation, 22
- Aquatic toxicity, 2329–2330
 - of personal care products, 2247
 - of pharmaceuticals, 2250–2251
- Aquatic toxicity index (ATI), 2249, 2251
- Armchair graphene ribbon model, 1486
- Artificial light-harvesting systems, 1934–1935
- Artificial neural networks (ANNs), 2312

- Assumption, 22, 25
Asymptotic correction, 306
Atom(s), 78–79
Atom-atom approach, 32
Atom-atom potential, 26, 32
 functions, 32
Atom-atom scheme, 33
Atom-atom term, 26
Atomic balance, 891
Atomic charge, 26, 2043
Atomic force microscope (AFM), 1649
Atomic mean field (AMF), 918
 approach, 917
 molecular dynamics simulations, 339
Atom-pair (AP) descriptors, 2081, 2307
Atoms in molecules (AIM), 783, 2045, 2056–2057
Atmospheric degradability index, 2240
Au clusters, 1366, 1385–1391
Aufbau/abbau method, 1366
Auger recombination, 1284–1291
Autocorrelation descriptors, 2079–2082
Auxiliary basis, 307
Auxiliary density functional theory (ADFT), 796, 806–808
Auxiliary density perturbation theory (ADPT), 798, 808–810
Average electrostatic potential, 2056
Axial birefringence, 543
Axialstrain, 1326
- B**
- Back-donation effect, 1053
Band topologies, 944–946
Bare cations, 1829–1834
Barostats, 378–380
Barysz–Sadlej–Snijders (BSS) approach, 914
Base pairing, 1808
Basin hopping algorithm, 1462
Basin hopping method, 1367
Basis set cage effects, 687
Basis Set Superposition Error (BSSE), 302, 315, 608, 709, 718–720, 752, 1144, 1811
 scheme, 1831
Basolo–Pearson autocatalytic mechanism, 1843
B-DNA structure, 1808
Bead-spring coarse-grained model, 350
Beeman’s algorithm, 412–413
Benchmarking studies, 227
Bending, 1326
“Benign by design” strategy, 2233
Benzamides, 2347
Benzorods, 1481
Benzotriazoles, 2238, 2249
Berendsen barostat, 378–379
Berendsen thermostat, 374–375, 413–414
Biasing potential replica exchange MD (BP-REMD) simulations, 1689
Bicycle pedal (BP), 1706
Bicyclic azoalkane
 fluorescence quenching, 1965–1971
 photodenitrogenation, 1962–1965
Bimolecular systems photoinduced reactions in, 707
BiNChE plain enrichment analysis, 2174
Biological macromolecules, 2184, 2196
Biological system applications, 1930–1935
Biomolecule, 33, 974–976
Bio-ontologies, 2165
Biopolymer(s), 30, 399
 structure, 22
Birefringences, 542–544
Bismuth vanadate (BiVO₄)
 electronic active centers of, 996–997
 electronic properties of, 1009–1011
 nanostructures, 1021–1023
 optical behaviour of, 997–998
 structural features of, 1006
 structural properties of, 994–996
Bloch function, 1000, 1427
Blood brain barrier (BBB), 2280–2282
B3LYP method, 1778, 1856
Bohr’s formula, 887
Bond angle bending, 25, 37
Bond critical point, 2057
Bonded interactions, 361
Bond length, 25
Bond length alteration (BLA), 1704, 1977
Bond-order parameter, 1379
Bond stretching, 30, 37
Born–Oppenheimer (BO) approximation, 24, 134, 188–190, 208–213, 215–216, 340–343, 642, 647, 1000, 1362, 1632, 1919–1922
Born–Oppenheimer molecular dynamics (BOMD), 343, 380–382, 797
Born–Oppenheimer treatment, 289
Born–von Kármán (BvK) zone, 1418, 1427
Boron nitride, 1487
Bose–Einstein condensates (BECs), 296
Boundary atoms, 349
Breathing mode relaxation, 1473
Breit operator, 893
Breit–Pauli Hamiltonian, 500
Brillouin theorem, 157, 162

- Bromofullerenes, 1073
5-Bromouracil (5-BrU), 1784–1785
Buckingham potential, 26
Buckminsterfullerene, 1179
Burden matrix, 2075
- C**
CAM-B3LYP, 789
Cambridge Serial Total Energy Package (CASTEP), 1006
CAMCASP, 327, 783
Canonical spin orbitals, 654
Carbo index, 2055
Carbon nanotubes applications, 1240, 1242
 computational approaches, 1202
 discovery, 1195–1200
 mechanical properties, 1204
 model structures, 1196
 structural forms, 1195
 tube diameter and chiral angle, 1197
Carbonyl out-of-plane deformation, 1711
Cardiotonic drugs, 2053
Car–Parrinello MD, 343–344, 382–384, 1922–1923
CASPT2, 1949, 1954
CASPT2//CASSCF protocol, 1949
CASSCF energy profile, 1961
CASSCF level, 1704, 1949, 1960
CASSCF/AMBER trajectory study, 1985
Category QSAR, 2314–2315
Cation- π interactions, 2046
Cauchy moments, 527
Cavitation energy, 762
Cavity field factor, 749, 772, 776
Cavity size, 762
CCSD. *See* Coupled-cluster singles-and-doubles (CCSD)
Cellular uptake prediction, 2211–2214
Channelrhodopsins (ChRs), 1932–1933
Characteristic temperature, 1389
Charge asymmetry, 213–214
Charge penetration effect, 762
Charge transfer, 1965–1971
CHARMM, 46
Chebyshev metric, 2100
ChemAxon software, 2306
Chemical entities of biological interest (ChEBI), 2204
 ontology, 2168–2170, 2173, 2176, 2177
Chemical graph, 2008, 2022–2023
Chemical identifier resolver (CIR), 2204
Chemical information (CHEMINF) ontology, 2174, 2175
Chemically induced dynamic nuclear polarization (CIDNP), 1567
Chemical markup language (CML), 2069
Chemical ontology, 2168–2169
 enrichment analysis, 2172–2173
 semantic similarity, 2170–2172
Chemical potential, 2050
Chemical-probe experiments, 1371
Chemical reactions, 622
Chemicals of emerging concern (CEC), 2233
Chemical space (CS), 2004–2006
Chemical text mining, 2170
Cheminformatics, 2042, 2173–2175
Chemistry Development Kit (CDK), 2203
Chemoinformatics, 2140
 chemical compound structure, 2007
 chemical reactions, 2028
 chemical space, 2004–2006
 databases, 2015, 2016
 deductive learning, 2000
 description, 1998
 drug design, 2030–2034
 Enalos KNIME nodes and Enalos Cloud platform, 2205
 fragment-based descriptors, 2028
 history, 1999–2000
 inductive learning, 2000
 IUPAC, 2002
 molecular descriptor, 2021
 molecular physicochemical properties, 2029
 organic synthesis, 2035
 physicochemistry-based descriptors, 2028
 property prediction, 2018
 software tools for, 2203
 structure elucidation, 2034–2035
Chirality molecular topological descriptors (CMTDs), 2307
Chromophore, 1881
 molecular environment, 1971–1974
Cinnamic acid, 2347–2348
Circular birefringence, 543
Circular dichroism, 547–550
Cisplatin, 1860
City block distance, 2100
Clamped nuclei approximation, 70
Clamped nuclei electronic Hamiltonian symmetries, 90–92
Clamped nuclei Hamiltonian, 86–89
 approximate eigenfunctions, 113–119
Classical force fields, 355–356
Classical fullerene, 1185

- Classical molecular dynamics (classical MD),
344–347, 1918
simulation, 368–370
- Classical polarizable electrostatical models,
359–361
- Classical valence bond theory, 653
- Classification and regression tree (CART),
2238
- Classification QSAR model, 2314
- Cleri–Rosato potential, 1365
- Cliffs, 2309
- Cluster geometry, 1340
- Cluster growth, 1383
- Cluster (hyper) polarizabilities, 1117
- Clustering
definition, 2098
hierarchical, 2102–2113
similarity and distance, 2099–2102
clusters, 1098
definition, 1461
Fe, 1473
germanium, 1468
ion mobility, 1466
Ir, 1472
mass spectroscopy, 1462
Mn, 1470
Mo, 1475
Ni, 1475
Pb, 1465, 1468
Pd, 1473
potential energy surface, 1461
Pt, 1472
Ta, 1471
- Cluster property
spin configurations, 1348
fundamental importance, 1340
Landau–Ginsburg–Devonshire theory, 1341
magnetic features, computation, 1348
magnetic storage devices, 1340
melting temperatures, 1354
quantum mechanical procedure, 1349
tight-binding molecular dynamics
methodology, 1342
- CMPSim, 2172
- Coarse-grained molecular dynamics, 350–354
simulations, 339, 1478
- Co clusters, 1469
- Collaborative drug design (CDD), 2017
- Collinear magnetic approximation
atomic MMs directions, 1344
SO interaction, 1345
TB matrix elements, 1345
ZT-TBMD method, 1347
- COLUMBUS, 875
- Comparative field analysis, 2057
- Comparative molecular field analysis
(CoMFA), 2053
descriptors, 2307
- Complete active space (CAS), 1717
- Complete-active-space self-consistent-field
(CASSCF)
method, 1949
approximation, 165
- Complete basis set (CBS) extrapolation
methods, 1811
- Complete linkage method, 2102
- Computational chemistry, 3–17
- Computational microscope, 1643
- Computational toxicity model, 2325
- Computations
interpretation, 1817–1820
level of, 1811–1812
quantum chemical techniques, 1806
- Computer-assisted structure elucidation
(CASE), 2034–2035
- Computer-assisted synthesis design (CASD),
2035
- Conductor-like screening model for real
solvents (COSMO-RS), 760, 2270
- Configuration interaction (CI), 662, 2058
- Configuration interaction-singles (CIS), 662
- Configuration-state functions (CSFs), 572
- Conformation, 30
- Confusion Matrix, 2209
- Conical intersection (CIs), 701–706, 1945
charge transfer, 1965–1971
photofragmentation, 1962–1965
point, 1947
- Conjugate Gradient method, 371–372
- Connection table (CT), 2011
- Connectivity matrix, 2075
- Constitutional descriptors, 2070–2071, 2307
- Continuous QSAR, 2313
- Continuum model, 399–402, 729–732,
750, 768
- Converse piezoelectric effect, 1450–1454
- Coproporphyrinogen III oxidase (CPO),
1560
- Core-excited shape resonances, 1750
- Core orbitals, 165
- Core/shell particles, 1398
- Correlation-consistent basis sets, 630–631
- Correlation time, 1606
- Corrosion inhibitors organic materials as,
2214–2221
- Cost-sensitive learning, 2315
- Cotton–Mouton effect (CME), 551–556
- Coulomb Hamiltonian, 70

- Coumarin(s), 2346
Coumarin in solution, absorption spectra, 1981–1983
Counterpoise correction (CP), 608–609
Coupled cluster (CC), 559
 method, 165
 theory, 1523–1531
Coupled-cluster singles-and-doubles (CCSD), 626
 method, 302, 323, 1812
 model, 2058
Coupled-cluster techniques for computational chemistry (CFOUR), 874–875
Coupled Kohn–Sham (CKS), 305, 306
Coupled perturbed Kohn–Sham (CPKS) methodology, 798
CPMD program, 1936
Crossover operation, 1463
Cross term, 25, 26
Cryogenic EPR/ENDOR technique, 1568
Crystal, 28
 structure, 32
 structure prediction, 325–326
Cu–Ag clusters, 1377
Cu clusters, 1366, 1375
Cumulative atomic multipole moments (CAMP), 770
Cyclomatic number, 2070
Cyclopurine lesion, 1908
CYP enzymes, 2283
Cytosine dimer, 707
Cytochrome oxidases, 1566–1569
Cytosine, 1720
Cytosine-guanine DNA base pair, deactivation mechanism, 1980–1981
- D**
DALTON, 778
Dalton program, 868
DALTON 2.0 program, 328
Damped linear response function, 279
Damping functions, 314
Data sharing, 2017
DBO, 1966
 3db-PSB11, 1959
 4db-PSB11, 1960, 1962
Deactivation mechanism, cytosine-guanine DNA base pair, 1980–1981
555777 Defect, 1482
Defect models, 1503
Deformation energy, 297
Dehalogenation, 1904
deMon program, 797
Dendrogram
 plot, 2104
 types, 2104
de novo design, 2223–2225
 k-means method, 2113
 medoids, 2116
 modes, 2116
 partitioning algorithm, 2113–2116
Density-functional based tight-binding method (DFTB), 1384
Density functional models, 270
Density functional theory (DFT), 23, 226, 271, 303, 602, 796, 799, 986, 1001–1005, 1298, 1922, 1951, 2058
 applicability, 227
 catastrophe, 1455
 complementary, 226
 computational performance, 241–243
 development, 227
 orbitals, 258–260
 electron pairs, 232–235
 holes, 232–235
 methods, 1813, 1905
 model proteins molecular mechanics, 1596
 performance, 227
 protein interactions ligand binding, 1611
 philosophy, 227
 protein hydration, 1609–1610
 theoretical chemistry, 228–232
 and weight derivatives, 487–491
Density functional zoo, 238
Density matrix, 1423
Density-fitting (DF), 307
Density-functional methods, 1384
Deoxyheme, 1575
2-Deoxyribonolactone, 1909
Deposition, 1381
Deprotonation, 1710
Detour matrix, 2074
DFT-D methods, 23
Diagonal Born–Oppenheimer correction, 78
Diamagnetic magnetizability operator, 508
Diamagnetic shielding operator, 508
Diamagnetic spin-orbit operator, 508
Diamond, 1176
Diatomic molecules, 189–190
 homonuclear, 190
 rotations and vibrations of, 191–196
2,3-Diazabicyclo[2.2.1]hept-2-ene (DBH), 1962
1,2-Dichloroethane, 471–487
Dichroism, 542–544
 circular, 547–550
Dielectric constant, 776, 788

- Differential shell approach (DSA), 756, 758
5,6-Dihydropyrimidines, 1902
Dioxygen reactions, 1565–1566
Dipole-dipole interactions, 35
Dipole hyperpolarizability, 523–525
Dipole interaction model, 774
 Applequist, 757, 774
 second hyperpolarizabilities, 774
Dipole moment, 516, 1420–1427, 2043, 2046
Dipole polarizability, 521–523
Dirac–Coulomb–Breit (DCB) Hamiltonian, 893
Dirac electrons, 1492
Dirac equation, 890, 894
 free-particle Dirac Hamiltonian, 901
 free-particle Foldy–Wouthuysen transformation, 901
Direct piezoelectric effect, 1450
Directed walks (DWs), 404
Disconnected structures, 2088–2089
Discrete reaction field (DRF), 747, 765, 773–777, 782, 786, 789
Disk diffusion studies, 2343
Dispersion interaction, 28
Dispersion term, 23
Dispersion-corrected atom-centered potential (DCACP), 605
Dissociation, 1399
Dissociative electron attachment (DEA), 1748
Distance metrics, 2101
Distributed dispersion coefficients, 320–322
Distributed electric properties, 783
Distributed hyperpolarizability, 782–783
Distributed multipole, 316, 782–783
Distributed multipole analysis (DMA), 783–784
Distributed polarizability, 317, 782
Distributing scheme, 751
Divisive method, 2103
DNA, 33, 1804
Docking
 molecular, 2134, 2147
 nanostructured materials, 2151
 protein-ligand, 2142
Double bonds in pentagons (DBIPs), 1068
Double-strand breaks (DSBs), 1745, 1775, 1780, 1901
Douglas–Kroll–Hess (DKH) method, 911–917
3D QSAR, 2134
 advantages, 2135
 developments in, 2141
 ligand-based, 2135
 nanostructured materials, 2151–2156
Drug half-life, 2287
Drug metabolizing enzymes (DMEs), 2278, 2290
Dyes sensitized solar cells (DSSC), 988
- E**
ECEPP. *See* Empirical conformational energy program for peptides (ECEPP)
Ecotoxicity, 2247
Edge adjacency matrix, 2075
Effective properties, 750, 762, 767, 771, 777
 DRF, 775
EFISH. *See* Electric field induced second harmonic generation (EFISH)
Ehrenfest mean-field dynamics, 1926–1927
Eigenfunction, 127
Eigenvalue equation, 127, 133
Elastic collision, 1748
Elbow method, 2121
Electric dipole polarizability frequency-dependent, 518–521
Electric field
 exciting, 745
 local, 745, 747, 769, 777
 permanent, 747, 772, 783
 total, 747
 macroscopic, 744–750, 777
 effective, 775, 777
 optical, 772
Electric field gradient (EFG), 529–530
Electric-field-gradient-induced birefringence, 553–555
Electric field induced second harmonic generation (EFISH), 748, 773, 776, 787–789
Electric field, Maxwell, 763
Electric multipole moment, 515–521, 747
Electric polarizabilities definitions and theory, 1114–1128
Electric properties, 515
Electromechanical measurements, 1325
Electron affinity (EA), 1899, 2043
Electron delocalization interaction, 29
Electron density difference map (EDDM), 937
Electronic structure DFL affinities, 249
 atomization energies, 250
 energy barriers, 251–253
 heats formation, 250–251
Electron paramagnetic spectroscopy (EPR), 989, 991, 992, 996
Electron repulsion integrals, 803
Electron spin, 1561

- Electron spin resonance (ESR), 571–581,
1762, 1774, 1784
 effective spin Hamiltonian, 574–576
 electronic *g*-tensor, 576–577
 hyperfine coupling tensors, 579–581
- Electron transmission spectroscopy (ETS),
1746
- Electron-trap, 1903
- Electronegativity, 2052
- Electronic circular dichroism (ECD), 547–549
- Electronic couplings computation, 720
- Electronic excitations, 641
- Electronic ground state, 271
- Electronic Hamiltonian, 135, 1632
- Electronic motion, influence of, 200–202
- Electronic Numerical Integrator And Computer
(ENIAC), 1630
- Electronic stability, 1900
- Electronic states of anionic systems, 686–689
- Electronic structure calculations, 649
 basis sets, 660–661
 excited states, 662–669
 general overview, 649–653
 methods, advantages, and drawbacks,
 653–660
- Electronic structure DFL ionization potentials,
249
- Electronic structure theory, 226
- Electronic transitions, 1521
- Electrophilic superdelocalizability, 2049
- Electrostatic effects, 1830
- Electrostatic energy, 311–313
- Electrostatic interactions, 359
- Electrostatic potential (ESP) , 2045
 fitting, 973
- Electrostatic shielding, 757
- Elimination of small component (ESC),
894, 900
- Embedded atom models (EAM), 1364, 1465
- Empirical conformational energy program for
peptides (ECEPP), 50
- Empirical valence bond (EVB) method,
1974
- Enalos InSilicoNano Cloud platform,
2206–2211
- Enalos InSilico platform, 2206–2211, 2213
- Enalos iron oxide platform, 2210
- Enalos KNIME node, 2205
- Enalos QNAR iron oxide toxicity platform,
2211
- Endocrine disruptors (EDs), 2245–2247
- Endohedral metallofullerenes (EMFs)
 metal-cage bonding, 1053
 molecular structures, 1062, 1063
 relative energies, 1061, 1062
 violation of isolated pentagon rule,
 1065–1066
- Energy and multiple moments, 514–515
- Engineered NPs (ENPs), 2212
- Enthalpy, 2049
- Entropy, 1374, 2049
- Environmental behavior, 2236–2253
- Environmental effects adding, 679
- Environmental half-lives, 2243
- Enzyme mechanisms, 1617–1621
- Equation-of-motion coupled-cluster theory,
1534–1538
- Estrogen receptor (ER) binders, 2246
- Estrogen-like chemicals, 2245
- Euclidean distance, 2099
- Exact enumeration technique, 407–408
- eXact-2-Component (X2C) method, 914–917
- Excess energy, 1379
- Exchange and correlation energy, 1001
- Exchange correlation energy, 232
- Exchange-correlation functional (XCF),
931–932
- Exchange-correlation kernel, 809
- Excitation energy, 530, 677
- Excited state(s)
 molecular dynamics, 1983–1986
 PCM-CC wave functions, 1535–1538
 BSSE for, 718–720
 properties, 778
 non-adiabatic dynamics, 1716
 MD, 1919
- Exciton, 1396
 in common semiconductor, 1269
 Frenkel, 1271
 nanostructures, 1266–1270
 Wannier–Mott, 1272
- Exciton-polariton, 1260
 photonic crystals, 1275–1276
 polarization and optical properties,
 1270–1275
- Exothermicity, 1230
- Experimental descriptors, 2066
- Explicit solvent model, 764
- Extensive property, 1421
- External forces geometry optimization,
464–465
- External magnetic field effects (MFE), 1579
- External validation, 2234, 2318–2319
- F**
- Faraday effect, 550–551
- Far-bond functions, 309

- Fe clusters, 1473
Fe(II)-Porphyrin-NH₃-O₂ model, 1572
Feed-forward back-propagation neural networks, 2312
Fermi contact (FC) operator, 508
Feshbach resonances, 1750
Firefly, 866–867
First dipole hyperpolarizability, 523–525
First-order optical transitions, 1311
First-order regular approximation (FORA), 897
First-principles plane wave calculations, 1501
Fitness parameter, 1463
Flame Retardants, 2256
Flash vacuum pyrolysis (FVP) technique, 1184
Flavanones, 2348
Flavin adenine dinucleotide (FAD), 1559–1561, 1570
Fluorescence, 940–942
Fluorescent proteins (FPs), 1645, 1881–1885
Fock operator, 150
Folding table (FT), 1707
Foldy–Wouthuysen transformation, 901, 902, 913
ForceBalance, 43
Force fields (FFs), 1363, 1635, 1639–1640
Allinger's, 48–49
AMBER, 52–55
AMOEBA, 46
calibration, 2186–2187
CHARMM, 56–58
data, 2059
ECEPP-05, 50
GROMOS, 56
molecular dynamics simulations model, 339
MMFF94, 49
OPLS-AA (all atom), 56
Scheraga's, 50
Formamide, 1722–1725
Fragment-based descriptors, 2028
Franck–Condon factors, 645
Franck–Condon point, 1945, 1951
Franck–Condon principle, 644
Franck–Condon region, 1699
Freely jointed chain (FJC) model, 399
Free-particle Dirac Hamiltonian, 901
Freeze-and-thaw cycles, 780
Frenkel exciton, 1271
Frequency-dependent density susceptibility (FDDS), 304, 307
Frequency-dependent electric dipole polarizability, 518–521
Frequent subgraph mining approach, 2308
Frontier molecular orbitals, single-walled carbon nanotubes, 1222
Frontier orbital densities, 2045
Frozen-density embedding (FDE), 747, 765, 779–782
Frozen localized orbitals, 349–350
Fukui functions, 2050
Fullerene(s), 842–847
applications, 1239
canonical spiral, 1036
classification, 1185
¹³C NMR spectroscopy, 1048, 1184
computational approaches, 1202–1205
computational studies of isomers, 1210–1214
definition, 1034–1036
DFT calculation, 1043
discovery, 1180
Euler's theorem, 1034
giant, 1214–1215
giant fullerene structures, 1186
hexagonal indices, 1041
IPR isomers, 1042
isolated pentagon rule, 1037, 1191
molecular structures, 1045
nano-capsules, 1189–1190
natural abundance, 1188–1189
 π orbital axis vector analysis, 1039, 1040, 1215
POAV pyramidalization angle, 1041, 1042
pyramidalization angle, 1215–1218
relative energies, 1038, 1044, 1047
Schlegel diagram, 1188, 1189
semiempirical method, 1047
spiral algorithm, 1035
steric strain, 1040
Stone–Wales defects, 1193, 1218
structure, 1183
synthesis, 1184
thermodynamic stability, 1043
total isomers, 1037
vacancy defects, 1194, 1219–1221
X-ray crystallographic studies, 1050
Fullerene derivatives
addition patterns, 1069
bond lengths, 1066
DFT calculations, 1070
energy isomers, 1067
Full quantum chemical studies, 732–733
Fully directed walks (FDWs), 404
Full width at half maximum (FWHM), 995
Functionalized amino acids (FAAs), 2319
Funnel, 1946
sculpting, 2189–2190

Furan derivatives, 2349–2350
Furocoumarins, 671
Furthest neighbor method, 2102

G

Gallium arsenide (GaAs) clusters, 1099, 1109–1112
Gauge, 1419–1420
Gauge invariance, 285–288
Gaussian chain model, 399
Gaussian-type functions (GTFs), 660
GDIIS, 438–439
GDMA, 783
Ge clusters, 1399, 1468
Gene Ontology (GO), 2165
General Atomic and Molecular Electronic Structure System (GAMESS-US), 864–866
General Atomic and Molecular Electronic Structure System (GAMESS-UK), 867–868
Generalized gradient approximation (GGA), 236, 363, 602, 1298
Generalized solvation free energy density (G-SFED) model, 2271
Generalized Tight-Binding Molecular Dynamics (GTBMD), 1342
Generalized valence bond method in the perfect pairing approximation (GVB-PP), 1717
Genetic algorithm (GA), 1385, 1463, 2233
Geometrical descriptors, 2082–2085
Geometrical distance matrix, 2082
Geometries, 1814–1817
Geometry optimization and symmetry, 439–440
Geranylgeranyltransferase I inhibitors (GGTIs), 2322–2323
GETAWAY descriptors, 2084
Giant fullerenes, 1214–1215
Gibbs free energy, 2049
Global descriptor, 2008
Global Half Life Index (GHLI), 2241
Global search algorithm(s), 1462
Global search algorithm of minima (GSAM), 766, 767
Glucosyltransferase model, 1570
Goepfert–Mayer (GM), 1878
Golden cages, 1388
Gold standard, 1812
Graphene, 1485–1505
 applications, 1238
 band gap, 1207

 chemisorption energy, 1209
 computational and experimental studies, 1180
 computational approaches, 1202–1205
 defects in, 1180
 description, 1178
 edges chemistry, 1312
 hydrogenation, 1209
 molecular dynamics simulations, 1208
 Stone–Wales defect, 1205
Graphene nanoribbons (GNRs), 1300, 1303, 1311, 1486
 electronic and mechanical properties, 1492–1500
 electronic properties, 1486
 model structure, 1487
 molecular dynamics computer simulations, 1487
 structures, 1487
Graphene quantum dots, 1503
Graph invariants, 2071
Graphitic systems
 edge effects, 1318–1323
 quantum confinement, 1315
Graph-theoretical matrices, 2072
Green chemistry, 2258
Green fluorescent proteins (GFP), 1601, 1971–1974
Green's function approach, 662
GRIND descriptor, 2308
GRINDOL, 771
GROMACS, 55
GROMOS, 55–58
Ground state structures Mn clusters, 1470
 Ta clusters, 1471
Ground-state (GS), 928–929
 structure, 1099–1101
Gupta potential, 1365, 1465
Gut physiological model, 2289–2291

H

HAIO, 1391–1396
5-Halouracils, 1905
Halonucleobases, 1904
Hamiltonian
 matrix, 133
 scalar potential in, 504–505
 vector potential in, 506
Hamilton operator, 127, 133, 141, 162
Hard and soft acid-base (HSAB), 2051
Hard sphere, 24, 27
Harmonic approximation, 194–196, 1373
Harmonic potential, 25

- Hartree–Fock (HF)
 level, 863
 mean field approach, 1503
 method, 147–152, 626, 653–660, 1072
 potential, 654
 theory, 600–602
- Hazard of chemical, 2257
- HD molecule, 213–214
- Heat capacity, 1374, 1389
- Heat of formation of H_2SO_2 , 634–636
- Hellmann–Feynman theorem, 11
- Helmholtz free energy, 972
- Helmholtz free enthalpy, 1374
- Hermitian operators, 127–129, 132, 159
- Hessian matrix and Hessian updates, 430–431
- Hexagonal indices, 1040, 1041
- Hierarchical cluster analysis (HCA), 2236
- Hierarchical clustering
 advantages and disadvantages, 2103
 definition, 2102
- Hierarchical optimization, 2190–2193
- Higher fullerenes
 addition to C_{70} and, 1078–1082
 definition, 1032
- Higher-order response functions, 281–283
- Highest occupied molecular orbital (HOMO), 2045
- High-performance liquid chromatography (HPLC), 1780
- High throughput screening (HTS), 2344
- High throughput virtual screening, 2210
- Histone deacetylase, 2051
- Hohenberg–Kohn theorem, 231, 800
- HOMO-LUMO gaps, 1317
- Homoisoflavanones, 2348
- Homology modelling, 1599–1600
- Homonuclear diatomic molecules, 190
- Homotops, 1376
- Horseradish peroxidase (HRP), 1566
- HSAB principle, 1396
- Hubbard models, 1004
- Human intestinal absorption (HIA) assay, 2278, 2279
- Hund’s rule, 1562
- Hybrid quantum mechanics/molecular mechanics (QM/MM), 765, 1929–1930
 molecular dynamics, 347–350
 simulations, 340
- Hydrate lattice, 1162, 1167
- Hydrated alkaline earth cations, 1836–1838
- Hydrated copper cations, 1838–1840
- Hydrated electrons, 1897
- Hydrated magnesium cation, 1837
- Hydrogen adsorption, 1504
- Hydrogen atom transfer, 1907
- Hydrogen bond(ing), 26, 300, 596, 1836, 1865
 acidity, 2052
 basicity, 2052
 and stacking interactions, nucleic acids, 1809–1814
- Hydrophobicity, 2049
- Hyper-functionals, 237
- Hyper-Rayleigh scattering (HRS), 788
- Hypermagnetizabilities, 551–553
- Hypernetted chain approximation, 966
- Hyperpolarizability, 518, 1126–1128, 1449, 2045
 distributed, 782–783
 effective, 763
 DRF effective 789
 effective, 750
 solute, 748, 750, 763
- Hypersurface deformation method, 1462
- Hypothetical perfect computer, 7–10
- Hypoxic conditions, 1899
- Hückel’s method, 2057
- I**
- Impact ionization and Auger recombination, 1284–1291
- Implicit solvent model, 759–764, 1834–1835
- In vitro* assays, 2266
- In vivo* assays, 2266
- Index of epitaxy, 1382
- Indigo, 2204
- Induced dipole moment, 2049
- Infinite-order two component (IOTC), 913, 918, 919
- Infrared (IR), 1444
- In silico* models, 2266, 2267
- Insubria graph, 2256
- Integrase inhibitors, 2056
- Intensive property, 1421
- Interaction, 26–27
 dispersion, 752, 779
 electrostatic, 752, 765, 768
 energy exchange-dispersion, 752
 exchange, 752
 exchange-induction, 752
 induced properties, 751–753
 induction, 752
 long-range electrostatic, 766
 polarization, 765

- properties distance dependence, 753–754
- many-body effects, 754–756
- short-range, 765
- short-range repulsion, 779
- 1–4 Interactions, 27
- Interaction energy, 598–599
 - charge-transfer, 308–310
 - components, 301
 - definition, 297
 - density-fitting, 307
 - DFT, 303
 - higher-order contributions, 308
 - many-body contributions, 314–316
 - multipole expansion, 310–314
 - SAPT(DFT), 303, 304
 - two-body interaction energy, 297
- Interaction-site model, 968
- Intermediate neglect of differential overlap (INDO), 2057
- Intermolecular hydrogen transfer, 1965–1971
- Intermolecular interactions, 32
 - BSSE, 1144
 - classification, 1141
 - electron correlation, 1143–1144
 - extrapolation, 1145–1146
 - orbital basis set, 1143–1144
- Intermolecular proton-transfer reaction, 1849
- Internal energy, 1374
- Intersystem-crossing (ISC), 642, 1946
- Intramolecular interaction, 31
- Intrinsic reaction coordinate (IRC), 648
- Intrinsically disordered protein (IDP), 1644
- Inverse coordination number, 1471
- Inverse Monte Carlo (IMC) algorithm, 353
- Inverse Stone–Wales (ISW) defect, 1180
- Ion channels, 1647–1648
- Ionization constant, 2271–2272
- Ionization potential (IP), 1899, 2043
- Ir clusters, 1472
- Iron oxide nanoparticles, 2206
- Isogyric reaction, 625
- Isolated chromophores, 1956–1971
- Isolated pentagon rule (IPR)
 - definition, 1037
 - endohedral metallofullerenes, 1065–1066
 - fullerenes, 1191
- Iterative Boltzmann inversion (IBI), 352
- J**
- Jablonski diagram, 642, 643
- Jahn–Teller distortions, 1471
- Jellium model, 1383
- K**
- Kaiser criterion, 2121
- K-Cs clusters, 1366
- k-nearest neighbor methods, 2312
- k-Nearest Neighbors (kNN) technique, 2217
- k.p* method, 1263
- KcsA channel, 975
- Kernel trick, 2311
- Kidney physiological model, 2294
- Kinetic stability, 1042
- Kiteplatin, 1860
- Kohn–Sham approximation, 781
- Kohn–Sham density functional theory, 799–801
- Kohn–Sham matrix elements, 816
- Kohn–Sham variational principle, 2058
- Konstanz Information Miner (KNIME), 2203
 - design and prediction, 2221–2225
- Koopman’s theorem, 2050, 2123
- Kovalenko–Hirata closure, 969
- Kronecker symbol, 130
- L**
- Lagrangian dynamics, 1922
- Landau–Zener transition probability, 1701
- Langevin dipole model (LD), 735
- Langevin dipoles (QM/LD/MC), 765, 769
- Langevin thermostat, 377, 414
- Langevin–Born-type effect, 553
- LANL2DZ basis set, 1851
- Laplacian matrix, 2073
- Lash’s model of CPO oxidation, 1571
- Lattice models, 403–404
- Layered perovskites, 1453
- LCGTO Kohn–Sham method, 801–806
- LDF approach to exact exchange (LDF-EXX), 817
- Leaching of pesticides, 2236
- Lead structure, 2030
- Leap-frog algorithm, 346, 411
- Lennard–Jones, 26
 - cluster, 1360
 - potential, 779, 1464
- Lhasa Limited nodes, 2204
- Ligand binding, 1610
- Ligand field (LF) excitations, 946
- Ligand-centered (LC) transitions, 947
- Ligand-to-ligand CT (LLCT), 947
- Ligand-to-metal CT (LMCT) excitations, 946
- Light-harvesting systems, 1934–1935
- Line notation, 2009, 2010
- Linear birefringence, 542

- Linear Combination of Atomic Orbitals (LCAO) approach, 651
- Linear combination of Gaussian-type orbitals (LCGTO), 801
- Linear discriminant analysis, 2350
- Linear free energy relationship (LFER)-based model, 2269, 2272
- Linear interpolation in internal coordinates (LIIC), 699
- Linear methods, QSAR, 2311
- Linear response function
- anomalous dispersion, 278
 - conjugation relations, 281
 - eigenvalue problem solving, 276
 - electronic electric dipole, 275
- Linkage
- definition, 2101
 - types, 2102
- Lipinski rule, 2268
- Lipophilicity (logP), 2273
- Liver physiological model, 2292–2293
- Liver tissue, 2293
- Local densities, 235
- Local density approximation (LDA), 363, 602, 1002, 1004
- Local density-fitting (LDF), 817
- Local hardness, 2051
- Local properties (LoProp), 784–786
- Local softness, 2050
- Local vertex invariants, 2076
- Local-field tensor, 749
- Localized dipole–interaction model (LIM), 757
- Long-range correction (LC) scheme, 1003
- Long-range dispersion interaction coefficients, 528–529
- LoProp, 784–786
- Lorentz local field model, 773, 777
- Lorentz-factor tensor, 748, 749, 772
- Lossless dielectric constant QD dimer systems, 1282–1283
- Low-energy electron (LEE)-induced DNA damage
- basis set, 1760
 - 5-BrU and radiosensitization activity, 1784–1785
 - electron-molecule interaction events, 1747–1748
 - excited states, 1773
 - gas-phase, electron attachment, 1756–1758
 - hydration of uracil-5-yl radical, 1785–1787
 - ionizing radiation, 1779–1780
 - mechanism of, 1745–1747
 - shape resonances of, 1752–1756
 - solvation, effect of, 1760–1766
 - 5-substituted uracil, 1788–1789
 - thermodynamic properties and physical observables, solvated electron, 1763
 - TNI formation, 1748
- Lowest unoccupied molecular orbital (LUMO), 1754, 1760, 2045
- Lowest-energy isomers, 1469
- L-shaped GNR models, 1499
- ## M
- M-DNA crosslink stabilization, 1853
- Machine learning molecular dynamics, 1923–1924
- Mackay transition, 1470
- Madin–Darby Canine Kidney (MDCK) cell assay, 2278
- Magic number, 1369, 1383, 1462
- Magic sizes, 1480
- Magnetic circular dichroism, 555–556
- Magnetic dipole operator, 507
- Magnetic moments, 1473
- Magnetic optical rotatory dispersion (MORD), 550
- Magnetic transition metal clusters, 1349
- Magnetic vector potential, 505–506
- Magnetizability, 535–539
- Magnetization, 1469, 1470
- Manhattan distance, 2100
- Many-body perturbation theory (MBPT), 659
- MARCH-INSIDE method, 2351
- Marcus and Pekar partitions, 1521
- Markov model, 2351
- Martini model, 352
- Mass abundance experiment, 1384
- Massively Parallel Quantum Chemistry (MPQC) program, 875
- Materials Studio software, 1008
- Maximum electroshock (MES) test, 2320
- Maximum likelihood optimization, 2193–2195
- Mean quantization error, 2130
- Mechanistic model, 23
- Metal-cage bonding, endohedral metallofullerenes back-donation effect, 1053
- electrostatic potential, 1051
 - formal and actual charge, 1055
 - ionic model, 1052
 - kinetic energy density, 1054

- Kohn–Sham MO levels, 1052
QTAIM, 1053, 1055
- Metal coordination
with base pairs, 1847–1855
ions in DNA quadruplexes, 1863–1865
metal cross-linked DNA structures,
1858–1863
with platinum complexes, 1856–1858
- Metal-to-ligand CT (MLCT) excitations, 946
- Metalation
exocyclic amino group, 1842
Pt-N7 energies, 1857
preferred sites, 1830
Watson–Crick base pairing energy, 1850
- Metallocoordinated, 1406
- Metcars, 1405–1406
- Micelle-free drugs, 2290
- Microelectromechanical systems (MEMS),
1323
- Microsolvation, 766
- Mid-bond functions, 309
- Minimization, 440–443
- Minimum cumulative variance criterion, 2121
- Minimum energy conical intersections
(MECI), 1946
- Minimum energy path (MEP), 648, 698, 1946,
1958, 1960
- Minimum inhibitory concentration (MIC)
assays, 2343
- Mismatched DNA, 1909
- Mixed quantum-classical approaches,
1700–1702
- MMFF94, 49
- MM-like terms, 23
- Mn clusters, 1470
- Mo clusters, 1475
- Model applicability domain, 2256
- Model proteins
homology modelling, 1599
structure prediction, 1599–1600
- Modern theory of polarization, 1419, 1433
- Modified neglect intermediate neglect of
differential overlap (MINDO), 2057
- Modified Newton–Raphson step, 435–438
- MOLCAS, 785
- Molecular Breit–Pauli Hamiltonian,
500–504
- Molecular clusters, 463–464
- Molecular crystal, 30–33
- Molecular descriptor(s), 2097
antimicrobial activity, 2345
chemical graph, 2008, 2021
drug design, 2030–2034
generation, 2344
mapping function, 2345
organic synthesis, 2035
selection of, 2344–2345
structure elucidation, 2034
- Molecular design, 2030
- Molecular dynamics (MD), 1381, 1629, 2043,
2192, 2195
CHARMM, 1638, 1640
development of, 1641
enzymatic activity, mechanism of,
1646–1647
force fields, 1635, 1639
functionally important motions, 1645–1646
GROMACS code, 1639
history, 1630–1631
IDP, 1644
molecular diseases, 1649
NAMD code, 1639
protein–DNA interactions, 1648–1649
protein–drug interactions and docking,
1644–1645
protein folding studies, 1643
simulations, 338, 409, 410, 1647, 1918
single molecule AFM experiments, 1650
spectroscopy experiments, 1645
steps in, 1636
structure optimization/energy
minimization, 1635
transport phenomena in proteins, 1647
- Molecular dynamics (MD) simulations, nucleic
acids, 1676
conformational transitions, 1686
DNA bending, 1683
induced conformational changes, 1677
REMD, 1679
replica-exchange and umbrella sampling
simulations, 1682
RNA, kink-turn motif in, 1685
molecular dynamics simulation, 27
- Molecular electrostatic potential (MEP), 2043,
2053–2056
- Molecular fingerprints, 2086–2087
- Molecular graph, 2010, 2071–2072
- Molecular Hamiltonian, 499–510
- Molecular mechanics (MM), 2043, 2059
Allinger’s force fields and parameter,
48–49
AMBER force fields, 52
application, 27
CHARMM force field, 56–58
description, 22
formulae and parameters, 41–44
intermolecular and non-bonded
interactions, 26

- Molecular mechanics (MM) (*cont.*)
intramolecular contributions, 25–26
mathematical expressions, 24
MMFF94, 49
molecular biology, 33–37
molecular crystal, 31–33
nucleic acid, 37
organic molecules, 30
Scheraga's force field, 50
water molecule, 44
- Molecular operating environment (MOE), 2306
- Molecular orbital configuration of dioxygen, 1564
- Molecular orbitals (MOs), 640
- Molecular properties DFL
bond angles, 246–247
bond energies, 253
bond lengths, 244–246
excited states, 257–258
hydrogen bonding, 253–254
spin states, 256–257
vibrational frequencies, 247–249
weak interactions, 254–256
- Molecular Quantum Numbers (MQNs), 2214
- Molecular recognition, 14–16, 34
- Molecular relativistic corrections, 215–217
- Molecular representation, 2068
- Molecular response, 510–514
- Molecular structure, 425–426
curation, 2069–2070
- Molecular surface, 13–14
- Molecular vibrations, 465–471
- Molecules, 79–86
- Møller–Plesset (MP), 273
- Møller–Plesset perturbation theory (MPPT), 152–159, 600–601, 659
- Møller–Plesset second-order perturbation (MP2), 2058
- MOLPRO 2008, 328
- Monomer-centered plus, 310
- Monte Carlo sampling, 27
- Monte Carlo simulation, 408–409, 2043, 2059
- Motion of nuclei, 178–181
- MP2 method, 1187
- Multi criteria decision making (MCDM), 2248
- Multi-configuration self-consistent field (MCSCF), 863, 2058
- Multi-electron molecule, 175–178
- Multi photon absorption theory, 1877–1878
- Multi-reference model, 2058
- Multi-walled carbon nanotubes
applications, 1240
structure, 1196
- Multiconfigurational methods, 651, 679–680
- Multiconfigurational quantum chemistry, 1948–1954
- Multiconfigurational time-dependent Hartree approach (MCTDH), 1715
- Multiconfiguration (MC) SCF method, 164
- Multiconfiguration self-consistent field (MCSCF), 1700
wave function, 651
- Multidrug Resistance Protein 1 (MDR1), 2286
- Multilayer dielectric continuum model, 757
- Multiphoton process, 1283–1284
- Multiple damage site (MDS), 1743
- Multiple moments, energy and, 514–515
- Multiple-carrier processes, 1283–1291
- Multiple-spawning, 1947
- Multipolar components, 2046
- Multipole(s)
atomic, 774
distributed, 782–783
localized, 785
- Multipole expansion method (MPE), 760, 763–764, 786
- Multipole moments, 778
- Multireference configuration interaction (MRCI), 662, 1700
- Multireference coupled-cluster (MRCC) methods, 663
- Multiscale methods, 1478
- Multitasking QSAR models (mt-QSAR), 2350–2351
- Multivariate explorative methods, 2235–2236
- Murrell–Mottram potential, 1365
- N**
- Na clusters, 1385
- Nanoalloys, 1377
- Nanobelts, 1489
- Nano-capsules, 1189–1190
- Nanoclusters, 1098, 2059
- Nanoelectromechanical (NEMS), 1324
- Nanomaterials, 2089, 2151
- Nanoparticles (NPs), 1461–1464, 2206–2211
engineered, 2212
properties, bioactivity and predictions, 2207
via Enalos InSilicoNano platform, 2211–2214
- Nanoribbons, 1476
electronic properties, 1486
graphene, 1476
- Nanoscale, rules of, 16
- Nanoscience, 1098

- Nanostrips, 1489
Nanostructure(s), 2151–2156
Nanostructure modeling
 empirical potential, 1464–1466
 Lennard–Jones potential, 1465
 local optimization, 1466–1475
 one-dimensional, 1475
 zero-dimensional, 1461
Nanotubes, 1476
 elastic properties, 1476–1479
 SiC, 1482
 zigzag, 1495
Nanowires, 1476
 elastic properties, 1476–1479
 structural properties, 1479–1481
Natural atomic orbitals, 2045
Natural bond orbital, 2045
Natural (electronic) circular dichroism, 547–550
Natural products, 2342
Natural selection, 1463
Natural transition orbitals (NTO), 934
Negative electron affinities, 690–693
Neighborhood function, 2126
NEMS applications
 carbon nanotubes, 1324–1327
 graphene nanoribbons, 1327–1329
Newton equations, 1630
Newton–Raphson step, 428–430
Newton-X program, 1935–1936
NH stretching, 1711
Ni–Ag clusters, 1377
Ni clusters, 1366, 1367, 1375, 1475
Ni–Cu clusters, 1366
NiTiO₃
 electronic properties of, 1012–1013
 nanostructures, 1017–1021
 structural features of, 1006
Non-adiabatic molecular dynamics, 1924–1930
Non-adiabatic transition probabilities, 1919
Non-classical fullerene, 1185
Non-covalent interactions, 595
Non-local spin-density approximation (NLS-D), 1002
Non-relativistic quantum chemistry program
 commercial software, 876
 free software, 864–876
Non-rigid bicycle pedal (NRBP), 1707
Non-rigid rotor, 198–200
Non-symmetric Algebraic Riccati Equation (nARE), 914, 916
Nonbonded energy, 24
Nonbonded interaction, 26, 356–361
Noncollinear spin formalism, 1475
Nonlinear methods, QSAR, 2311
Nonradiative T-S transitions, 1565
Normalized elimination of small component (NESC), 898, 914
Normconserving pseudopotentials, 366–367
Nosé–Hoover thermostat, 375–377, 379–380, 414
np-apomorphine, 2056
NP-hard problem, 1360
Nuclear magnetic resonance (NMR) spectroscopy
 data acquisition, 1593
 effective spin Hamiltonian, 557–559
 indirect spin–spin coupling, 568–571
 properties, 778
 refinement, 1595
 resonance frequency assignment, 1594
 restraint collection, 1594
 sample preparation, 1593
 spin-spin coupling constants and shifts, 1833
 stability testing, 1593
 structure calculation, 1595
 structure validation, 1595
Nuclear magnetic shielding, 559
 spin-orbit corrections to, 566–567
Nuclear motion effects, 288–292
Nuclear quadrupole coupling constant, 529–530
Nuclear receptors (NRs), 2286
NuclearSpin-related properties, 556–571
Nuclear spin-rotation constants (NSRC), 567–568, 825
Nuclear vibrations
 anharmonic approximation, 197–198
 harmonic approximation, 194–196
Nuclear wavefunctions, 72
Nucleic acid, 22, 33
 hydrogen bonding and stacking interactions, 1809–1814
Nucleic acid bases (NABs), 690, 1886–1887
Nucleobase analogues, 1886
Nucleoside analogues
 computational investigations of, 1887–1889
 NWChem, 870–871
- O**
Octopus program, 1936
One-bond flip (OBF), 1703, 1704
One-dimensional structures, 1475
One-photon absorption (OPA), 530–533, 1876
ONIOM approach, 1203
ONIOM method, 1855

- Online Chemical Modeling Environment (OCHEM), 2204
- Onsager local field, 749
- Onsager–Lorentz cavity field factor, 763
- Ontology, 2164–2165
- chemical, 2168–2175
 - content and metadata, 2167–2168
 - languages and logical expressivity, 2165–2167
- OpenBabel cheminformatics package, 2204
- Optical absorption coefficient, 1503
- Optical rotation, 544
- Optimisation algorithms, 370–373
- Optimized Potentials for Liquid Simulations (OPLS), 55–58
- Orbital model, 12–13
- ORCA, 871–872
- Ordinary Least Squares (OLS), 2233
- Organic compounds, 2156
- Organic conjugated materials (OCMs), 1933
- Organic crystals, 324–325
- Organic materials as corrosion inhibitors, 2214
- Organization for Economic Cooperation and Development (OECD), 2208
- ORIENT, 327
- Origin invariance, 285–288
- Ornstein–Zernike equation, 966
- Osmium complexes, 1847
- Overplotting, 2124
- Oversampling, 2315
- Oxaliplatin, 1860
- 8-Oxoguanine, 1689
- Oxyclozanide, 2347
- Oxygen revolution, 1559
- Oxygen-dependent PUVA mechanism, 714
- Oxyheme model, 1572
- P**
- P-glycoprotein (P-gP), 2286
- p-Hydroxybenzylideneimidazolinone (HBDI), 1971
- Pair density distribution function, 966
- Parallel Artificial Membrane Permeability Assay (PAMPA), 2278
- Paramagnetic spin-orbit operator, 507
- Parameter, 25
- Pariser–Parr–Pople (PPP) method, 1753
- Partially molten clusters, 1354
- Particle mesh Ewald (PME) methods, 1674
- Partition coefficients, 2045
- Partition function, 1374
- Partitive approach, 2099
- Pauli principle, 91
- Pb clusters, 1468
- p-benzosemiquinone radical anion, 686–689
- PBT Index, 2255
- PCM cavity field factors, 763
- PCM-CC equations, 1529–1531
- PCM-CCSD analytical gradients, 1531–1534
- PCM-EOM-CCSD analytical gradients, 1538–1541
- PE-CC2, 778
- PE-CCSD, 778, 786
- PE-DFT, 777, 786
- PE-HF, 777
- PE-MCSCF, 777
- PE-MP2, 777
- PE-SOPPA, 777
- Penetration, 311
- Penrose–Hameroff orchestrated objective reduction (OOR) model, 1581
- Penta-3,5-dieneiminium cation (PSB3), 1704–1706
- Pentagonal dodecahedron, 1162
- Peptide, 28
- Perfluoroalkylation of fullerenes, 1075
- PERI-CC2, 778
- Periodic boundary conditions, 765, 771, 1298–1300
- Permutational symmetry, 92
- Persistence bioaccumulation toxicity (PBT), 2253–2257
- Persistence in environment, 2239–2240
- Persistent organic pollutants (POPs), 2239
- Persistent organic pollutants (POPs) data set
- k*-means method, 2114
 - clustering, 2105, 2111
 - PCA, 2121
- Perutz model, 1562
- Pharmacokinetic (PBPK) models, 2288–2295
- Phenathrene-based tylophorine derivatives (PBTs), 2321
- Phenolic compounds, 2348–2349
- Phosphorescence, 573–574, 949–951
- Photoactive semiconducting oxides
- BiVO₄, 993–994
 - cluster approach, nanocrystals and nanostructures, 1005–1006
 - computational parameters, 1006–1008
 - crystalline structures, modeling of, 998–1001
 - DFT methodology and functionals, 1001–1005
 - NiTiO₃, 1006–1008
 - TiO₂, 987–993
- Photoactive yellow protein (PYP), 1930–1931
- Photochemical funnel, 1946
- Photochemical reaction path, 1946

- Photodenitrogenation bicyclic azoalkane, 1962–1965
- Photodynamic therapy (PDT), 712
- Photoelectron spectroscopy (PES), 1187
- Photoelectronic excitation, 1699
- Photoexcitation of proteins, 1721–1722
- Photofragmentation
conical intersection, 1962–1965
DBH, 1964
- Photoinduced reactions in bimolecular systems, 707–718
- Photoisomerization retinal in different environments, 1983–1986
- Photonic dispersion QD dimer systems, 1276
- Photophysical and photochemical process, 642
- Photostability, nucleic acid, 1699
- Physicochemistry-based descriptors, 2028
- Physiologically based pharmacokinetic (PBPK) model, 2289
- Piezoelectricity, 1450–1454
- π orbital axis vector (POAV)
analysis, 1039, 1215–1218
pyramidalization angle, 1041
- Plane wave basis set, 365–366
- Plane wave-based GGA-DFT method, 1480
- Plasma protein binding (PPB), 2280
- PM3, 2058
- PM6, 2058
- Point groups and transformations, 97–111
- Polarizability, 1125–1126, 1449
distributed, 774, 782–784
interaction-induced, 756
interacting atomic, 774
intrinsic, 758
localized, 778, 785
many-body corrected (MBP), 757
solute intrinsic, 757
net, 757
transferable atomic, 775
- Polarizability per molecule (PPM), 756
- Polarizable continuum model (PCM), 679, 760, 761, 786, 932, 1518, 1522–1531, 1534–1538, 1887, 1905, 1982, 2270
- Polarizable embedding (PE), 747, 765, 777–779, 786
- Polarization, 29
catastrophe, 775
electric, 744, 746, 747
electronic, 770
energy, 299
expansion, 298
orientational, 770
- Polyfluorinated chemicals (PFCs), 2252
- Polymer and critical phenomena, 404–405
- Polypeptide, 34
- Pople–Nesbet equations, 656
- Pople-type basis sets, 628–629
- 6-12 Potential, 26
- Potential energy, 27, 30
- Potential energy curves (PECs), 708
- Potential energy hypersurfaces (PEHs), 641, 642
- Potential energy surface (PES), 5, 27, 1631–1633, 1748, 1769, 1921
of cluster, 1461
- Prehydrated electrons, 1898
- Principal component analysis (PCA), 2222, 2235
applications, 2118
correlation-covariance matrix, 2119
covariance, 2119
description, 2118
eigenvalues and explained variance, 2121
POPs, 2121
- Prioritisation, 2231–2257
- Prolateness parameter, 1467
- Propagator tensor
dipole-charge, 747
dipole-dipole, 747, 748, 772
- Protein, 22
folding, 33, 1643
hydration, 1609–1610
interactions protein-protein, 1613–1617
threading, 1600
- Protein crystallography
data collection, 1591
model building, 1592
refinement, 1592
sample preparation, 1591
structure determination, 1591
validation, 1592
- Protein dynamics
dynamical structures, 1608
and molecular dynamics, 1629
protein-ligand interactions, 1608
time-scales, 1605
- Protein structure(s)
electrostatics, 1602–1604
molecular graphics, 1600–1601
prediction, 2185, 2187–2189
solvent-accessible Surface, 1604–1605
- Proton transfer, 1847
- Protonated formamide, 1725
- Protonated Schiff base (PSB) models, 1703
- Protonated Schiff base of 11-cis-retinal (PSB11), 1947, 1956

- Protonation, 1710
vs. cation coordination, 1830
energy, 2049
Gibbs energy correction, 1855
vs. metalation, 1840
preferred sites, 1830
- PSB4 model, 1703, 1705–1708
- Pseudopotentials, 366–368
- PSI3, 872–873
- Psilocybin, 2171
- Psoralen, 671–676
- Psoralen +O₂, 712–718
- Pt clusters, 1472
- PTDE coupled schemes, 1526–1528
- PTE coupled schemes, 1526–1528
- PTED, 1520
- PUVA mechanism, 714
- Pyramidalization angle fullerenes,
1215–1218
single-walled carbon nanotubes,
1223
- Pyrazoline–dichloromethane model
pair, 1968
- Pyrimidine rings, 1712
- Pyroole, 1715–1719
- Q**
- QCFF/PI level of theory, 1037, 1038
- QD dimer systems
lossless dielectric constant, 1282–1283
photonic dispersion, 1276
- QM technique
ab initio, 1810
calculations, 1810
- Quantum mechanical/molecular mechanical
(QM/MM)
models, 733–734, 1955
method, 973
methodology, 1954–1956
technique, 2058
- QSAR. *See* Quantitative structure-activity
relationship (QSAR)
- QSARINS, 2233
- QSARINS-Chem, 2243
- Quadratic response, 1877, 1879, 1880
function, 281–283
- Quadrupolar solvents 1,4-dioxane, 788
- Quadrupole moment, 516–518, 2049
- Qualitative analysis of adiabatic
approximations, 184–188
- Quantitative nanostructure-activity
relationships (QNARs), 2206,
2208, 2211
- Quantitative structure-activity relationship
(QSAR), 2018, 2033, 2042, 2066,
2233–2234, 2267, 2268
antimicrobial compounds, 2346–2349
description, 2134, 2343
high throughput screening (HTS), 2344
stages, 2344
3D, 2134
uses, 2134
- Quantitative structure-activity relationship
(QSAR) modeling
anticancer agents, 2321–2322
anticonvulsant compounds, 2320
applicability domain, 2315–2317
aquatic toxicity, 2329–2330
category models, 2314–2315
chemical toxicity and safety, 2323–2324
classification, 2314
combinatorial, 2311
consensus model, 2331
continuous, 2313
data preparation, 2305
description, 2304
external validation, 2318–2319
geranylgeranyltransferase I inhibitors,
2322–2323
in vitro and *in vivo* relationships,
2325–2327
methods, 2310–2313
research strategies, 2332–2333
steps in, 2305
statistical characteristics, model
performance, 2330
workflow, 2310
- Quantitative structure in vitro-in vivo
relationship (QSIIR) modeling
description, 2327
rodent acute toxicity, 2327–2329
rodent carcinogenicity, 2327
- Quantum chemical methods, 1944
- Quantum chemical response theory, 270
- Quantum chemical standard state, 623
- Quantum chemistry, 4
- Quantum cluster theory and solvation,
1518–1552
- Quantum electrodynamics (QED) effects, 211
- Quantum mechanical (QM) approach, 1954
- Quantum mechanics, 43, 44, 48
- Quantum Monte Carlo, 600
- Quantum theory of atoms in molecules
(QTAIM)
atomic charges, 1055
electron density distribution, 1053
- Quantum topology, 2056

- Quartet-doublet (D) transition, 1568
Quasi-relativistic-energy-averaged Stuttgart pseudopotentials, 1836
- R**
- Radial breathing mode (RBM), 1199
Radial distance, 1370, 1381
Radiosensitizers, 1896
Radiotherapy, 1896
Ramachandran map, 27
Ramachandran plot, 34
Raman, 1444
Raman active modes, 995
Raman spectra, 1854
Random forest methods, 2312
Ranking indexes, 2234–2236
Rational function, 1446
Rb-Cs clusters, 1366
RDKit, 2204
Reaction dynamics (RD), 706
Reaction field, 757, 772, 782
 factor, 749
Reaction vectors approach, 2223–2225
Reciprocal-lattice vectors, 1492
Reciprocal matrices, 2074
Reciprocal space, 1263
Recursive partitioning (RP) methods, 2312
Reference interaction site model (RISM), 734–735, 969
Refractive index, 776, 788
Regression analysis, 1073
Rehybridization, 1039
Replica-exchange molecular dynamics (REMD), 1679
Repulsive intrinsic stacking energy terms, 1818
Rescale thermostats, 374
Resonance, 1416
RESP. *See* Restrained electrostatic potential (RESP)
Response functions theory for molecular solutes, 1541–1550
Response theory, 510–515
 electric multipole moments, 272
 electronic excited states, 275
 magnetic multipole moments, 272
 matrix elements, 274
 software packages, 274
 standard ground state calculation, 273
Restrained electrostatic potential (RESP), 51
Restricted open-shell Kohn–Sham (ROKS), 1923
 method, 1703
Retinylidene, 1703
Retinal proteins, spectral tuning, 1974–1979
Retinal protonated Schiff base (RPSB), 1702
Rhodopsin, 1931–1932
Rhodopsin chromophore model, photoisomerization, 1956–1962
Rigid body docking, 1611
Rigorous local field method (RLF), 746, 765, 771, 786, 789
Ring-opening, 1711
Ring-puckering, 1711
Ritz method, 133
RNA, 1804
Rodent toxicity, 2252
Rotational g Tensor, 539–541
Rule of thumb, 2108
Ruthenium complexes, 1843
- S**
- SAPT2008 program, 327
Sawtooth-like nanoribbon, 1487
Sawtooth potential, 1418
Scalar potential, 1420
 in Hamiltonian, 504
Scaled quantum mechanical force fields, 478–479
Schiff base cross-link product, 1909
Schlegel diagram, C_{70} , 1079
Schrödinger equation, 127, 134, 142, 643, 1263, 1631, 1632
Screening, 2232
Second dipole hyperpolarizability, 525–527
Second-order induction, 313
Segmented all-electron relativistically contracted (SARC) Gaussian basis set, 919
Self-consistent field (SCF), 151, 164
Self-consistent reaction field (SCRf), 760, 2270
Self-organizing maps (SOMs)
 description, 2124
 quality assessment, 2129–2130
 structure determination, 2124–2127
 training, 2127–2128
 visualization, 2128–2129
Self-purification of nanomaterials, 1484
Semantic web, 2173–2175
II–VI Semiconductors, 1396
III–V Semiconductors, 1396
Semi-empirical PM6-DH2 method, 1860
Semiconducting SWNTs, 1309

- Semiconductor clusters
 (hyper) polarizabilities, 1123–1125
 III-V and II-VI, 1108–1112
 II-VI, 1112–1114
 structural properties, 1099–1114
- Semiempirical PM3 method, 1833
- Semiempirical terms, 23
- Serial replica exchange molecular dynamics (SREMD), 1681
- Shape analysis, 1369
- Shape resonances, 1749, 1752
- Si₃₆, 1105–1108
- Si₆, 1105–1108
- Si clusters, 1399
- Si-Ge clusters, 1399
- Si-Ge interface, 1483
- Silhouette index, 2111
- Silicon clusters, 1103–1108
- Silver-mediated cytosine, 1853
- Similarity function, 1371, 1403
- Similarity metrics, 2099
- Simplification, 36
- Simulated annealing method, 1463
- Single-configuration methods, 650–651
- Single linkage method, 2102
- Single-members clusters, 2103
- Single-particle model, 1434
- Single point energy calculations, 1834
- Single-strand breaks (SSBs), 1745, 1746, 1767, 1769, 1777, 1901
- Single-walled carbon nanotubes, 1300–1302, 1308–1311
 armchair, zig-zag and chiral, 1197
 computational studies, 1221–1225
 covalent functionalization, 1225–1231
 defect formation, 1200
 DFT methods, 1203
 electron diffraction, 1199
 heptagon defects, 1200
 hydrogen chemisorption, 1227
 Raman and electronic spectroscopy, 1199
 single- and di-vacancy defects, 1236–1242
 Stone–Wales defects, 1201, 1233
 topological ring defects, 1235–1236
- Singly occupied molecular orbital (SOMO), 1760
- Slater determinants, 112
- Slater rules, 156
- SMILES, 2069
- Sodium, 1383
- Sodium clusters, 827–832
- Solid-state k-p theory, 1261
- Solubility (logS), 2273–2274
- Solute properties, 747, 749, 762, 767, 771, 772, 775
- Solute-solvent interactions, 1519
- Solvation, 53
 effects, 1835
 free energy, 760, 2268
- Solvation energy
 electric properties dispersion term, 762
 electrostatic term, 762
 Pauli repulsion term, 762
- Solvent accessible surface area (SASA), 1605
 model, 2269
- Solvent-phase pairing energies, 1860
- Sphere exclusion protocol, 2309
- Spherical cavity approximation, 750, 764
- Spin and point group symmetry, 111–112
- Spin-dipole (SD) operator, 509
- Spin inversion, 1562
- Spin neuroscience, 1581–1582
- Spin-orbit coupling (SOC), 929, 1504
 constant, 572–573
 definition, 1564
 in O₂ heme interaction, 1574–1578
 T→S transition, 1571
 on UV/Vis properties, 948–949
 weak, 1565
- Spin-orbit interaction, 1345
- Spin-polarized DFT method, 1475
- Spin-polarized first-principles, 1500
- Spin variables, 125, 126
- Spin-valence concept, 1561
- Spin-Zeeman operator, 508
- Splitting scheme, 751, 756–759
- Spurious solutions in anions, 689–690
- Stability function, 1368, 1377, 1390
- Stabilization energies
 adenine and guanine, 1841
 and binding energies, 1845
 formula, 1830
 vs. Gibb's free energies, 1839
 metal-base complexes, 1831
 Pt-cross-linked structures, 1858
- Stacking, 596–597
- Standardization, 2069, 2098
- Static response functions, 283–285
- Statistical mechanics, 964
- Steered molecular dynamics (SMD), 1649
- Steric effects, 2027
- Steric strain, 1040, 1041
- Stern–Gerlach experiment, 1469
- String, 1463
- Structural keys, 2086–2087
- Structural outliers, 2309
- Submolecule approach, 772–773

- Substances of very high concern (SVHC), 2253
- Sum-of-molecular fragments, 780
- Sum of squared errors (SSE), 2115
- Sum-over-states expression (SOS), 1877
- Superdelocalizability, 2045
- Supermolecular approach (SM), 751, 757, 759, 765, 767, 786
- Supermolecular method, 298, 301
- Supermolecule model, 728
- Superposition approximation, 1376
- Supervised machine learning, 2096
- Support vector machines (SVMs), 2312
- Surface(s), 1425
- Surface hopping, 1927–1929
- Susceptibility, 744–746, 749, 772, 776, 783, 786
- cascading, 749
- Sutton–Chen potential, 1365
- Symmetry-adapted perturbation theory (SAPT), 303, 752, 1813
- Systematic nomenclature scheme, 2009
- T**
- Ta clusters, 1471
- Tamm–Dancoff approach, 663
- Target functions
- category QSAR, 2314–2315
- classification QSAR model, 2314
- continuous QSAR, 2313–2314
- Tauc plot, 997
- Tautomeric equilibrium, platinum complex, 1841–1843
- TDDFT, 1952, 1954
- TDDFT/FDE, 782
- Temporary negative ion (TNI), 686
- Tetraikadecahedron (T-Cage), 1165–1166
- Theoretical descriptors, 2066
- Thermochemical calculations, 623–626
- accuracy, 624–626
- case study, 626
- chemical reactions, 623
- Thermochemistry of species with F–O bonds, 632–634
- Thermodynamic
- cycle, 2272
- limit, 1421
- properties, 1373–1376
- stability, 1042, 1900
- Thermostats, 373–377
- THG, 776
- Thioguanine base, 1837
- Thole damping, 774
- Thomas–Fermi–Dirac (TFD) approach, 2051
- Thomas–Reiche–Kuhn (TRK) sum rule, 532
- Thymine, 693–698, 1720
- Ti clusters, 1366
- Tight-binding Hamiltonian, 1493
- Tight-binding molecular dynamics (TBMD), 1344
- Time-dependent density functional theory (TD-DFT), 928–957
- adequate XCF, 931–932
- approach, 664
- atomic basis set selection, 930–931
- complex environments and biomolecules, 953–956
- electronic transition determination, 933–935
- emission and vibronic effects, 939–946
- experiments comparison, 935–936
- performance, 951–953
- solvent effects, 932–933
- vertical absorption, 930–939
- Time-dependent Hartree–Fock (TDHF) theory, 271, 532
- Time-dependent Schroedinger equation (TDSE), 1701
- Tissue distribution
- of CYP enzymes, 2283
- of phase II enzymes, 2284
- Tissue physiological model, 2291
- Titanium dioxide (TiO₂)
- electronic and optical features of, 988
- mesoporous TiO₂, synthesis of, 989–993
- morphology and structural organization, 987
- nanostructures, 1015–1021
- photoactivity of, 987–988
- structural features of, 1006–1008
- Topochemical indices, 2076
- Topographic error, 2130
- Topological distance matrix, 2072
- Topological indices (TIs), 2071–2082, 2345
- Topological representation, 2068
- Topological ring defects, 1235–1236
- Topostructural indices, 2076
- Torsion, 1326
- Torsional interactions, 361–362
- Torsion energy, 25
- Total energy, 28
- Total interaction energy, 300
- Toxicity, 2245–2247
- Traditional MD, 1918
- Transferability, 25

- Transient negative ion (TNI) formation
 molecular orbital approach, 1748
- Transition dipole moment (TDM), 644, 645
- Transition metal (TM), 1348
 complexes photophysics of, 946–953
- Transition moment, 530
- Transition state searches, 432–433, 451–458
 potential scans in, 458–460
- Transition temperature, 1376
- Translationally invariant Hamiltonian, 76
- Translational motion separation of, 74
- Transporter, 2285, 2286
- Transposition, 93
- Transthyrin (TTR), 1649
- Tropospheric oxidants, 2240
- Turbomole, 778
- Two-dimensional graphene sheet, 1303
- Two-level model, 1878–1879
- Two photon absorption (TPA), 533–535, 1876–1889
 applications, 1881–1889
 computational implementation of, 1879–1880
 theory, 1877–1880
 two-level model, 1878–1879
- Two-photon circular dichroism, 549–550
- Two-photon transition matrix, 1877
- U**
- Ultrafast photodeactivation mechanism, 1981
- U-matrix, 2128
- Umbrella integration (UI) method, 1862
- Undersampling, 2315
- Unified distance matrix, 2128
- United atom, 36
- Unnormalized elimination of the small component (UESC), 897–898
- Unoccupied molecular orbital (UMO), 1749, 1754
- UNRES, 51
- Unsupervised machine learning clustering, 2099
 data preparation, 2097–2098
 PCA, 2096
 self-organizing maps, 2124
 uses, 2096
- UV-Filters, 2249
- V**
- Vacancy defects, 1194
- Valence–Rydberg mixing, 684–685
- Vanderbilt ultrasoft pseudopotentials, 367–368
- van der Waals, 26, 2055
 atom radii, 34
 bonding, 300–301
 interaction, 26, 356–359, 598
 radii, 27
- van der Waals density functional (vdW-DF), 605
- Variational principle, 136
- Variational–perturbational scheme, 752
- VB theory, 653
- Vector potential, 1420, 1444
 in Hamiltonian, 506–510
 approach, 1419, 1443–1444
- Velocity Verlet algorithm, 346–347, 411–412
- Verlet algorithm, 345, 411
- Vernalis Research, 2204
- Vertical detachment energy (VDE), 1752, 1754, 1756, 1762, 1764, 1899
- Vertical ionization potential, 2123
- Vibrational circular dichroism (VCD) spectra, 1863
- Vibrational feshbach resonances (VFRs), 1750
- Vibrational heat capacity, 1389
- Vienna ab initio simulation package (VASP), 1204
- Virtual orbital energies (VOEs), 1754
- Virtual orbitals, 165
- Visual molecular dynamics (VMD), 1640
- Volatile organic compounds (VOCs), 2240–2241
- VolSurf descriptor, 2308
- Volume of distribution (VD), 2280–2281
- W**
- Wannier functions, 1435
- Wannier–Mott exciton, 1272
- Ward’s method, 2107
- Water clusters
 “all-surface” to “internally solvated”, 1155–1157
 electronic structure calculations, 1156–1157
 intermolecular interactions, 1141–1146
 2–10 molecules, 1146–1152
 pentagonal dodecahedron, 1162–1165
 $11 \leq n \leq 16$ range, 1152–1156
 spectroscopic signature, 1157
 structure, 1157
 tetraikaidecahedron (T-Cage), 1165–1166
 vibrational spectra, 1160
- Water liquid
 dc-Pockels effect, 786
 electric properties, 786

- Water model, 43
Water molecule, 43
Water radiolysis, 1897
Wave function
 coupled cluster approximation, 165–170
 Hartree–Fock method, 147–152, 159
 Møller–Plesset perturbation theory, 152–159
 one-electron approximation, 142–147
 perturbation calculus, 137–147
 quantum mechanics, 125–133
 variational principle, 136–137
Wave function theory (WFT), 226
Weak intermolecular interactions
 aromatic stacking, 596–597
 basis sets, 607–608
 benchmark set, 609–610
 counterpoise correction, 608–609
 coupled cluster, 599–600
 density functional theory, 602–603
 deoxyribonucleic acid, 595
 empirical performance considerations, 610–613
 Hartree–Fock (HF) theory, 602
 hydrogen-bond, 596
 π interaction, 596, 597
 interaction energy, 598
 molecular mechanics, 606
 Møller–Plesset perturbation theory, 600–601
 quantum Monte Carlo, 600
 semiempirical methods, 605–606
Web Ontology Language (OWL), 2165–2167, 2175
Weighted histogram analysis (WHAM)
 method, 1862
WHIM descriptors, 2084
Wide angle X-ray scattering (WAXS), 989
Wigner–Witmer rules, 678
Williams–Stone–Misquitta (WSM)
 distribution, 318–319
Worm like chain (WLC) model, 399

X
X-ray crystallographic studies, 1050
X-ray diffraction, 989
X ray-like, 1605

Y
Y randomization test, 2317
Yellow fever (YF) inhibitor, 2221–2223
Yellow fever flavivirus (YFV), 2221
Young’s modulus, 1476

Z
Z-cutoff value, 2317
Zener tunneling, 1417
Zeolites structure, 833–842
Zero temperature (ZT), 1344
Zero-dimensional nanomaterials, 1461
Zero-field splitting (ZFS), 577
Zero-point vibrational energy (ZVE), 644
Zeroth-order regular approximation (ZORA), 917
Zigzag graphene ribbon model, 1486
ZINC database, 2016
Zinc-group metal cations, 1836–1838
Zincblende, 1262
ZnO nanobelts, 1497
Z-transformation, 2098
ZT-TBMD method, 1347